Convergence of best response dynamics in extensive-form games

Zibo Xu[1]

Department of Economics, Stockholm School of Economics

20th June 2013

Abstract

We prove that, in all finite generic extensive-form games of perfect information, a continuous-time best response dynamic always converges to a Nash equilibrium component. We show the robustness of convergence by an approximate best response dynamic: whatever the initial state and an allowed approximate best response dynamic, the state is close to the set of Nash equilibria most of the time. In a perfect-information game where each player can only move at one node, we prove that all interior approximate best response dynamics converge to the backward induction equilibrium, which is hence the socially stable strategy in the game.

JEL classification: C73, D83.

Keywords: Convergence to Nash equilibrium, games in extensive form, games of perfect information, Nash equilibrium components, best response dynamics, fictitious play, socially stable strategy.

1 Introduction

Uncoupled dynamics require the adjustment of a player's strategy to be independent of the payoff functions of the other players, but it may depend on the other players' strategies, as well as on the payoff function of the player herself. The rules of behaviour in such situations are called *adaptive heuristic* in Hart (2005), as they are simple, natural and (possibly naively) payoff-improving. Examples are best reply dynamics, better reply dynamics and fictitious plays.

[1]The author is grateful to Sergiu Hart and Jorgen Weibull for many suggestions and discussions. The author also wishes to thank Carlos Alos-Ferrer, Itai Arieli, Larry Samuelson, Bill Sandholm, Tomas Sjostrom and Eyal Winter for their comments. The author would like to acknowledge financial support from the Knut and Alice Wallenberg Foundation.

Uncoupled dynamics are often applied to justify solution concepts in games. In doing so, one of the questions is which dynamic of process leads to which equilibrium in what kind of games. Hart and Mas-Colell (2005) have proved a general negative result that there exists no uncoupled dynamics which guarantee Nash convergence in all games. For special classes of games, some uncoupled dynamics can guarantee the converge to the set of Nash equilibria. See Hofbauer and Sandholm (2002) and Hart and Mas-Colell (2003) for the study on zero-sum games, potential games, dominance-solvable games, etc.

We study best response dynamics in this paper. In a game dynamic process, when populations are large and individuals have bounded reasoning ability and limited information of the whole dynamic process, they may simply refer to best replies according to their current states. On one hand, populations are large and a strategy deviation from a single player have little influence on the aggregate behaviour of the whole population at any time. On the other hand, because of the large populations and the complicated dynamic process, an individual may spare her thoughts on the feedback impact of her current strategy to her own future payoff. So each individual may prefer her current optimal strategy to other strategic behaviour. A best response dynamic is often viewed as a basic and natural process to study the long-run behaviour in a repeated game. We prove the following general result for extensive-form games.

Main result 1: *A continuous-time best response dynamic converges to a Nash equilibrium component in all finite generic extensive-form games of perfect information.*

(See Section A.2 for the generic assumption. Recall that a Nash equilibrium component is a maximal connected set of Nash equilibria.) We adopt the standard continuous-time best response dynamic formulated by a constant revision rate and myopic optimization, as in Gilboa and Matsui (1991), Matsui (1989) and Hofbauer (1995). It can be viewed as a continuous-time analog to a fictitious play. (See the start of Section 3.)

Such continuous best response dynamics have been extensively analysed in various classes of games in strategic (or normal) form. (See books Cressman (2003) and Sandholm (2010).) For extensive-form games of perfect information, Cressman (2003) gives an interesting example of a centipede game of length 4 on page 259 and shows that the subgame perfect Nash equilibrium component is not the maximal attractor (and hence not asymptotically stable) for a best response dynamic. Recall that a *maximal attractor* is an invariant set both forward and backward in time for the best response dynamic. Cressman shows a trajectory initially in the backward induction equilibrium component leads away from it for a period of time. However, we can observe that for the initial state given by Cressman, the distribution of local behaviour strategy at the last decision node in the example game is entirely on the dominated strategy, which will be corrected in the long run if the last node is reached. So, in either case of the last node ever reached or not, the dynamic should behave in the long run as in a centipede game of length 3, and the trajectory converges to the backward induction equilibrium component.

'Convergence of best response dynamics in extensive-form games of perfect information is an open problem,' Cressman (2003) writes, 'even in the case where there is a unique Nash equilibrium component.' We dispense with this condition, and prove the Main Result 1 above in any finite generic extensive-form game of perfect information.

Cressman and Schlag (1998) show in a finite generic extensive-form game of perfect information that a replicator dynamic converges to a Nash equilibrium. However, this result can only be applied to an interior dynamic. For any replicator dynamic with an initial state in a pure strategy profile, the trajectory will stay at the initial state forever. For the best response dynamic, we instead prove the convergence in the speed of exponential decay without such constraint.

Unlike a replicator dynamic, a best response dynamic is not a regular selection dynamic, and the solution trajectory may not be unique. To be more specific, the strategy adjustment rate may not be continuous, and there can be multiple best-reply strategies for a player at some time. When the state is in the basin of attraction to a Nash equilibrium component, the projection distribution of some player's strategies may keep unchanged in a subgame off the equilibrium path. If that distribution generates multiple best replies from other players, the strategy distribution of those players may 'drift' freely. (See the game in Figure 2 for an example.) Hence, we cannot guarantee the convergence of an (interior) dynamic to any particular Nash equilibrium in this case.

In our context, we can only guarantee the convergence of a best response dynamic to a Nash equilibrium component, which is a set of Nash equilibria with the same outcome. In the proof, we define in an extensive-form game a *quasi strategy* as a class of pure strategies such that the outcome is the same regardless of strategies played by other players. A quasi strategy only consists of the sequence of moves which the player may be called upon to play. If a player has already diverted away from the play through a node n at one of its predecessors, then the player knows that node will not be reached by any means. So she does not need to consider her moves after n. From a dynamic view, only the moves included in a quasi strategy matter for the movement of the trajectory. For perfect-information game without chance nodes, a quasi strategy is also a strategy in the semi-reduced normal-form game. (See Lemma A.1.)

To show that the convergence of a best response dynamic in Main Result 1 is robust, we consider an *approximate best response dynamic* adapted from the ϵ-accessible path defined in Gilboa and Matsui (1991). This is a weaker version of a best response dynamic, in the sense that players have a limitation on the ability of recognising the current state. So their change in the mixed strategy may not be directed towards their best response to the current state, but rather the best response to a different strategy profile very close to the current state. For this more general dynamic, it is not true that an approximate best response dynamic always converges to the set of Nash equilibria. (See the game in Figure 3 for an example.) However, we can show that the trajectory comes close to the set of Nash equilibria a high proportion of the time.

Main result 2: *An approximate best response dynamic converges in a weaker sense to the set of Nash equilibria in all finite generic extensive-form games of perfect information: most of the time the state is close to the set of Nash equilibria.*

We shall formalise the definition of the approximate best response dynamic in Section 4. This notion of weak convergence is proposed in a stochastic model in Benaim and Weibull (2003): the time fractions of convergence is called empirical visitation rate there. Young (2009) has also applied this notion to show that the behaviour in the so called interactive trial and error learning comes close to Nash equilibrium most of the time. We prove in our context that an approximate best response dynamic converges to the set of Nash equilibria in this weak sense whatever the initial state and the exact change direction in the specified neighbourhood of the current state at any time.

For the technical part, we show that the distribution of plays converges to a single play in a best response dynamic, and we further show that play is an equilibrium path. For an approximate best response dynamic, we apply the induction on the number of decision nodes in our proof. Note that the set of Nash equilibria may change in the induction step. We however show by induction that most of the time the distribution of plays concentrates on a single play, which is an equilibrium path in the game we are studying.

Since the backward induction equilibrium is traditionally regarded as the rational solution in an extensive-form game of perfect information, it is natural to ask whether the (approximate) best response dynamic converges to the backward induction equilibrium? The answer to the general situation is no, as we will see in the game in Figure 3. However, in the case that each player can only move at one node, the convergence of an interior dynamic in Main result 1 is indeed towards the backward induction equilibrium. To see this, note that the mixed strategy of a player is a distribution of all moves directed from the same node, say n. If, in all subgames after n, the projection distribution in the state is close to the backward induction equilibrium in the subgame, then the dominating move at n in the dynamic is the backward induction equilibrium move in the subgame rooted at n. This dominating move at n is the direction of the change for this player from then.

For a general finite generic extensive-form game of perfect information, given any initial state, we can also find an approximate best response dynamic such that the trajectory converges to the backward induction equilibrium. We observe that when the probability of any non-backward-induction move at some node is enough small, we can then assume that the player plays the backward-induction move at that node at that time in an approximate best response dynamic. We then apply the induction to complete the proof.

Gilboa and Matsui (1991) introduce the definitions of a *cyclically stable set* and a *socially stable strategy*. These stability notions incorporate the idea that, if the perturbation tendency of state recognition is observed, then people are likely to follow this behaviour pattern and the dynamic may turn to be a positive feedback process. Some Nash equilibria in an extensive-form game may not be stable (cyclically or socially) in an approximate best response dynamic. We show that, in any finite generic extensive-form game of perfect information, the backward induction equilibrium belongs to the cyclically stable set. If each player can only move at one node, then the backward induction equilibrium is the only socially stable strategy in the game.

2 The Model

We adopt the standard definition of a finite extensive-form game of perfection information. (See Kreps (1982), Hart (1992) and Ritzberger (2002) for reference.) The main technique contribution is the introduction of quasi strategy−a class of strategies with outcome equivalence.

Given a set N of finitely many nodes, we define a partial order binary relation $\prec$ on N that represents *precedence*. We further suppose an initial node n^0 as a predecessor of all other nodes in N. Such $(N, \prec)$ defines a *tree* T, and we call n^0 the *root* of T. We define the *immediate-predecessor* *function* $\psi : N \to N$ such that

$$\psi(n') = \max\{n : n \prec n'\}\ \forall n' \in N \setminus \{n^0\}$$

and $\psi(n^0) = \emptyset$. Let Ψ be the *predecessor function* $\Psi : N \to 2^N$ with

$$\Psi(n') = \{n \in N : n \prec n'\}.$$

We denote ψ^{-1} to be the *immediate-successor function*. Thus, $\psi^{-1}(n) = \{n' \in N : n = \psi(n')\}$ for all n in N. The *successor function* Ψ^{-1} can be similarly deduced. We call a node n a *terminal node* if $\psi^{-1}(n) = \emptyset$, and write $N_t := \{n \in N, \psi^{-1}(n) = \emptyset\}$.

We say that a sequence $\{n_1, ..., n_i\}$ of nodes is a *subplay* in the tree T if $n_{j-1} = \psi(n_j)$ for all $1 < j \leq i$. If $n_1 = n^0$ and $n_i \in N_t$, then it is called a *play*. Denote the set of all plays by H.

We define a k-player extensive-form game of perfect information on the finite tree $(N, \prec)$. Denote $\mathcal{N} = \{\Lambda^0, \Lambda^1, ..., \Lambda^k\}$ as a partition of $N \setminus N_t$, and call it the *assignment of decision nodes*. The members of Λ^0 are called *chance nodes*; for each $i \leq k$, the members of Λ^i are called the *nodes of player i*. Given a node $n \in N$, we put $\lambda(n)$ as the indicator of which player moves on this node. So $\lambda(n) = i$, if $n \in \Lambda^i$. For chance nodes, define $\tau : \psi^{-1}(\Lambda^0) \to [0, 1]$ to be a probability distribution function such that

$$\sum_{n' \in \psi^{-1}(n)} \tau(n') = 1\ \forall n \in \Lambda^0.$$

We define a vector $\mathbf{v} = (v^1, ..., v^k)$ such that each $v^i : H \to \mathbb{R}$ is a Bernoulli function of player i for all $1 \leq i \leq k$. Since there is an one to one correspondence between N_t and H, we may abuse the notation and write $v^i(n) = v^i(h)$ when $n \in h \cap N_t$. We call the quadruple $(T, \mathcal{N}, \tau, \mathbf{v})$ an extensive-form game Γ of perfect information.

For each player i, a (*pure*) *strategy* a^i assigns a successor to each node in Λ^i. So $\psi(a^i(n)) = n$ for all n in Λ^i. Denote the set of pure strategies of player i by A^i, and the set of *pure-strategy profiles* by $A = \prod_{i=1}^k A^i$. We denote the probability distribution of play in game Γ for a pure strategy profile a to be a function $\rho_a : H \to [0, 1]$ with $\sum_{h \in H} \rho_a(h) = 1$. (Note that H is finite.) Given a node $\bar{n}$, we denote $H_{\bar{n}} := \{h \in H : \bar{n} \in h\}$ and say that the node $\bar{n}$ is *connected* (or *reached*) under a pure strategy profile a if

$$\sum_{h \in H_{\bar{n}}} \rho_a(h) > 0. \tag{2.1}$$

When $\Lambda^0 = \emptyset$, given such a pure strategy profile a in A, we can find a play $h = \{n_0, n_1, ..., n_m\}$ such that $\rho_a(h) = 1$, $n_0 = n^0$ and

$$n_{i+1} = a^{\lambda(n_i)}(n_i) \; \forall 0 \leq i < m.$$

(So the last node $n_m \in N_t$.)

The set of *mixed strategies* for player i is defined as

$$X^i := \Delta(A^i) = \{\sigma = \left(\sigma^i\left(a\right)\right)_{a \in A^i} : \sigma^i\left(a\right) \geq 0 \; \forall a \in A^i \text{ and } \sum_{a \in A^i} \sigma^i(a) = 1\}. \tag{2.2}$$

So a mixed strategy x^i is a vector of probabilities assigned to each pure strategy in A^i. The set of *mixed-strategy profiles* is denoted as $X = \prod_{i=1}^k X^i$. We call the induced probability distribution of a mixed-strategy profile x over plays in T as the *outcome* of x. Note that a pure-strategy profile a generates a *payoff vector* $\mathbf{u}(a) = \sum_{h \in H} \rho_a(h)\mathbf{v}(h)$. We can linearly extend it to a mixed-strategy profile x:

$$\mathbf{u}(x) = \sum_{a \in \text{supp}(x)} \left(\prod_{a^i \in a} x^i\left(a^i\right) \right) \mathbf{u}(a). \tag{2.3}$$

A mixed-strategy profile x is a *Nash equilibrium* of the game Γ if

$$u^i(x) \geq u^i(y^i, x^{-i})$$

for every $i \leq k$ and every $y^i \in X^i$, where $x^{-i} := (x^j | 1 \leq j \leq k, j \neq i)$. We denote by NE the set of all Nash equilibria.

A *subtree* rooted at a node n is the truncated tree $(\Psi^{-1}(n) \cup \{n\}, \prec)$. A *subgame* rooted at node n is the corresponding subtree with the projection assignment of decision nodes and the payoff function. We denote this subgame by Γ_n, and denote the set of all nodes in Γ_n by $N(\Gamma_n)$. A Nash equilibrium is a *backward induction equilibrium* (also called *subgame-perfect equilibrium*) if it induces a Nash equilibrium in all subgames. Kuhn proved in Kuhn (1953) that there always exists a pure backward induction equilibrium, constructed from the terminal nodes and going towards the root. In this paper, we consider a *generic* finite k-player game Γ in extensive form with perfect information. Under the generic assumption, the backward induction equilibrium in Γ is unique.

For a mixed strategy profile x, we say that a node $\bar{n}$ is connected under x if there exists a pure strategy profile a with non-zero probability in x such that node $\bar{n}$ is connected under a. A realised play of a Nash equilibrium is also called an equilibrium path.

In an extensive-form game, two different pure strategies for the same player always induce the same probability distributions over plays, if they differ only at disconnected nodes (cf. Proposition 4.1 in Ritzberger (2002)). This observation suggests a lower-dimensional representation of an extensive-form game. We call two pure strategies a^i_1 and a^i_2 for player i *outcome equivalent* and write $a^i_1 \sim a^i_2$ if, with every combination a^{-i} of strategies for the other players, the outcome generated by these two strategies are always the same, i.e.,

$$\rho_{(a^i_1, a^{-i})}(h) = \rho_{(a^i_2, a^{-i})}(h) \ \forall h \in H \ \forall a^{-i} \in A^{-i}. \tag{2.4}$$

Such relationship of outcome equivalence generates for each player i a partition B^i of the set A^i. That means

1. The union of all sets in B^i equals to A^i;
2. Given any b^i in B^i, for any two strategies $a^i_1, a^i_2 \in b^i$, (2.4) holds.

Thus, each b^i is an equivalence class, and we call B^i the set of *pure quasi strategies* of player i, and the set of *pure quasi strategy profiles* is defined as $B := \prod_{i=1}^k B^i$. Given a pure quasi strategy profile $b = (b^1, ..., b^k)$, we can find a pure strategy profile $a = (a^1, a^2, ..., a^k)$ with $a^i \in b^i$ for all $1 \leq i \leq k$, and we define the payoff vector of profile b as $\mathbf{u}(b) := \sum_{h \in H}(\rho_a(h))v(h))$. The set of *mixed quasi strategies* and the payoff vector of a mixed quasi strategy profile can be defined analogously to (2.2) and (2.3), respectively. Nash equilibria can also be defined with quasi strategies. We use the definition of quasi strategy in this paper. When no ambiguity, we may simply refer to a quasi strategy as a strategy, and write the set of mixed quasi strategies of player i as X^i.

For instance, given the game $\bar{\Gamma}$ in Figure 1, for a pure strategy of player I which includes the move α_1 at the root, it must also specify the move she would play at the bottom node. We, however, do not specify it for a quasi strategy, as it is impossible to reach the bottom node in that case. Hence, in this one-player game, there are only three quasi strategies corresponding to α_1, α_2 and α_3, respectively, in our framework. See Appendix A.1 for comparison between the partition generated from outcome equivalence and the standard representation of reduced normal form for an extensive-form game: outcome equivalence is defined on the realisation of outcome, while a reduced-normal-form strategy is concerning the payoff equivalence.

If we would like to emphasize that a notation is with respect to a game G, then we add (G) after the notation, e.g., $N(G)$, $N_t(G)$, etc.

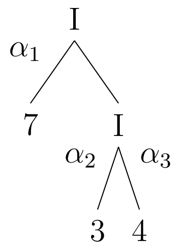

Figure 1: Game $\bar{\Gamma}$

3 Continuous best response dynamics

Given a player i in a k-player extensive-form game Γ of perfect information, for any mixed quasi strategy profile

$$x = (x^i)_{1 \le i \le k} \in \times_{1 \le i \le k} \Delta(B^i),$$

let $BR^i(x)$ be the set of best response in pure quasi strategy for player i, i.e.,

$$BR^i(x) = \arg\max_{b^i \in B^i} u^i(b^i, x^{-i}). \tag{3.1}$$

The sequence $((x_t^i)_{1 \le i \le k})_{t \in \mathbb{N}}$ is a *discrete-time fictitious play* if

$$x_1^i \in \Delta(B^i) \quad \text{and} \quad x_{t+1}^i = \frac{t x_t^i + g_t^i}{t+1}$$

for all $1 \le i \le k$ and all $t \ge 1$ where $g_t^i \in BR^i(x_t)$. Cressman proves that a discrete-time fictitious play converges to NE in a finite generic extensive-form game of perfect information (see Theorem 8.2.9 in Cressman (2003)).

From (3.1), it follows that $x_{t+1} - x_t = (g_t - x_t)/(t+1)$. If we equate this difference to $\dot{x}|_t$, we obtain a nonautonomous system due to the factor $1/(t+1)$. We can ignore this factor and define the (continuous) best response dynamic as a process $(x_t)_{t \ge 0}$ with

$$\dot{x}^i = g^i - x^i \tag{3.2}$$

for some $g^i \in BR^i(x)$. So $\dot{x}^i \in BR^i(x) - x^i$. Up to a rescaling of time, which does not affect the shape of the trajectory, we reach the so called continuous fictitious play:

$$\dot{x}^i \in \frac{BR^i(x) - x^i}{t}.$$

Note the two continuous dynamic processes above are differential inclusions. Since the best response correspondence is upper-semi continuous with closed and convex values, a solution through given initial state exists, though not necessarily unique. (cf. Aubin and Cellina (1984))

Remark: In the literature, it is also common that $BR^i(x)$ in (3.1) is defined to be a subset of X^i. It can be shown that all conclusions in this paper still hold for the more general definition.

We prove the following general convergence result. For a subset $S \subseteq X$ of mixed quasi strategy profiles, denote the ϵ-neighbourhood of S by $S[\epsilon]$. We apply the supremum norm here. That is, $x \in S[\epsilon]$ if and only if there exists a $y \in S$ such that, for all i with $1 \le i \le k$ and all $b^i \in B^i$,

$$|x^i(b^i) - y^i(b^i)| \le \epsilon.$$

Theorem 3.1 *Given a finite generic extensive-form game of perfect information without chance nodes, any best response dynamic $(x_t)_t$ or continuous fictitious play converges to a Nash equilibrium component. That is, for any $\epsilon > 0$, any initial state x_0 and a best response dynamic $(x_t)_{t\ge 0}$, there exists a Nash equilibrium component EC and a time $\tau > 0$ such that for all $t > \tau$*

$$x_t \in EC[\epsilon].$$

Remark 1: A best response dynamic converges to only one Nash equilibrium component. To see this, take a sufficiently small ϵ in the above theorem.

Remark 2: For simplicity, we only consider extensive-form games without chance nodes here. If we drop this condition, the result still holds under the generic assumption.

In the extensive-form game Γ, if a move is included in any best-reply quasi strategy for only finite period of time, then in the solution trajectory, $x_t(b)$ is decreasing to zero for any quasi strategy b including that move. So we only need to focus on a play $(n_1, n_2, ..., n_i)$ such that each move (n_j, n_{j+1}) along the play is part of a best-reply quasi strategy for infinitely long time, and we call such a play *a perpetually realised play.* We can show a perpetually realised play exists and is unique. So the trajectory of the best response dynamic converges to the set of strategy profiles which generate that special play. We then prove by contradiction that the perpetually realised play is a path of a Nash equilibria component.

3.1 Notations and preliminary results

Suppose that there are $\kappa := |N_t|$ terminal nodes in the generic k-player extensive-form game Γ. We enumerate the attached payoffs to each player in order, respectively, as $(v^i_{(1)}, ..., v^i_{(\kappa)})$ for all i with $1 \le i \le k$ such that the superscripts specify the players and $v^i_{(m)} > v^i_{(n)}$ for all $m < n$, $1 \le i \le k$. We can find a positive number δ such that for all i with $1 \le i \le k$ for all $j \le \kappa$

$$(1-\delta)v^i_{(j)} + \delta v^i_{(1)} < (1-\delta)v^i_{(j-1)} + \delta v^i_{(\kappa)}. \tag{3.3}$$

That is, player i's preference over terminal nodes keeps unchanged, even if there is uncertainty of probability δ over which terminal node is finally reached. We fix such a δ for the proof later.

For a pure quasi strategy b^i of player i, we can denote its domain as $\bar{N}_{b^i}$. So $n \in \bar{N}_{b^i}$, if, for all nodes $n' \in \Psi(n) \cap \Lambda^i$, $b^i(n') \in \Psi(n) \cup \{n\}$. That is, if a player i moves towards a node n at every its predecessor node where she could play, then player i should also specify her move at node n. So, a pure quasi strategy b^i of player i in Γ can be represented by a sequence of moves as

$$S(b^i) := \left((n, b^i(n)) : n \in \bar{N}_{b^i}\right). \tag{3.4}$$

Given a player i and a pair of nodes (n_1, n_2) with $n_1 \in \Lambda^i$ and $n_2 \in \psi^{-1}(n_1)$, we denote $B^i(n_1, n_2)$ to be the set including all pure quasi strategies b^i with the property $b^i(n_1) = n_2$. (Note that some strategy of player i may require $n_1 \notin \bar{N}_{b^i}$.) Given any such pair of nodes above, for a mixed quasi strategy $x^i = (x^i(b^i))_{b^i \in B^i}$ (recall $x^i(b^i) \ge 0$ and $\sum_{b^i \in B^i} x^i(b^i) = 1$), define

$$y(n_1, n_2) := \sum_{b^i \in B^i(n_1, n_2)} x^i(b^i) \tag{3.5}$$

and $y(n_1) := \sum_{n_2 \in \psi^{-1}(n_1)} y(n_1, n_2)$. Thus $y(n_1)$ is the probability that player i plays at node n_1, and $y(n_1, n_2)$ is the probability that player i moves from node n_1 to node n_2. Let $x(n_1, n_2) := y(n_1, n_2)/y(n_1)$, when $y(n_1) > 0$. For a node n_1 with $y(n_1) > 0$, we call the distribution $(x(n_1, n_2))_{n_2 \in \psi^{-1}(n_1)}$ the *local behaviour strategy of player i at node* n_1. For a play $h = (n_0, n_1, ..., n_{|h|})$ with $n_{j-1} = \psi(n_j)$ for all $0 < j \le |h|$, we can denote the probability of h followed in game Γ for a quasi strategy profile x to be

$$\rho_x(h) := \begin{cases} \prod_{0<j\le|h|} x(n_{j-1}, n_j) & \text{if } x(n_{j-1}, n_j) \text{ defined } \forall j \le |h| \\ 0 & \text{otherwise.} \end{cases}$$

In a dynamic process, we sometimes write $\rho_t(h)$ as a shorthand for $\rho_{x_t}(h)$.

Given a subtree Γ_n with root n, for a pure quasi strategy b^i of player i, the projection quasi strategy $b^i(\Gamma_n)$ in Γ_n is the subsequence

$$\left((n', b^i(n')) : n' \in \bar{N}_{b^i} \cap N(\Gamma_n)\right)$$

in $S(b^i)$. (Recall that $N(\Gamma_n)$ contains all nodes in the subtree Γ_n.) We can then define a projection mixed quasi strategy as a distribution in $\{b^i(\Gamma_n)\}_{b^i \in B^i}$.

Given a pair of nodes (n, n') with $n = \psi(n')$, we define the Lebesgue integral

$$f(n, n') := \int_0^\infty \mathbb{1}_{(n,n') \in S(g_t^i)} \mathrm{d}t \tag{3.6}$$

where player $i = \lambda(n)$ and g_t is defined in the best response dynamic (3.2). (Recall that $\mathbb{1}$ is the indicator function.)

We give a straightforward lemma below without proof.

Lemma 3.2 *Given a finite generic extensive-form game Γ of perfect information, we suppose that $(x_t)_{t\geq 0}$ is a best response dynamic. We further assume that for a pair of nodes (n, n') with $\psi(n') = n$ where player i plays at node n, $n \notin \bar{N}_{g_t^i}$ or $g_t^i(n) \neq n'$ for all best response strategy $g_t^i \in BR^i(x_t)$ at time t. Then*

$$\dot{x}^i(\Gamma_{n'})|_t = 0, \ \forall t \geq 0, \tag{3.7}$$

where $x^i(\Gamma_{n'})$ is the projection distribution of quasi strategy of player i in the subgame $\Gamma_{n'}$. If the assumption condition on (n, n') is valid for all $t \geq 0$, then

$$x_t(n_1, n_2) = x_0(n_1, n_2)e^{-t} \tag{3.8}$$

for any pair of nodes (n_1, n_2) with $n_1 \in (\Psi^{-1}(n') \cup \{n'\}) \cap \Lambda^i$ and $n_1 = \psi(n_2)$ and for all $t > 0$.

If $n \in \bar{N}_{g_{\tilde{t}}^i}$ and $g_{\tilde{t}}^i(n) = n'$ at any $\tilde{t} \geq 0$, then for all nodes $\bar{n}$ in $\psi^{-1}(n)\backslash\{n'\}$

$$\left(\frac{x_t(n, n')}{x_t(n, \bar{n})}\right)'_{\tilde{t}} = \frac{1}{x_{\tilde{t}}(n, \bar{n})}. \tag{3.9}$$

3.2 Proof of Theorem 3.1

Under a best response dynamic, we would like to find and focus on a play that is followed with probability close to 1 in the long run.

Lemma 3.3 *For a pair of nodes $(\bar{n}, n')$ with $\bar{n} = \psi(n')$ and $f(\bar{n}, n') = +\infty$, there is a set $N^{\bar{n}} \subseteq \psi^{-1}(\bar{n}) \setminus \{n'\}$ with $f(\bar{n}, n) < +\infty$ for all $n \in N^{\bar{n}}$. Then, for every $\epsilon > 0$, there exists a time $t_{\bar{n}}$ such that*

$$\forall t > t_{\bar{n}}, \ \sum_{n \in N^{\bar{n}}} x_t(\bar{n}, n) < \epsilon x_t(\bar{n}, n'). \tag{3.10}$$

Proof. Denote the player who moves at node $\bar{n}$ by player i. From the definition of f in (3.6), it follows that for every $\mu > 0$ there exists a time $t(\mu)$ such that

$$\sum_{n \in N^{\bar{n}}} \int_{t(\mu)}^{\infty} \mathbb{1}_{\{t:(\bar{n},n) \in S(g_t^i)\}} \mathrm{d}t < \mu.$$

The desired result follows from (3.7) and (3.9). □

Lemma 3.4 *There exists a play $h = (n_1, ..., n_{|h|}) \in H$ (with $n_{j-1} = \psi(n_j)$) such that $f(n_i, n_{i+1}) = +\infty$ for all i with $1 \leq i < |h|$. We call such a play a perpetually realised play.*

Proof. This follows from the definition of a quasi strategy (see $S(b^i)$ defined in (3.4)) and the finiteness of the game tree. □

Lemma 3.5 *There is only one perpetually realised play $h_r = (n_1, ..., n_{|h_r|})$ in Γ. Given an $\epsilon > 0$, we denote $\hat{\epsilon} = (\delta\epsilon)/(8|N|)$. For every $\epsilon > 0$ there exists a time $t(\epsilon)$ such that*

$$\rho_t(h_r) > 1 - \hat{\epsilon} \tag{3.11}$$

for all $t > t(\epsilon)$.

Proof. We first prove that there is only one perpetually realised play by contradiction. Assume that there are $j > 1$ perpetually realised plays. Then we can find two perpetually realised plays h^1 and h^2 such that for some integer c

1. $n_i^1 = n_i^2$ for all $1 \leq i \leq c$; we let $n_c := n_c^1 = n_c^2$;
2. $n_{c+1}^1 \neq n_{c+1}^2$;
3. for any other perpetually realised play h_r, $(h_r \setminus \{n_c\}) \cap N(\Gamma_{n_c}) = \emptyset$.

(Recall that $N(\Gamma_{n_c})$ denotes the set of all nodes in the subgame rooted at node n_c.) From Lemma 3.3, we know that there exists a finite time $\bar{t}$ such that

$$\forall t > \bar{t},\ \forall j = \{1, 2\}, \forall i \geq c,\ x_t(n_i^j, n_{i+1}^j) > 1 - \hat{\epsilon}.$$

Denote the player who moves at node n_c by player l. For any $t > \bar{t}$, for a pure quasi strategy b^l in B^l with $b^l(n_i^1) = n_{i+1}^1$ for all $n_i^1 \in \Lambda^l$,

$$N\hat{\epsilon}v_{(\kappa)}^l + (1 - N\hat{\epsilon})v^l(h^1) < u^l((b^l, x_t^{-l})(\Gamma_{n_c})) < N\hat{\epsilon}v_{(1)}^l + (1 - N\hat{\epsilon})v^l(h^1), \tag{3.12}$$

where $u^1(b^l, x_t^{-l}(\Gamma_{n_c}))$ is the expected payoff of player l for strategy b^l at time t conditional on reaching node n_c.

Similarly, for any $t > \bar{t}$, for a pure quasi strategy $\tilde{b}^l$ in B^l with $\tilde{b}^l(n_i^2) = n_{i+1}^2$ for all $n_i^2 \in \Lambda^l$,

$$N\hat{\epsilon}v_{(\kappa)}^l + (1 - N\hat{\epsilon})v^l(h^2) < u^l((\tilde{b}^l, x_t^{-l})(\Gamma_{n_c})) < N\hat{\epsilon}v_{(1)}^l + (1 - N\hat{\epsilon})v^l(h^2). \quad (3.13)$$

Since $N\hat{\epsilon} < \delta$, we infer from (3.12), (3.13) and (3.3) that

$$u^l(b^l, x_t^{-l}(\Gamma_{n_c})) \neq u^l(\tilde{b}^l, x_t^{-l}(\Gamma_{n_c}))$$

for all $t > \bar{t}$. So the definition of best reply strategy g_t^i requires that $f(n_c, n_c^1) = +\infty$ and $f(n_c, n_c^2) = +\infty$ cannot both hold. So there are at most $j - 1$ perpetually realised plays. Contradiction.

If there is only one perpetually realised play, then (3.11) follows straightforwardly. □

Lemma 3.6 *Given any $\epsilon < 1/2^{|N|}$, for the perpetually realised play $h_r = (n_1, ..., n_{|h_r|})$, there are infinitely many times $t > t(\epsilon)$, where $t(\epsilon)$ is defined in Lemma 3.5, such that*

$$\dot{x}(n_j, n_{j+1})|_t \geq 0 \ \forall \ j \text{ with } 1 \leq j < |h_r|. \quad (3.14)$$

Proof. We apply Lemma A.4, which considers a more general definition of μ-dynamic. A best response dynamic is a 0-dynamic. For the formal definition of a μ-dynamic, please refer to Definition 4.1. □

Lemma A.2 and Lemma A.3 in Appendix show that, if a play h is not an equilibrium path, then there exists at least one node n in h such that, in any state where all players follows the play h, a deviation from the node n can give the player who plays at n an extra payoff bounded away from zero.

We show below that, if the distribution of the outcome in any state is enough close to and moving towards a single play, then that play is an equilibrium path.

For the k-player game Γ, let

$$c := \min_{i=1}^{k} \left\{ \min_{l=1}^{\kappa-1} \left\{ v_{(l)}^i - v_{(l+1)}^i \right\} \right\}, \quad (3.15)$$

and

$$\bar{\epsilon} := \min \left\{ c, \frac{c}{\max_{i=1}^k \left(2v_{(1)}^i - 2v_{(\kappa)}^i \right)} \right\}.$$

Lemma 3.7 *For any*

$$\epsilon < \bar{\epsilon}, \tag{3.16}$$

if both (3.11) and (3.14) hold for the perpetually realised play h_r in Γ at some time t, then h_r is an equilibrium path in Γ.

Proof. Given the play $h_r = (n_1, ..., n_{|h_r|})$, we define $X(h_r)$ to be the set of mixed strategy profiles x with $x(n_j, n_{j+1}) = 1$ for all $1 \leq j < |h_r|$. (Recall that $x(n_j, n_{j+1})$ is the probability of moving from n_j to n_{j+1} in the local behaviour strategy at node n_j.)

We consider a time t when both (3.11) and (3.14) hold. We modify x_t to an $\hat{x} \in X(h_r)$ such that

1. $\hat{x}(n_j, n_{j+1}) = 1$ for all $1 \leq j < |h_r|$;
2. $\hat{x}(n', n'') = x_t(n', n'')$ for all (n', n'') with $n' \notin h_r$, $x_t(n', n'')$ defined, and $\lambda(n') \neq \lambda(\psi^{\bar{q}}(n'))$ in which $\bar{q} := \min\{q : \psi^q(n') \in h_r\}$. (Roughly speaking, the last condition above means the player who moves at node n' is not the player who can deviate from h_r towards node n'.)

So the behaviour strategy of each player in $\hat{x}$ is well defined, and $\hat{x}$ is in $X(h_r)$.

Suppose h_r is not an equilibrium path, we pick a deviation node n_j in h_r as defined in Lemma A.2, and denote the player who moves at n_j by player i. We then take a pure quasi strategy $g_{\hat{x}^{-i}}$ of player i defined in (A.3), which is a best response of player i against $\hat{x}^{-i}$ under the constraint of a move at node n_j towards a node different from n_{j+1}. From (3.11) , we may infer

$$u^i(x_t) < \epsilon v^i_{(1)} + (1 - \epsilon) v^i(h_r); \tag{3.17}$$

$$u^i(g_{\hat{x}^{-i}}, x_t^{-1}) > \epsilon v^i_{(\kappa)} + (1 - \epsilon)\hat{u}(\hat{x}^{-i}), \tag{3.18}$$

where $\hat{u}(\hat{x}^{-i})$ is the payoff to player i for $(g_{\hat{x}^{-i}}, \hat{x}^{-i})$. We also deduce from (A.5) in Lemma A.3 and (3.15) that

$$\hat{u}(\hat{x}^{-i}) - v^i(h_r) \geq c. \tag{3.19}$$

From (3.16), it follows

$$\epsilon < \frac{c}{(2v^i_{(1)} - v^i(h) - v^i_{(\kappa)})}.$$

So

$$\epsilon \left(v^i_{(1)} + \hat{u}(\hat{x}^{-i}) - v^i(h_r) - v^i_{(\kappa)} \right) < c. \tag{3.20}$$

From (3.17), (3.18), (3.19) and (3.20), it follows that

$$u^i(x_t) < u^i(g_{\hat{x}^{-i}}, x_t^{-1}).$$

Hence (3.14) does not hold at time t. Contradiction. □

Rigorously speaking, the convergence of the distribution of plays to an equilibrium path is not equal to the convergence of the trajectory to the corresponding equilibrium component. One needs to show that, from some time on, the behaviour-strategy distribution off the equilibrium path either makes the state in the basin of attraction to that equilibrium component, or not far away from the projection distribution of a strategy profile in the equilibrium component.

If (3.14) always holds from some time t, then it follows from (3.7) and (3.8) that both (3.11) and (3.14) hold for h from that t on. Even if (3.14) is sometimes not true, it is true most of the time and the state is always in the ϵ-neighbourhood of the equilibrium component whose path is the perpetually realised play.

Proof of Theorem 3.1: Without loss of generality, we take an $\epsilon < \min\{1/2^{|N|}, \bar{\epsilon}\}$. From Lemma 3.5, Lemma 3.6 and Lemma 3.7, there exists a time $\bar{t} > t(\epsilon)$ such that both (3.11) and (3.14) hold for an equilibrium path h. Denote the equilibrium component with equilibrium path h by EC. So $x_{\bar{t}} \in EC[\hat{\epsilon}]$.

If there exists a time $\tilde{t}$ as the infimum of time $t > \bar{t}$ such that (3.14) does not hold at time t, then there must exist a time $\hat{t}$ as the infimum of time $t > \tilde{t}$ such that (3.14) holds at time t. Recall that for any edge (n_j, n_{j+1}) in h, if $\dot{x}(n_j, n_{j+1}) < 0$, then $\dot{\rho}(h) < 0$. (c.f. Lemma A.4.) In that case, by (3.8) as well as (A.8) and (A.9) in the proof of Lemma A.4 (for the case of $\mu = 0$), we find

$$\dot{\rho}(h)|_t < -(1 - 2|N|\hat{\epsilon}) = -(1 - \epsilon/4) \tag{3.21}$$

for all t with $\tilde{t} < t < \hat{t}$. (Recall that the number of nodes in Γ is $|N|$.) We may also infer from (3.21) and Lemma 3.5 that

$$\left(1 - \frac{\epsilon}{4}\right) e^{-(\hat{t}-\tilde{t})} > 1 - \frac{\epsilon}{4} - \hat{\epsilon}. \tag{3.22}$$

To see this, note that $\rho_t(h) > 1 - \hat{\epsilon}$ for any $t \in [\tilde{t}, \hat{t}]$, and that for the function $y = (1 - \epsilon/4)e^{-x}$, $\dot{y}|_x > -(1 - \epsilon/4)$ for all $x > 0$. We compare the function $\rho_t(h)$ and y and hence reach (3.22).

We then infer from (3.22) and the proposition of the best response dynamic that

$$|x^i_{t_1}(b^i) - x^i_{t_2}(b^i)| < \frac{\epsilon}{4} + \hat{\epsilon} < \frac{\epsilon}{2} \;\forall \tilde{t} < t_1 < t_2 < \bar{t} \;\forall b^i \in B^i \;\forall 1 \le i \le k.$$

So $x_t \in EC[\epsilon]$ for all t with $\tilde{t} \le t \le \hat{t}$. We observe $x_{\hat{t}} \in EC[\hat{\epsilon}]$. Q.E.D.

Comment 1: It is not always true that an interior best response dynamic converges to a single Nash equilibrium. (An interior dynamic requires $y_0(n) > 0$ for all non-terminal nodes n in Γ. See the definition of $y(n)$ just below (3.5).) Consider the game in Figure 2. Suppose at time $t = 0$,

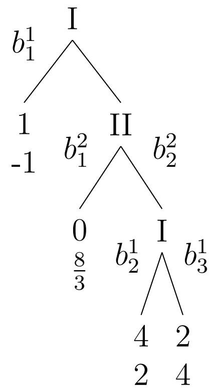

Figure 2: Convergence only to a component

$$x^1(b_1^1) < 1, \quad \frac{x^1(b_2^1)}{x^1(b_2^1) + x^1(b_3^1)} = \frac{2}{3} \quad 0 < x^2(b_1^2) \le \frac{1}{8}.$$

Then $\{b_1^1\} = BR^1(x_0)$ and $\{b_1^2, b_2^2\} = BR^2(x_0)$. We can choose $(g_t^2)_t$ such that the $(x_t)_t$ dynamic has the following properties for all $t \ge 0$.

1. $\{b_1^1\} = BR^1(x_t)$,

2. $$\frac{x_t^1(b_2^1)}{x_t^1(b_2^1) + x_t^1(b_3^1)} = \frac{2}{3}$$

   (this follows from (3.7)),

3. $\{b_1^2, b_2^2\} = BR^2(x_t)$,

4. $x^2(b_2^2) \le 1/8$.

In this way, $(x_t)_t$ converges to the Nash equilibrium component, but x_t^2 may not be fixed in the dynamic process, and hence $(x_t)_t$ may not converge to any particular Nash equilibrium.

Comment 2: The finite time τ in Theorem 3.1 may not be bounded. For instance, see the game in Figure 8.3.3 on page 259 in Cressman (2003), which shows that NE is not the maximal attractor for the best response dynamic. We assume the initial state of a μ-dynamic is $(D, (\frac{4}{5}, 0, \frac{1}{5}))$. Then, we can let $g_t^1 = D$ and $x_t = x_0$ until an arbitrary finite time $\bar{t}$. If $g_{\bar{t}}^1$ turns to AA and $x_{\bar{t}}^1(AA) > 0$, then the trajectory enters $NE[\epsilon]$ in bounded time starting from $\bar{t}$.

Comment 3: We use Lebesgue integral in (3.6) for general dynamic processes. For instance, in the game in Figure 2, we can require in some period $[t_1, t_2]$, $g_t^2 = b_1^2$ when t is a rational number, and $g_t^2 = b_2^2$ when t is an irrational number.

Comment 4: Readers may find that the proof above has similar flavour with the proof in Cressman and Schlag (1998). We show here why we cannot imitate their proof for a best response dynamic process.

1. Their proof is not designed for the convergence to a Nash equilibrium component. For example, in the proof to Lemma 2 in Cressman and Schlag (1998), they claim that 'Since v_1 is a final decision node of Γ', by the definition of e_i^1, $u_1(e_i^1, f_i) \geq u_1(e_i, f_i)$ for any $f_j \in F_2$.' In the game in Figure 2, we can make $1/16 \leq x_t^2(b_2^2) \leq 1/8$, and we know $u^1(b_1^1, b_2^2) = 1 < u^1(b_2^1, b_2^2) = 4$, and $\limsup_t x_t^1(b_2^1) = 0$. Here b_1^1 can be viewed as e_i^1, b_2^2 as f_j and e_1 as b_2^1 in their paper. So their claim is not always valid for a best response dynamic.

2. They focus on the integral $\int_0^\infty x_t^i(b^i)dt$, which leaves out some basic information of best-reply propositions in our context.

   Even if $\int_0^\infty x_t^i(b^i)dt < \infty$, it is possible that $x^i(b^i) > 0$ holds less and less frequently, but $x^i(b^i)$ is bounded from below when it holds. We instead turn to the integral of the period that a strategy is a best response in (3.6).

3. In (3.2), $\dot{x}$ can be discontinuous at some time, which cannot happen in a replicator dynamic process.

4 Approximate best response dynamics

Given a mixed strategy profile x and a positive number ϵ, we abuse the notation and write the ϵ-neighbourhood of x as $x[\epsilon] = \{x\}[\epsilon]$.

Definition 4.1 *Consider a k-player extensive-form game Γ of perfect information with the set of pure quasi strategies B^i for each player i, $1 \leq i \leq k$. We call a continuous-time dynamic process $(x_t)_{t \geq 0}$ a μ-dynamic on Γ if every x_t is a k-dimensional vector $(x_t^i)_{1 \leq i \leq k} \in \times\Delta(B^i)$ and the dynamic satisfies $\dot{x}^i = g^i - x^i$ for some $g^i \in BR^i(x[\mu])$ for all i with $1 \leq i \leq k$.*

We show an example that a μ-dynamic $(x_t)_t$ may not converge to the set of Nash equilibria (denoted as NE), in the following extensive-form game of perfect information.

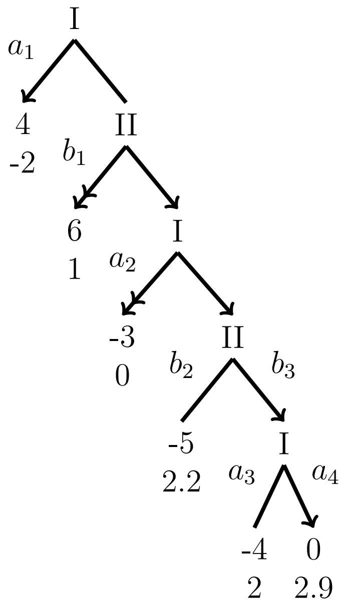

Figure 3: Example in which μ-dynamic may not converge to NE

For ease of exposition, we may also write the set of best replies for player i at time t as $BR^i(x_t^{-i}[\mu])$, since it does not depend on the component x_t^i.

In the game in Figure 3, the backward induction equilibrium is (a_1, b_3) and the alternative pure Nash equilibrium is (a_2, b_1). (The backward induction moves are single arrowed, and the final moves in the alternative equilibrium are double arrowed.) We denote the corresponding two Nash equilibrium components by BC and NC, respectively. If x_t ever approaches to NE, then there exists a time t_0 such that $x_{t_0} \in BC[\mu]$ or $x_{t_0} \in NC[\mu]$. Without loss of generality, we let $\mu \leq 1/100$.

1. $x_{t_0} \in NC[\mu]$: we define a mixed quasi strategy $\tilde{x}^2$ of player 2 with $\tilde{x}^2(b_1) = 1 - \mu$ and $\tilde{x}^2(b_3) = \mu$. Then for any $x \in NC$, $\tilde{x}^2 \in x^2[\mu]$. We assume
$$x_t \in NC[\mu] \ \forall t \geq t_0, \tag{4.1}$$
which we show below in fact not true for some μ-dynamic. Under the assumption, we can let
$$g_t^1 \in BR^1(\tilde{x}^2) \ \forall t \geq t_0$$

in a μ-dynamic and hence

$$g_t^1 = a_4 \ \forall t \geq t_0.$$

So there exists a time $t_1 \geq t_0$ such that $x_{t_1}^1(a_3) + x_{t_1}^1(a_4) > 2/3$. So at some time before t_1, g^2 becomes b_3, and then g_1 can still be set as a_4. Hence, the trajectory is moving away from NC towards BC. We then infer that the assumption (4.1) cannot hold for every μ-dynamic.

2. $x_{t_0} \in BC[\mu]$: we define a mixed quasi strategy $\bar{x}^1$ of player 1 with $\bar{x}^1(a_1) = 1 - \mu$ and $\bar{x}^1(a_3) = \mu$. Then for any $x \in BC$, $\bar{x}^1 \in x^1[\mu]$. We assume

$$x_t \in BC[\mu] \ \forall t \geq t_0, \tag{4.2}$$

which we show below in fact not true for some μ-dynamic. Under the assumption, we can let

$$g_t^2 \in BR^2(\bar{x}^1)$$

and hence $g_t^2 = b_2$ in a μ-dynamic from t_0 until the time $t_1 \geq t_0$ when

$$x_{t_1}^2(b_2) > 3/4. \tag{4.3}$$

We define another mixed quasi strategy $\hat{x}^1$ of player 1 with $\hat{x}^1(a_1) = 1 - \mu$ and $\hat{x}^1(a_2) = \mu$. Then for any $x \in BC[\mu]$, $\hat{x}^1 \in x^1[\mu]$. Under the assumption (4.2), we can then let

$$g_t^2 \in BR^2(\hat{x}^1)$$

and hence $g_t^2 = b^1$ from t_1 until the time $t_2 \geq t_1$ when $x_{t_1}^2(b_1) > 4/5$. When $x_t^2(b^1) > 4/5$, $g_t^1 \neq a_1$. From(4.3) and (3.7), it follows $x_t^2(b_2)/x_t^2(b_3) > 3$ for all t with $t_1 \leq t \leq t_2$. So at some time before t_2, g^1 becomes a_2, and then g_2 can still be set as b_1. (Note that the best reply of player 2 against a_2 is b_1.) Hence, the trajectory is moving away from BC towards NC. We then infer that the assumption (4.2) cannot hold for every μ-dynamic.

From another direction, we can show why the proof of Theorem 3.1 cannot be applied to all approximate best response dynamic processes. We check Lemma 3.5 for the game in Figure 3. In a best response dynamic (0-dynamic), if $f(a^4) = +\infty$ (f defined in (3.6), then

$$\lim_{t \to \infty} \frac{x_t^1(a_3)}{x_t^1(a_4)} = 0.$$

So the best reply of player 2 cannot be b_2 after sufficiently long time, and hence the play with the last move b_2 cannot be a perpetually realised play. Given a μ-dynamic which transits between $NC[\epsilon]$ and $BC[\epsilon]$ infinitely often, we find that both the plays with the last move b_2 and a_4, respectively, are perpetually realised plays. This is because when $x^1(a_4) < \mu$, a μ-perturbation on $x^1(a_3)$ may decide whether b^2 or b^3 is the current best reply of player 2.

In the game in Figure 3, although the μ-dynamic $(x_t)_t$ does not converge to NE, the relative proportion of the time spent outside NE is small in the long run, if μ is small. That means most of the time the state is very close to some Nash equilibrium component, and, if it transits to another component, then the transition time is very small compared with the period when it is close to a component.

For simplicity, we only consider extensive-form games without chance nodes here. If we drop this condition, the result still holds under the generic assumption. Recall that an interior dynamic requires $y_0(n) > 0$ in the initial state for all non-terminal nodes n in Γ.

Theorem 4.2 *Given any finite generic extensive-form game Γ of perfect information without chance nodes, for every $\epsilon > 0$ there exists a number $\mu(\epsilon) > 0$ and $T(\epsilon) > 0$ such that for every interior $\mu(\epsilon)$-dynamic $(x_t)_t$ on Γ with any initial state x_0*

$$\frac{1}{T}\int_0^T \mathbb{1}_{\{t:x_t\in NE[\epsilon]\}}\mathrm{d}t \geq 1-\epsilon$$

for all $T \geq T(\epsilon)$.

From this theorem, which will be proved in Appendix, we have the following main theorem.

Theorem 4.3 *Given any finite generic extensive-form game Γ of perfect information without chance nodes, for every $\epsilon > 0$ there exists a number $\mu(\epsilon) > 0$ such that for every $\mu(\epsilon)$-dynamic $(x_t)_t$ on Γ with any initial state x_0*

$$\lim_{T\to\infty}\left(\frac{1}{T}\int_0^T \mathbb{1}_{\{t:x_t\in NE[\epsilon]\}}\mathrm{d}t\right) \geq 1-\epsilon.$$

Proof. If the μ-dynamic is an interior dynamic, then the desired result follows from Theorem 4.2.

If the μ-dynamic is not an interior dynamic, we denote the set $N_y := \{n \in N_d : \exists t : y_t(n) > 0\}$, where N_d is the set of all non-terminal nodes in Γ. So there exists a finite time $\bar{t}$ such that $y_t(n) > 0$ for all $n \in N_y$ and all $t \geq \bar{t}$. We remove from Γ all non-terminal nodes (and the subgames rooted there) not in N_y, and apply Theorem 4.2 in the resulted game from time $\bar{t}$. □

5 Convergence to the backward induction equilibrium component

Given an extensive-form game of perfect information, we denote the backward induction equilibrium component in the game by BC. Denote the backward induction equilibrium by BIE. We show below that, if each player can only move at one node in the game, then an interior μ-dynamic with enough small μ always converges to BC.

Theorem 5.1 *We suppose that in a finite generic extensive-form game Γ of perfect information, every player can only move at one node. Then, for every $\epsilon > 0$ there exists a number $\mu(\epsilon) > 0$ and $T(\epsilon) > 0$ such that for every interior $\mu(\epsilon)$-dynamic $(x_t)_t$ on Γ with any initial state x_0*

$$x_t \in BIE[\epsilon]$$

for all $t \geq T(\epsilon)$.

Remark 1: This theorem can be applied to the case $\mu = 0$, i.e., a best response dynamic.

Remark 2: Note that the μ-dynamic converges to the (exact) backward induction equilibrium, which is stronger than the convergence to the backward induction equilibrium component.

Given an extensive-form game of perfect information, the convergence of a μ-dynamic with small μ to the backward induction equilibrium is always possible.

Theorem 5.2 *Given a finite generic extensive-form game Γ of perfect information and an $\epsilon > 0$, there exists a number $T(\epsilon) > 0$ such that for any initial state x_0 and any $\mu > 0$, there exists a μ-dynamic $(x_t)_t$ with the property*

$$x_t \in BIE[\epsilon]$$

for all $t > T(\epsilon)$.

Comment 1: One may be tempted to prove the following claim. Given a finite generic extensive-form game Γ of perfect information and an $\epsilon > 0$, there exists a number $T(\epsilon) > 0$ such that for any initial state x_0, there exists a best response dynamic $(x_t)_t$ with the property $x_t \in BC[\epsilon]$ for all $t > T(\epsilon)$. However, one can give a counterexample of the game Γ in Figure 3 to this claim. Consider an initial state $x_0 = (\hat{x}^1, \hat{x}^2)$ with $\hat{x}^2(b_2) > 3/4$,

$\hat{x}^1(a_1) = 99/100$ and $\hat{x}^1(a_2) = 99/10000$. Following the analysis of $\hat{x}^1$ in Step 2 for Γ (the game in Figure 3), we find that the trajectory of the best response dynamic converges to the non-backward-induction equilibrium component.

Comment 2: The convergence property in Theorem 5.2 may not hold for the non-backward-induction equilibrium component. Consider Theorem 5.1 and any extensive-form game of perfect information with more than one pure Nash equilibrium and that each player can only move at one node.

6 Further Comments

1. Noldeke and Samuelson (1992) study a model of evolutionary process in extensive-form games, and they characterise each player by her strategy and conjecture. Here a conjecture is essentially the belief of the player on the local behaviour strategy at all unreached decision nodes. A best response of a player is against the behaviour-strategy distribution at all reached decision nodes and her conjecture at all other decision nodes. At any time, all players can observe the behaviour-strategy distribution at all reached decision nodes.

   If we apply such characterisation in our model, then we only need to consider non interior dynamics. This is because a conjecture is irrelevant in the case of interior dynamics, and once a node is reached at any time t, it will be reached at all finite time after t. For a non interior dynamic, by the similar argument in the proof of Theorem 4.3, we can focus on the set of nodes which are eventually reached. From Lemma A.2 and Lemma 3.7, we may infer that, for this version of best response dynamics, the distribution of plays in a finite generic extensive-form game of perfect information converges to an equilibrium path, and the trajectory converges to the set of self-confirming equilibria.

2. One cannot be overoptimistic about the convergence of best response dynamics. As stated in Hart and Mas-Colell (2005) and Theorem 8.6.1 in Hofbauer and Sigmund (1998), there is no natural or reasonable dynamic process that leads to an equilibrium in every (normal-form) game.

3. For a μ-dynamic in this paper, we assume that a player may not fully recognise the current state, and his best response is against a point in a μ-neighbourhood of the current state. We may also consider another related weak best response dynamic: a player can always recognise the current state, but cannot fully control her behaviour in the sense that

the direction of change can only be guaranteed in a neighbourhood of the best response strategy. So it is a trembling-hand version of the best response dynamic, and we may formulate it as

$$\dot{x}^i \in \left(BR^i\left(x\right)\right)[\mu] - x^i \tag{6.1}$$

for each player i in the game.

Note that the limit of the differential inclusion in (6.1) and the one in Definition 4.1 are the same when μ decreases to zero. We conjecture that Theorem 4.3 is still true for the dynamics described in (6.1). Note that, for this dynamic, Lemma 3.7 cannot be applied directly, and (3.7) may not hold.

4. We constrain ourselves within the processes of continuous-time best response dynamics in this paper. One may also study a discretisation of such continuous dynamics, as called (stochastic) *best-reply dynamics* in Xu (2013). A mixed strategy of a player can be approximated by a large population of individuals where the proportion of individuals playing a pure strategy represents the probability of the pure strategy being played in the mixed strategy. We construct a best-reply dynamic as in Kandori et al. (1993), Kandori and Rob (1995) and Hart (2002): in each period, the revision opportunity arrives independently across individuals and time, with probability bounded from below. Each active individual then myopically replaces her strategy by a currently best reply, when selection occurs; she randomly chooses a strategy, when mutation happens with small probability.

   In such a best-reply dynamic, a state is called *evolutionarily stable* if its long-term relative frequency of occurrence is bounded away from zero as the mutation rate decreases to zero and populations grow to infinity. (See Hart (2002).) The combined results in Hart (2002) and Gorodeisky (2006) show that the backward induction equilibrium is the only evolutionarily stable state in a finite generic extensive-form game of perfect information where each player can only move at one node. (Hart calls such a game a *gene-normal form* game.) We have proved in Theorem 5.1 that the backward induction equilibrium is the only socially stable strategy in this kind of games. The analogy between these two results may be from that the dynamics in (6.1) is an iterative limit, $\lim_{\mu\to 0}\lim_{m\to\infty}$, of a best reply dynamic, where μ is the mutation rate and m is the population size. (Recall that the limit of the differential inclusion in (6.1) and the one in Definition 4.1 are the same.) The stability result of the backward induction equilibrium applies to both deterministic and stochastic dynamics, and both continuous and discrete time models. An appealing analysis approach to general stability notion and results of best response dynamics is a direction for further research.

A Appendix

A.1 Terminology of reduced normal form

There are two other well-known normal form representations of an extensive-form game. (cf. Ritzberger (2002).) A normal-form game is *semi-reduced* (so called *pure-strategy reduced* or *quasi-reduced*) if for all pair (s_1^i, s_2^i) of strategies of player i and for all players $i = 1, ..., k$

$$\mathbf{u}(s_1^i, s^{-i}) = \mathbf{u}(s_2^i, s^{-i}),\ \forall s^{-i} \Rightarrow s_1^i = s_2^i.$$

A normal-form game is *mixed-strategy reduced* (or simply called *reduced*) if for all strategies s^i of player i and for all players $i = 1, ..., k$

$$\mathbf{u}(s^i, s^{-i}) = \sum_{\bar{s}^i \in S^i} \sigma^i(\bar{s}^i)\mathbf{u}(\bar{s}^i, s^{-i}),\ \forall s^{-i} \Rightarrow \sigma^i(s^i) = 1.$$

That is, no s^i is payoff equivalent to any convex combination of the elements in $S^i \setminus \{s^i\}$. (Recall σ^i in (2.2).) Given an extensive-form game Γ, we denote the set of strategies for player i in semi-reduced normal form and mixed-strategy reduced normal form by S^i and M^i, respectively. Recall the set of quasi strategies for player i is B^i. From their definitions, we could let $B^i \supseteq S^i \supseteq M^i$ for any player i in any extensive-form game Γ.

Recall that $\Lambda^0 = \emptyset$ implies no chance node.

Lemma A.1 *Given a generic extensive-form game of perfect information with $\Lambda^0 = \emptyset$, for each player i, $B^i \supseteq S^i$ implies $B^i = S^i$.*

Proof. When $\Lambda^0 = \emptyset$, for every player i and every two different quasi strategies $b_1^i, b_2^i \in B^i$, there exist two different plays h_1, h_2 and one strategy combination b^{-i} of other players such that

$$\rho_{(b_1^i, b^{-i})}(h_1) = \rho_{(b_2^i, b^{-i})}(h_2) = 1,$$

i.e., (b_1^i, b^{-i}) and (b_2^i, b^{-i}) generate different plays h_1 and h_2, respectively, in Γ. (Rigorously, given a quasi strategy profile $b = (b^1, ..., b^k)$ and a pure strategy profile $a = (a^1, ..., a^k)$ such that $a^i \in b^i$ for all i with $1 \leq i \leq k$, we have $\rho_b(h) = \rho_a(h)$ for all play h in H.) Since Γ is generic, it follows that

$$\mathbf{u}(b_1^i, b^{-i}) \neq \mathbf{u}(b_2^i, b^{-i}).$$

It is true for every pair of quasi strategies $b_1^i, b_2^i \in B^i$. Recall that the set of quasi strategies is a partition of pure strategies, and we reach the conclusion $B^i = S^i$. $\square$

Even if $\Lambda^0 = \emptyset$, it is not true that $B^i = M^i$ for all players i in all generic extensive-form game of perfect information. In the one-player game Γ' below, we denote $B^1 = \{b_1, b_2, b_3\}$ where $u^1(b_1) = 3$, $u^1(b_2) = 0$ and $u^1(b_3) = 6$. However, strategy b_1 can be replaced by a mixed strategy $x^1 = \{\sigma^1(b_1) = 0, \sigma^1(b_2) = 1/2, \sigma^1(b_3) = 1/2\}$. Thus, $M^1 \neq B^1$.

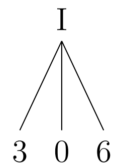

Figure 4: Γ' with $B^i \neq M^i$

For the general case without the constraint on Λ, the conclusion in Lemma A.1 is no longer true. In the one-player game Γ'' below, we denote $B^1 = \{b_1, b_2\}$. The strategy b_1 and b_2 leads to the same expected payoff, and hence can be replaced by each other. So $B^1 \neq S^1$.

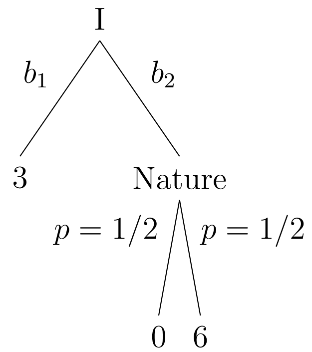

Figure 5: Γ'' with $B^i \neq S^i$

A.2 Generic assumption

We introduce the generic assumption to rule out any payoff tie which may emerge through any backward induction process in our model. The definition of a generic game discussed here is from Cressman (2003).

An extensive-form game Γ is *generic* if no two pure strategy profiles that yield different outcomes generate the same payoff for any player. For extensive-form games without chance nodes this is equivalent to the property that no two terminal nodes have the same payoff for any player. To put it formally, given an extensive-form game Γ, the set of all possible payoffs assigned to terminal nodes for which Γ is generic is a set whose complement is a closed set with Lebesgue measure zero.

A.3 Preliminary lemmas for Theorem 3.1

Given a play $h = (n_1, ..., n_{|h|})$, we define $X(h)$ to be the set of mixed strategy profiles x with $x(n_j, n_{j+1}) = 1$ for all $1 \leq j < |h|$. (Recall that $x(n_j, n_{j+1})$ is the probability of moving from n_j to n_{j+1} in the local behaviour strategy at node n_j.)

Lemma A.2 *If a play $h = (n_1, ..., n_{|h|})$ is not an equilibrium path in Γ, then there exists one node n_j in h such that for any mixed quasi strategy profile x in $X(h)$, there exists a pure quasi strategy $b^i \in B^i$ such that*

$$b^i(n_j) \neq n_{j+1} \text{ and } u^i(b^i, x^{-i}) > u^i(x) = v^i(h) \tag{A.1}$$

where $i = \lambda(n_j)$. We call such a n_j a deviation node.

Proof. Note that each Nash equilibrium component of Γ consists of a pure strategy Nash equilibrium together with other Nash equilibria with the same equilibrium path.

We prove the desired conclusion by contradiction. For each node n_j in h with $j < |h| - 1$, we transform the subgame Γ_{n_j} to a game Γ^{n_j} by replacing the node n_{j+1} with a terminal node n_h which has the same payoff vector as for the terminal node $n_{|h|}$ in the original play h. (Hence the subgame $\Gamma_{n_{j+1}}$ is not included in Γ^{n_j}.) Note that $\Gamma^{n_{|h|-1}} = \Gamma_{n_{|h|-1}}$.

If the assumption (A.1) is not true, then for any n_j and the generated game Γ^{n_j} there exists a distribution $x^{-\lambda(n_j)}(\Gamma^{n_j})$ such that the best reply of player $\lambda(n_j)$ can be a pure quasi strategy of the move (n_j, n_h). Back to the original game Γ, for each node $n \notin h$ there exists a unique j such that $n \in \Gamma^{n_j}$. We generate a mixed quasi strategy $\bar{x}$ from $(x^{-\lambda(n_j)}(\Gamma^{n_j}))_{j<|h|}$ such that

1. $\bar{x} \in X(h)$;

2. for every $n \notin h$, the local behaviour strategy of $\bar{x}$ at node n agrees with the one in $x^{-\lambda(n_j)}(\Gamma^{n_j})$ where n is in Γ^{n_j}.

In this way, we can see that play h is the path of a Nash equilibrium $\bar{x}$ in Γ. Contradiction. □

Given a node n_j in a play $h = (n_1, ..., n_{|h_r|})$, we denote

$$B_{n_j} := \{b^{\lambda(n_j)} \in B^{\lambda(n_j)} : n_j \in \bar{N}_{b^{\lambda(n_j)}} \text{ and } b^{\lambda(n_j)}(n_j) \neq n_{j+1}\}. \tag{A.2}$$

Suppose that player i plays at node n_j. For every x^{-i} in $X(h)^{-i}$, let $g_{x^{-i}}(n_j)$ be a pure quasi strategy of player i which is a best response within B_{n_j} for the strategy profile x^{-i}, i.e.,

$$g_{x^{-i}}(n_j) \in \underset{b^i \in B_{n_j}}{\operatorname{argmax}}\ u^i(b^i, x^{-i}). \tag{A.3}$$

If no ambiguity, we drop (n_j) in $g_{x^{-i}}(n_j)$. Let

$$\hat{u}(x^{-i}) := \max_{b^i \in B_{n_j}} u^i(b^i, x^{-i}), \tag{A.4}$$

so $\hat{u}(x^{-i}) = u^i(g_{x^{-i}}, x^{-1})$.

Lemma A.3 *We consider a k-player generic extensive-form game Γ of perfect information with κ terminal nodes. If a play $\bar{h} = (n_1, ..., n_{|\bar{h}|})$ is not an equilibrium path in Γ, then we take a deviation node n_j defined in Lemma A.2, and suppose player i moves at the node n_j. Then,*

$$\min_{x \in X(\bar{h})} \left(\hat{u}(x^{-i}) - v^i(\bar{h})\right) \geq \min_{i=1}^{k} \left\{ \min_{l=1}^{\kappa-1} \left\{ v^i_{(l)} - v^i_{(l+1)} \right\} \right\}, \tag{A.5}$$

where $\hat{u}(x^{-1})$ is defined in (A.4) with respect to n_j in $\bar{h}$, and $v^i_{(l)}$ is defined at the start of Section 3.1.

Note that it is straightforward to see that the left hand in (A.5) is bounded away from 0. We define a function $f' : X(\bar{h})^{-i} \to \mathbb{R}$ such that $f'(x^{-i}) = \hat{u}(x^{-i}) - v^i(\bar{h})$. From Lemma A.2, $f'(x^{-i}) > 0$ for all x^{-i} in $X(\bar{h})$. Since f' is continuous and $X(\bar{h})^{-i}$ is compact, f' is bounded away from 0.

Proof. For every x in $X(\bar{h})$, we define below a sequence $(x_0^{-i}, x_1^{-i}, ..., x_l^{-i})$ such that

1. $x_s^{-i} \in X(\bar{h})^{-1}$ for all s with $0 \leq s \leq l$;
2. $x_0^{-i} = x^{-i}$;
3. $g_{x_l^{-i}} = g_{x_{l-1}^{-i}}$;
4. $\hat{u}(x_s^{-i}) \geq \hat{u}(x_{s+1}^{-i})$ for all $1 \leq s < l$.

Given an x_s^{-i} with $s \geq 0$, we take a $g_{x_s^{-i}}$. If the set

$$H_s := \{h : \rho_{(x_s^{-i}, g_{x_s^{-i}})}(h) > 0\} \tag{A.6}$$

contains multiple elements, we take an h in $\underset{h \in H_l}{\operatorname{argmin}}\ v^i(h)$ and denote it as $h_s = (n_1^s, ..., n_{|h_s|}^s)$. We now set x_{s+1}^{-i} as follows. For every non-terminal node n not played by player i in the subgame Γ_{n_j}, i.e., in the set $N_d(\Gamma_{n_j}) \cap \Gamma^{-i}$, we check the distance of node n from h_s, i.e., $\min\{q : \psi^q(n) \in h_s\}$.

1. If $q = 0$, i.e., node n is in the play h_s, we let $x(n, n') = 1$ where $n' \in h_s \cap \psi^{-1}(n)$, i.e., it assigns a move along the play h_s with probability 1.

2. If $q > 0$, i.e., node n is not in the play h_s, we assign the same local behaviour strategy at node n as in x_s^{-1}, if n and $\psi^q(n)$ are played by different players and x_s^{-1} requires a move at node n.

So at any node n in $h_s \cap \Lambda^i \cap N_d(\Gamma_{n_j})$, if we cut the edge from node n to its successor and predecessor node in h_s respectively, then the projection distribution of x^{-i} in this generated game is the same between x_s^{-i} and x_{s+1}^{-i}. From the definition of $g_{x_s^{-i}}$, it follows $\hat{u}(x_s^{-i}) \geq \hat{u}(x_{s+1}^{-i})$. We repeat the process above and obtain $g_{x_{s+2}^{-i}}$ and so on.

Because Γ has finitely many nodes and Γ is generic, we will come to an x_l^{-i} which can generate a $g_{x_l^{-i}}$ the same as $g_{x_{l-1}^{-i}}$, so the set H_l defined in (A.6) is a singleton. Note that at every node not played by player i in the only play h_l in H_l, x_l^{-1} requires the move along h_l. So $\hat{u}(x_l^{-i}) = v^i(h_l) > v^i(\bar{h})$. □

A.4 Notations, operations and preliminary results for Theorem 4.2

Denote the set of decision nodes in a finite extensive-form game by $N_d(G)$ and let $r(\Gamma) := |N_d(\Gamma)|$ for the game Γ. If no ambiguity, we use the shorthand r for $r(\Gamma)$.

Lemma A.4 *Given a μ-dynamic $(x_t)_t$ on the k-player game Γ with $\mu < 1/(2^{r+2})$ and given a positive number $\epsilon < 1/(2^{r+1})$, suppose there exists a play $h := (n_1, ..., n_{|h|})$ and a time t_0 such that $\rho_{t_0}(h) \geq 1 - \epsilon$ and $\dot{\rho}(h)|_{t_0} \geq 0$. Then $\dot{x}(n_j, n_{j+1})|_{t_0} \geq 0$ for all j with $1 \leq j < |h|$.*

Proof. Firstly note that, when $r = 1$, i.e., Γ contains only one non-terminal node, the result is trivial.

We prove by contradiction for the case $r > 1$. Suppose the desired conclusion is not true. Then, at any t with

$$\rho_t(h) \geq 1 - \epsilon > 1 - 1/2^{r+1}, \tag{A.7}$$

there exists at least one n_j such that $\dot{x}(n_j, n_{j+1})|_t < 0$.

Without loss of generality, we assume that $\lambda(n_i) \neq \lambda(n_l)$ for all $i \neq l$, i.e., each non-terminal node in h is played by a different player. It follows from $\rho_t(h) > 1 - 1/2^{r+1}$ that $x_t(n_j, n_{j+1}) > 1 - 1/2^{r+1}$ for all j with $1 \leq j < |h|$. From

$$\left(\prod_{j=1}^{|h|-1} x(n_j, n_j + 1) \right)' = \sum_{l=1}^{|h|-1} \left(\dot{x}(n_l, n_{l+1}) \prod_{j \neq l} x(n_j, n_{j+1}) \right) \tag{A.8}$$

and the definition of μ-dynamic, we may infer that

$$\begin{aligned} & \dot{\rho}(h)|_t \\ &= \left(\prod_{j=1}^{|h|-1} x(n_j, n_j + 1) \right)' \Bigg|_t \\ &\leq (\epsilon + \mu)\,(|h| - 2) - (1 - \epsilon - \mu)^{|h|-1} \qquad \text{(A.9)} \\ &\leq \left(\frac{1}{2^{r+1}} + \mu \right) (r - 1) - \left(1 - \frac{1}{2^{r+1}} - \mu \right)^r . \qquad \text{(A.10)} \end{aligned}$$

Since $\mu < 1/(2^{r+2})$,

$$\left(\frac{1}{2^{r+1}} + \mu \right) (r - 1) \leq \frac{5}{16} \tag{A.11}$$

and

$$\left(1 - \frac{1}{2^{r+1}} - \mu \right)^r \geq \left(1 - \frac{5}{2^{r+2}} \right)^r \Bigg|_{r=2} = \frac{121}{256} \tag{A.12}$$

for all r with $r > 1$.

From (A.10), (A.11) and (A.12) it follows that $\dot{\rho}(h)|_t < 0$. Contradiction. □

Recall that the set of plays is denoted by H. Given an $\epsilon > 0$ and a dynamic process $(x_t)_{t \geq 0}$, we denote

$$S(\epsilon) := \{t : \rho_t(h) > 1 - \epsilon \text{ and } \dot{\rho}(h)|_t > 0 \text{ for some } h \in H\}.$$

and

$$W(\epsilon) := \{t : \rho_t(h) > 1 - \epsilon \text{ for some } h \in H\}. \tag{A.13}$$

From Lemma A.4 and Lemma 3.7 with (3.16) replaced by

$$(r+1)\epsilon < \bar{\epsilon}, \tag{A.14}$$

(this adjustment is for an ϵ-dynamic, c.f. (3.17) and (3.18)) we have the following corollary.

Corollary A.5 *Given the game Γ, we take any positive $\epsilon < 1/(2^{r+1})$. For a μ-dynamic $(x_t)_{t\geq 0}$ on Γ with $\mu < \epsilon/2$ and with the property (A.14), it follows*

$$x_t \in S(\epsilon) \Rightarrow x_t \in NE[\epsilon].$$

We then have the following theorem.

Theorem A.6 *Given the game Γ, suppose there exists a positive number ϵ with the following properties.*

1. $\epsilon < 1/(2^{r+1})$;
2. *(A.14) holds;*
3. $$\frac{r\epsilon}{(1-\epsilon)^r - (r-1)\epsilon} < \frac{2r\epsilon}{1-2r\epsilon}. \tag{A.15}$$

Given a positive number T', if an $\epsilon/2$-dynamic satisfies

$$\frac{1}{T}\int_0^T \mathbb{1}_{W(\epsilon)} \mathrm{d}t \geq 1-\epsilon$$

for all $T > T'$, then

$$\frac{1}{T}\int_0^T \mathbb{1}_{\{t: x_t \in NE[\epsilon]\}} \mathrm{d}t \geq 1-(2r+1)\epsilon \tag{A.16}$$

for all $T > T'$.

Proof. For an $\epsilon/2$-dynamic on Γ, suppose there exists a time t_1 and a play $h := (n_1, ..., n_{|h|})$ with $\rho_{t_1}(h) \geq 1-\epsilon$ and $\dot{\rho}(h)|_{t_1} > 0$. Then, by Lemma A.4, $\dot{x}(n_j, n_{j+1})|_{t_1} \geq 0$ for all j with $1 \leq j < |h|$. Denote

$$t_2 := \min\{t > t_1 : \rho_t(h) = \rho_{t_1}(h)\}. \tag{A.17}$$

We know that, when $\rho_t(h) \geq 1-\epsilon$ at any t, $x_t(n_j, n_{j+1}) \geq 1-\epsilon$ for all $j < |h|$. From the definition of an approximate best response dynamic and (A.9), we know that when $\dot{\rho}(h)|_t \geq 0$, $\dot{\rho}(h)|_t < r\epsilon$; when $\dot{\rho}(h)|_t < 0$, $\dot{\rho}(h)|_t <$

$-(1-\epsilon)^r+(r-1)\epsilon$. So, it follows from(A.17) and (A.15) that, if t_2 is finite, then

$$\int_{t_1}^{t_2} \mathbb{1}_{\{t:\dot{\rho}(h)|_t<0\}}\mathrm{d}t < (2r\epsilon)(t_2-t_1).$$

We then apply Corollary A.5 to complete the proof. □

Tree operations: Recall that $N_d(G)$ is the set of decision nodes in the game G. Given a generic extensive-form game G of perfect information with $|N_d(G)| = r(G) > 1$, define the set of pre-terminal nodes in G as

$$L(G) := \{n \in N_d(G) : \Psi^{-1}(n) \subseteq N_t(G)\}, \tag{A.18}$$

i.e., $L(G)$ is the subset of decision nodes whose successors are all terminal nodes. We take a node $\tilde{n} \in L(G)$ and introduce below two operations on game G.

1. Given the node $\tilde{n}$, we denote the last node played by player $\lambda(\tilde{n})$ before $\tilde{n}$ by $\hat{n}$. So
$$\Psi(\tilde{n}) \cap \Lambda^{\lambda(\tilde{n})} = \{\hat{n}\} \cup (\Psi(\hat{n}) \cap \Lambda^{\lambda(\tilde{n})}).$$
(If such $\hat{n}$ does not exist, then we will not implement this operation on G in the proof.) We then denote the node n' to be the immediate successor node of $\hat{n}$ in the play to $\tilde{n}$, i.e., $n' \in \psi^{-1}(\hat{n}) \cap (\Psi(\tilde{n}) \cup \{\tilde{n}\})$. We define game $\tilde{G}$ to be the game with the edge $(\hat{n}, n')$ together with the subgame $G_{n'}$ removed from G. So $|N_d(\tilde{G})| \leq r(G) - 1$.

   Given the game $\tilde{G}$, for each player i, we derive the set $\tilde{B}^i$ of pure quasi strategies and the set $\tilde{X}^i$ in $\tilde{G}$ from the original game G as follows. For each player i, we define a surjective function $q^i : B^i \to \tilde{B}^i$ such that, for every b^i in B^i, $q^i(b^i)$ and b^i have the same sequence of moves in $\tilde{G}$, i.e.,
$$\tilde{b}^i(n) = b^i(n) \ \forall n \in N_d(\tilde{G}) \cap \bar{N}_{b^i} \tag{A.19}$$
when $\tilde{b}^i = q^i(b^i)$. (Recall the definition of $\bar{N}_{b^i}$ for a pure quasi strategy b^i in the second paragraph in Section 3.1.) If no ambiguity, we may drop the superscript i in q^i. We define $\tilde{X}^i$ as in (2.2). We may also view each $\tilde{b}^i$ in $\tilde{B}^i$ as a class of pure quasi strategies in Γ, and thus define
$$x^i(\tilde{b}^i) := \sum_{b^i:q(b^i)=\tilde{b}^i} x^i(b^i). \tag{A.20}$$

2. Denote $\psi(\tilde{n})$ by $\bar{n}$. In $\psi^{-1}(\tilde{n})$, denote the terminal node with the maximum payoff to the player who moves at node $\tilde{n}$ by $\tilde{n}'$, i.e.,
$$v^{\lambda(\tilde{n})}(\tilde{n}') = \max_{n\in\psi^{-1}(\tilde{n})} v^{\lambda(\tilde{n})}(n). \tag{A.21}$$

We define a new game $\bar{G}$ with $|N_d(\bar{G})| = r(G) - 1$ by the following two steps.

(a) Remove subgame $G_{\tilde{n}}$ and the edge $(\bar{n}, \tilde{n})$. We arbitrarily index all remaining non-terminal nodes as $n^1, n^2, ..., n^{|r(G)-1|}$. For each node n^j, we add two terminal nodes n_1^j and n_2^j to n^j with the property that
$$v^i(n_1^j) = v^i_{(|N_t(G)|)} - j \tag{A.22}$$
for all players i; $v^i(n_2^j) = v^i_{(1)} + j$ for all players $i \neq \lambda(n^j)$ and $v^{\lambda(n^j)}(n_2^j) = v^{\lambda(n^j)}_{(|N_t(G)|)} - r(G) - j$. (Recall the definition of $v^i_{(l)}$ in the game G at the start of Section 3.1.)

(b) Add a new terminal node $\tilde{n}$ to $\bar{n}$ with $\mathbf{v}_{\tilde{n}} = \mathbf{v}_{\tilde{n}'}$, where $\tilde{n}'$ is the terminal node denoted in (A.21) in G.

We can see that the generated game $\bar{G}$ is a generic game.

A.5 Proof of Theorem 4.2

We prove Theorem 4.2 by induction on the number of non-terminal nodes.

Induction hypothesis: Given a natural number $\eta \geq 1$ and any finite generic extensive-form game G of perfect information with $r(G) \leq \eta$ and without chance nodes, for every $\epsilon > 0$ there exists a number $\mu_G(\epsilon)$ with $0 < \mu_G(\epsilon) < \epsilon$ and a number $T_G(\epsilon) > 0$ such that for every interior $\mu_G(\epsilon)$-dynamic $(x_t)_t$ on G with any initial state x_0
$$\frac{1}{T}\int_0^T \mathbb{1}_{\{t: x_t \in NE[\epsilon]\}} \mathrm{d}t \geq 1 - \epsilon \tag{A.23}$$
for all $T \geq T_G(\epsilon)$. In the proof, we also apply a weaker version with (A.23) replaced by
$$\frac{1}{T}\int_0^T \mathbb{1}_{W(\epsilon)} \mathrm{d}t \geq 1 - \epsilon. \tag{A.24}$$
(Recall $W(\epsilon)$ in (A.13).) Note that (A.23) trivially holds when $r(G) = 1$.

For the game Γ in Theorem 3.1, we let $r(\Gamma) = \eta + 1 > 1$ and assume that the induction hypothesis above is true for all games G with $r(G) \leq \eta$. We pick a node $\tilde{n}$ in $L(\Gamma)$ (defined in (A.18)) and suppose that player i plays at node $\tilde{n}$.

Step I [when the dynamic in the subgame $\Gamma_{\tilde{n}}$ can be ignored due to (A.25) below]

Recall the definition of $\bar{N}_{b^i}$ for a pure quasi strategy b^i in the second paragraph in Section 3.1.

Lemma A.7 *Given a μ-dynamic $(x_t)_{t\geq 0}$ on Γ with $\mu < 1/2$, suppose*

$$\sum_{b^i:\tilde{n}\in\bar{N}_{b^i}} x_t^i(b^i) < \mu \ \forall t \geq 0. \tag{A.25}$$

For the game $\tilde{\Gamma}$ generated by Tree Operation 1 with respect to node $\tilde{n}$, there exists a 4μ-dynamic $(\tilde{x}_t)_t$ on $\tilde{\Gamma}$ such that for all players $j \neq i$, $x_t^j(\tilde{b}^j) = \tilde{x}_t^j(\tilde{b}^j)$ for all $\tilde{b}^j$ in $\tilde{B}^j$ and for all $t \geq 0$ (recall $x_t^j(\tilde{b}^j)$ in (A.20)); for player i,

$$|x_t^i(\tilde{b}^i) - \tilde{x}_t^i(\tilde{b}^i)| \leq 3\mu$$

for all $\tilde{b}^i$ in $\tilde{B}^i$ and for all $t \geq 0$. Moreover, for all players $j \neq i$, $\tilde{g}_t^j = q(g_t^j)$ for all $t \geq 0$, and, for player i, if $\tilde{n} \notin \bar{N}_{g_t^i}$ at time t, then $\tilde{g}_t^i = g_t^i$.

Proof. Firstly note that $\tilde{B}^i$ in $\tilde{\Gamma}$ is a subset of B^i in Γ.

We define an assistant process $(z_t^i)_{t\geq 0}$ on G with the property that $z_0^i = x_0^i$ and $\dot{z}^i = \tilde{g}^i - z^i$ where the function $\tilde{g}_t^i := g_t^i$ when $\tilde{n} \notin \bar{N}_{g_t^i}$ and $\tilde{g}_t^i$ is arbitrarily defined in $\tilde{B}^i$ when $\tilde{n} \in \bar{N}_{g_t^i}$. We show from the following observation that, in any such process $(z_t^i)_{t\geq 0}$, for any $t \geq 0$,

$$|x_t^i(\tilde{b}^i) - z_t^i(\tilde{b}^i)| \leq \mu \tag{A.26}$$

for all $\tilde{b}^i \in \tilde{B}^i \subseteq B^i$.

If for a period of time $[t_1, t_2]$ whenever $\tilde{g}_t^i \neq g_t^i$, $\tilde{g}_t^i$ is the same, say, $\tilde{g}_t^i = \hat{b}^i \in \tilde{B}^i$ for all time t in $[t_1, t_2]$ when $\tilde{g}_t^i \neq g_t^i$. If $z_{t_1}(\hat{b}^i) = x_{t_1}(\hat{b}^i)$, then $z_{t_1}(\hat{b}^i) > x_{t_1}(\hat{b}^i)$. (This holds despite point 2 may hold at some t in $[t_1, t_2]$.) We may further infer that $z_t(\tilde{b}^i) \geq x_t(\tilde{b}^i)$ for all $\tilde{b}^i \in \tilde{B}^i$ and all $t \geq 0$. So we reach (A.26).

At $t = 0$, for all players $j \neq i$, we let $\tilde{x}_0^j(\tilde{b}^j) = x_0^j(\tilde{b}^j)$ for all $\tilde{b}^j$ in $\tilde{B}^j$, and, for player i, we normalise x_0^i in $\tilde{X}^i$ in the way that

$$\forall \tilde{b}^i \in \tilde{B}^i,\ \tilde{x}_0^i(\tilde{b}^i) = \frac{x_0^i(\tilde{b}^i)}{\sum_{\tilde{b}^i\in\tilde{B}^i} x_0^i(\tilde{b}^i)}.$$

So

$$\tilde{x}_0^i(\tilde{b}^i) < \frac{x_0^i(\tilde{b}^i)}{1-\mu} < (1+2\mu)x_0^i(\tilde{b}^i)\ \forall \tilde{b}^i \in \tilde{B}^i. \tag{A.27}$$

If in $(x_t)_{t\geq 0}$, $\tilde{n} \notin \bar{N}_{g_t^i}$ for all time $t \geq 0$, then the desired conclusion trivially holds. For the general case, consider the assistant process $(z_t^i)_t$ above, and from (A.26) and (A.27), we can always set a 4μ-dynamic $(\tilde{x}_t)_t$ such that for each player $j \neq i$, $\tilde{g}_t^j = q(g_t^j)$ for all $t > 0$. Thus, we are allowed to set $\tilde{x}_t^i$ such that $|x_t^i(\tilde{b}^i) - \tilde{x}_t^i(\tilde{b}^i)| \leq 3\mu$ for all $\tilde{b}^i$ in $\tilde{B}^i$ and for all $t > 0$. $\square$

Lemma A.8 *Given the game* Γ *and an* $\epsilon > 0$*, there exists a* $\mu > 0$ *and a* $\bar{T} > 0$ *with the following property. For any interior* μ*-dynamic in* Γ *with the property (A.25) for some node* $\tilde{n}$ *in* $L(\Gamma)$*, (A.16) holds for all* $T \geq \bar{T}$ *in the* μ*-dynamic.*

Proof. Take an ϵ_Γ which satisfies Conditions 1 to 3 in Theorem A.6. Denote $\min\{\epsilon/4, \epsilon_\Gamma\}$ by ϵ_1. Since $|N_d(\tilde{\Gamma})| < |N_d(\Gamma)|$, we apply the induction hypothesis on $\tilde{\Gamma}$ and obtain a $\mu_{\tilde{\Gamma}}(\epsilon_1) < \epsilon_1$. Consider any positive $\mu \leq \mu_{\tilde{\Gamma}}(\epsilon_1)/4$. By (A.24) in the induction hypothesis with respect to a 4μ-dynamic in $\tilde{\Gamma}$ and Lemma A.7, we can find $T_{\tilde{\Gamma}}(\epsilon_1)$ such that in the interior μ-dynamic in Γ

$$\frac{1}{T}\int_0^T \mathbb{1}_{W(\epsilon_1+3\mu)}\mathrm{d}t > 1 - \epsilon_1$$

for all $T \geq T_{\tilde{\Gamma}}(\epsilon_1)$. Note $\epsilon_1 + 3\mu < \epsilon$. By Theorem A.6, we complete the proof. □

Step II [when $\tilde{n}$ can be viewed as a terminal node due to (A.28) and (A.29)]

Recall that node $\tilde{n}$ we picked in Γ is played by player i and the game $\bar{\Gamma}$ obtained by Tree Operation 2. (We can tell in the proof whether $\tilde{n}$ refers to the decision node in Γ or the terminal node in $\bar{\Gamma}$ from the context.) We define a function $\hat{q} : X \to \bar{X}$ with the property that, if $\hat{q}(x) = \bar{x} \in \bar{X}$, then $x(n_1, n_2) = \bar{x}(n_1, n_2)$ for all edges (n_1, n_2) in both Γ and $\bar{\Gamma}$ when $x(n_1, n_2)$ is defined in Γ. (When we say $x(n_1, n_2)$ is defined, we mean there exists a pure quasi strategy $b^{\lambda(n_1)}$ such that $x^{\lambda(n_1)}(b^{\lambda(n_1)}) > 0$ and $n_1 \in \bar{N}_{b^{\lambda(n_1)}}$. Please refer to Section 3.1 for more details.) Thus, $\hat{q}(x)$ is the canonical projection of the mixed quasi strategy profile x in $\bar{\Gamma}$. Moreover, $\bar{x}(n, n_1) = \bar{x}(n, n_2) = 0$, if $\bar{x}(n, n_1)$ and $\bar{x}(n, n_2)$ is defined, where n_1 and n_2 are the terminal nodes denoted in step (a) in the Tree Operation 2.

Lemma A.9, A.10 and A.11 discuss, if $\bar{x} \in \hat{q}(x)$, then under what condition $BR^l(x) \subseteq BR^l(\bar{x}[\mu])$ for a player $l \neq i$.

Recall the constant δ defined for Γ in (3.3). Denote the player who plays at node $\bar{n}$ by player j.

Lemma A.9 *Given a mixed quasi strategy profile* x *in* Γ *and a positive number* μ*, we suppose there exists a node* $\tilde{n} \in L(\Gamma)$ *with* $\tilde{n}'$ *as defined in (A.21) and the property*

$$0 < \sum_{n \in \psi^{-1}(\tilde{n}) \setminus \{\tilde{n}'\}} \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}, b^i(\tilde{n}) = n} x^i(b^i) \leq \frac{\mu^4}{4} \tag{A.28}$$

$$\frac{\mu^2}{\delta} \leq \sum_{b^i:\tilde{n}\in\bar{N}_{b^i}} x^i(b^i) < 1. \tag{A.29}$$

We denote in game $\bar{\Gamma}$ (obtained by Tree Operation 2 with respect to node $\tilde{n}$) a mixed quasi strategy profile $\bar{x} := \hat{q}(x)$. Consider any $g^j \in BR^j(x[\mu^4/4])$ in Γ and $\bar{g}^j \in BR^j(\bar{x})$ in $\bar{\Gamma}$. The following two conditions cannot hold at the same time.

1. *In $\bar{\Gamma}$, there are two subplays $(n_1^1, n_2^1, ..., n_{k_1}^1)$ and $(n_1^2, n_2^2, ..., n_{k_2}^2)$ with $n_{k_1}^1 = \tilde{n} \neq n_{k_2}^2$, $n_{k_2}^2 \in N_t(\bar{\Gamma})$ and $n_1^1 = n_1^2$. All decision nodes in these two subplays in $\bar{\Gamma}$ are played by player j. (Recall that player j plays at node $\bar{n}$.)*

2. *For $i = 1$ or 2, $n_l^i \in \bar{N}_{g^j}$ and $g^j(n_l^i) = n_{l+1}^i$ for all $l < k_1$; $n_l^{3-i} \in \bar{N}_{\bar{g}^j}$ and $\bar{g}^j(n_l^{3-i}) = n_{l+1}^{3-i}$ for all $l < k_{3-i}$.*

Proof. It follows from (A.28) and (A.29) that the local behaviour strategy of player i at node $\tilde{n}$ in Γ satisfies

$$x'(\tilde{n}, \tilde{n}') \geq \frac{\frac{\mu^2}{\delta} - \frac{\mu^4}{2}}{\frac{\mu^2}{\delta}} > 1 - \frac{\delta}{2}.$$

for all $x' \in x[\mu^4/4]$. We complete the proof by the definition of δ in (3.3). □

Recall that the decision node $\tilde{n}$ is played by player i in Γ. Note that $\bar{B}^l = B^l$ for all $l \neq i$.

Lemma A.10 *Given any x in X in Γ, we obtain $\bar{x} := \hat{q}(x)$ in $\bar{\Gamma}$. For any player $l \neq i$, we consider a best reply $g^l \in BR^l(x)$ and a $\bar{g}^l \in BR^l(\bar{x})$ in two games Γ and $\bar{\Gamma}$, respectively. If $g^l \neq \bar{g}^l$, then there is a node n^l such that, $n^l \in \bar{N}_{g^l} \cap \bar{N}_{\bar{g}^l}$, and, for all n in $\Psi^{-1}(n^l) \cap \Lambda^l$, $n \notin \bar{N}_{g^l} \cap \bar{N}_{\bar{g}^l}$. For any x, we can always find a pair of $(g^l, \bar{g}^l)$ such that, if $g^l \neq \bar{g}^l$, then there is at most one such node n^l and*

$$\Psi^{-1}(n^l) \ni \tilde{n} \tag{A.30}$$

in $\bar{\Gamma}$.

Remark: A node n^l is where g^l and $\bar{g}^l$ are deviating from each other.

Proof. This follows from $\bar{x} = \hat{q}(x)$ and that the projection distribution of x in $\Gamma_{\tilde{n}}$ is different from the projection of $\bar{x}$ in $\bar{\Gamma}_{\tilde{n}}$. □

Lemma A.11 *We continue with the setting in Lemma A.10 and assume only one node n^l exists with the property (A.30). We define a set $\hat{N}$ with the property $\hat{N} \subseteq (\Psi^{-1}(n^l) \cap \Lambda^{-l})$, and for each $\hat{n} \in \hat{N}$ all nodes n in $\Psi(\hat{n}) \cap \Psi^{-1}(n^l)$ are in Λ^l.*

If $\hat{N} \neq \emptyset$ and both (A.28) and (A.29) hold, then there exists a $\bar{z} \in \bar{x}[\mu^2/4]$ and a $\bar{g}^l(\bar{z}) \in BR^l(\bar{z})$ such that $\bar{g}^l(\bar{z}) = g^l$.

Remark: $\hat{N}$ is the subset in $\Psi^{-1}(n^l)$ which contains all first nodes not played by player l after n^l. For any two nodes $n^1, n^2 \in \hat{N}$, $n^1 \notin \Psi(n^2) \cup \Psi^{-1}(n^2)$.

Proof. Firstly note that one and only one of the following statements is true:

$$g^l(n^l) \in \Psi(\tilde{n}) \cap \{\tilde{n}\} \text{ and } \bar{g}^l(n^l) \notin \Psi(\tilde{n}) \tag{A.31}$$

or

$$\bar{g}^l(n^l) \in \Psi(\tilde{n}) \cap \{\tilde{n}\} \text{ and } g^l(n^l) \notin \Psi(\tilde{n}). \tag{A.32}$$

We first assume (A.31) holds, and we generate a $\bar{z} \in \bar{x}[\mu^2/4]$ as follows. For the local behaviour strategy at every node $n \in N_d(\bar{\Gamma}) \setminus \hat{N}$, we let $\bar{z}(n, n') = \bar{x}(n, n')$ whenever $\bar{x}(n, n')$ is defined in $\bar{\Gamma}$.

If $\hat{N} \cap \Psi(\tilde{n}) \neq \emptyset$, then from the remark above there is only one node $\hat{n}$ in $\hat{N} \cap \Psi(\tilde{n})$, and we let $\bar{z}(\hat{n}, \hat{n}_2) = \mu^2/4$, where $\hat{n}_2$ is defined in step (a) in the Tree Operation 2; $\bar{z}(\hat{n}, \hat{n}') = (1 - \mu^2/4)\bar{x}(\hat{n}, \hat{n}')$ for all nodes $\hat{n}'$ in $\psi^{-1}(\hat{n}) \setminus \{\hat{n}_2\}$.

For every node $\hat{n}$ in $\hat{N} \setminus \Psi(\tilde{n})$, we let $\bar{z}(\hat{n}, \hat{n}_1) = \mu^2/4$, where $\hat{n}_1$ is defined in step (a) in the Tree Operation 2; $\bar{z}(\hat{n}, \hat{n}') = (1 - \mu^2/4)\bar{x}(\hat{n}, \hat{n}')$ for all nodes $\hat{n}'$ in $\psi^{-1}(\hat{n}) \setminus \{\hat{n}_1\}$.

If (A.32) holds, we apply the similar process to obtain $\bar{z}$ but with $\hat{n}_1$ and $\hat{n}_2$ swapped.

We can check that there exists a such $\bar{z}$ with $\bar{z} \in \bar{x}[\mu^4/4]$ and $\bar{g}^l(\bar{z}) = g^l$. □

Lemma A.12 *Given a $\mu^4/4$-dynamic $(x_t)_{t\geq 0}$ on Γ, we suppose (A.28) and (A.29) hold at all $t \geq 0$. Then there exists a $\mu^2/2$-dynamic $(\bar{x}_t)_{t\geq 0}$ in $\bar{\Gamma}$ (obtained by Tree Operation 2) with the property*

$$\bar{x}_t = \hat{q}(x_t). \tag{A.33}$$

at all $t \geq 0$.

Proof. We define a $\mu^2/2$-dynamic $(\bar{x}_t)_{t\geq 0}$ on $\bar{\Gamma}$ with the following propositions. Recall the step (a) in Tree Operation 2. For any decision node n in $\bar{\Gamma}$, we denote $S_n^1 := \{b \in \bar{B}^{\lambda(n)} : n \in \bar{N}_b,\ b(n) = n_1\}$ and $S_n^2 := \{b \in \bar{B}^{\lambda(n)} : n \in \bar{N}_b,\ b(n) = n_2\}$. We require that (A.33) holds at time $t = 0$. So for each decision node n in $\bar{\Gamma}$ and the associated two terminal nodes $n_1, n_2 \in \psi^{-1}(n)$, it follows

$$\sum_{b \in S_n^1} \bar{x}_0^{\lambda(n)}(b) = \sum_{b \in S_n^2} \bar{x}_0^{\lambda(n)}(b) = 0.$$

From the propositions of node n_1 and n_2 and the definition of μ-dynamics, we know that

$$\sum_{b \in S_n^1} \bar{x}_t^{\lambda(n)}(b) = \sum_{b \in S_n^2} \bar{x}_t^{\lambda(n)}(b) = 0$$

for all $t \geq 0$. We show below that there exists a such $\mu^2/2$-dynamic $(\bar{x}_t)_{t\geq 0}$ on $\bar{\Gamma}$ which satisfies (A.33) for all $t \geq 0$.

When $\bar{x} = \hat{q}(x_t)$, it is straightforward to see that $BR^i(x_t[\mu^4/4])$ in Γ a subset of $BR^i(\bar{x}[\mu^4/4])$, where player i moves at the node $\tilde{n}$ in Γ. Recall that player j moves at node $\bar{n}$. For a player $l = j$ and with $\hat{N} \neq \emptyset$ (defined in Lemma A.11), or a player l with $l \neq j$ and $l \neq i$, which induces $\hat{N} \neq \emptyset$, we can always adjust $\bar{x}(n, n_1)$ and $\bar{x}(n, n_2)$ in some local behaviour strategies as shown in the proof of Lemma A.11, and make $BR^l(x_t[\mu^4/4])$ in Γ a subset of $BR^l(\bar{x}[\mu^2/2])$ in $\bar{\Gamma}$. (The final adjustment consists of two parts: imitation to the $x \in x_t[\mu^4/4]$ and the adjustment shown in Lemma A.11.)

Note that, if player $l = j$ and $\hat{N} = \emptyset$, then we cannot adjust by that way to make $BR^j(x_t[\mu^4/4]) \subseteq BR^j(\bar{x}[\mu^2/2])$, since player j's best response cannot include a move to her worst payoff. However, $\hat{N} = \emptyset$ leads to Condition 1 in Lemma A.9, and it follows that any $g^j_t \in BR^j(x_t[\mu^4/4])$ is also a best reply to some $\bar{x}' \in \hat{q}(x_t)[\mu^4/4]$ in $\bar{\Gamma}$. That completes the proof. □

Lemma A.13 *Given the game* Γ *and an* $\epsilon > 0$*, there exists a* $\bar{\mu}$ *with* $0 < \bar{\mu} < \epsilon$ *and a* $\bar{T} > 0$ *with the following property. For any interior* $\mu^4/4$*-dynamic* $(x_t)_t$ *in* Γ *with* $\mu < \bar{\mu}$ *and the property (A.28) and (A.29) at some node* $\tilde{n} \in L(\Gamma)$ *for all* $t \geq 0$*, (A.16) holds for all* $T \geq \bar{T}$*.*

Proof. Take an ϵ_Γ which satisfies Conditions 1 to 3 in Theorem A.6. Denote $\min\{\epsilon/2, \epsilon_\Gamma\}$ by ϵ_2. Since $|N_d(\tilde{\Gamma})| < |N_d(\Gamma)|$, we apply the induction hypothesis on $\bar{\Gamma}$ and obtain a $\mu_{\bar{\Gamma}}(\epsilon_2) < \epsilon_2$. Take $\bar{\mu} := (2\mu_{\bar{\Gamma}}(\epsilon_2))^{1/2}$, and consider any $\mu < \bar{\mu}$. Note that, for the dynamic $(x_t)_t$ with the property (A.28) and (A.29) and the associated dynamic $(\bar{x}_t)_t$ studied in Lemma A.12, if, $\bar{\rho}_t(\bar{h}) = p$ in $(\bar{x}_t)_t$ for the play $\bar{h}$ up to the terminal node $\tilde{n}$ in $\bar{\Gamma}$, then for $(x_t)_t$, $\rho_t(h) \geq p(1 - \mu^2/4)$ for the play $h = \bar{h} \cup \{\tilde{n}'\}$ (in Γ). By (A.24) in the induction hypothesis with respect to a $\mu^2/2$-dynamic in $\bar{\Gamma}$ and Lemma A.12, we can find $T_{\bar{\Gamma}}(\epsilon_2)$ such that for any interior $\mu^2/4$-dynamic in Γ

$$\frac{1}{T}\int_0^T \mathbb{1}_{W(\epsilon_2+\mu^2/4)}\mathrm{d}t > 1 - \epsilon_2$$

for all $T > T_{\bar{\Gamma}}(\epsilon_2)$. From $\mu^2/4 < \mu_{\bar{\Gamma}}(\epsilon_2)/2 < \epsilon_2/2 < \epsilon/4$, it follows

$$\frac{1}{T}\int_0^T \mathbb{1}_{W(\epsilon)}\mathrm{d}t > 1 - \epsilon/2$$

By Theorem A.6, we complete the proof. □

Step 3: [Find a small μ.]

We first give a corollary from Lemma A.8 and Lemma A.13.

Corollary A.14 *Given the game* Γ *and an* $\epsilon > 0$*, there exists a* $\hat{\mu}$ *with* $0 < \hat{\mu} < \epsilon$ *and a* $\hat{T} > 0$ *with the following property. For any interior* $\mu^4/4$*-dynamic* $(x_t)_t$ *in* Γ *with* $\mu < \hat{\mu}$*, any initial state* x_0 *and either of the two conditions below at some node* $\tilde{n} \in L(\Gamma)$ *for all* $t \geq 0$

1. *(A.25)*
2. *the conjunction of (A.28) and (A.29),*

it follows

$$\int_0^T \mathbb{1}_{\{t:x_t \in NE[\epsilon]\}} \mathrm{d}t \geq T(1-\epsilon) \tag{A.34}$$

for all $T \geq \hat{T}$.

Proof of Theorem 4.2: Recall $\tilde{n} \in L(\Gamma)$, and two associated games $\tilde{\Gamma}$ and $\bar{\Gamma}$ defined by Tree Operation 1 and Tree Operation 2, respectively. Recall that in Γ node $\tilde{n}$ is played by player i. We firstly note that for any interior approximate best response dynamic $(x_t)_{t\geq 0}$ on Γ, for any possible best reply g_t^i of plater i at any time t, if $\tilde{n} \in \bar{N}_{g_t^i}$, then $g_t^i(\tilde{n})$ must be $\tilde{n}'$. So, for any $c > 0$, there exists a time $\bar{t}$ such that

$$\forall t \geq \bar{t}, \quad \sum_{n \in \psi^{-1}(\tilde{n}) \setminus \{\tilde{n}'\}} \ \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}, b^i(\tilde{n}) = n} x_t^i(b^i) \leq c \tag{A.35}$$

for all approximate best response dynamics $(x_t)_{t\geq 0}$.

We take a μ under the constraints:

1. $$\mu < \delta/2; \tag{A.36}$$
2. $$\hat{T} < \ln\delta - \ln(2\mu), \tag{A.37}$$
   where $\hat{T}$ is defined in Corollary A.14;
3. $\mu < \hat{\mu}$ in Corollary A.14.

We consider any interior $\mu^4/4$-dynamic $(x_t)_t$ with the μ taken under the above constraints, and then transform the original dynamic to a dynamic $(z_t)_t$ with $z_t = x_{t+\bar{t}}$ where (A.35) holds with c replaced by $\mu^4/4$ for all $t \geq \bar{t}$. We define a sequence of times $(t_k)_{k\geq 0}$ with $t_0 = 0$ such that for any $t \geq 0$

$$t_{2k+1} = \min\left\{ t \geq t_{2k} : \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}} z_t^i(b^i) \geq \mu \right\}$$

and

$$t_{2k+2} = \min \left\{ t \geq t_{2k+1} : \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}} z_t^i(b^i) < \frac{\mu^2}{\delta} \right\}.$$

Note that it is possible that there exists a natural number l such that $t_k = +\infty$ for all $k \geq l$. During $[t_{2k+1}, t_{2k+2})$ for any $k \geq 0$, we can apply the condition of (A.28) and (A.29) in Corollary A.14. From (A.36), it follows $\mu/2 > \mu^2/\delta$. From (3.8), we may further infer

$$\begin{aligned} t_{2k+2} - t_{2k+1} > \min & \left\{ t > t_{2k+1} : \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}, b^i(\tilde{n}) = \tilde{n}'} z_t^i(b^i) = \frac{\mu}{2} \right\} \\ & - \min \left\{ t > t_{2k+1} : \sum_{b^i : \tilde{n} \in \bar{N}_{b^i}, b^i(\tilde{n}) = \tilde{n}'} z_t^i(b^i) = \frac{\mu^2}{\delta} \right\} \\ & > \ln \frac{\delta}{\mu^2} - \ln \frac{2}{\mu} \\ & = \ln \delta - \ln(2\mu). \end{aligned} \tag{A.38}$$

From the definition of t_{2k+1}, it follows that, for any t with $t_{2k} \leq t < t_{2k+1}$,

$$\sum_{b^i : \tilde{n} \in \bar{N}_{b^i}} z_t^i(b^i) < \mu. \tag{A.39}$$

We can then apply the condition (A.25) in Corollary A.14 for any $[t_{2k}, t_{2k+1})$ with $k \geq 0$. Recall (A.37), and we may then infer from Corollary A.14 that, during $[t_{2k+1}, t_{2k+2})$ for any $k \geq 0$,

$$\frac{1}{t_{2k+2} - t_{2k+1}} \int_{2k+1}^{2k+2} \mathbb{1}_{\{t : z_t \in NE[\epsilon]\}} \mathrm{d}t > 1 - \epsilon.$$

Moreover, by Corollary A.14, we can see from (A.38) and (A.37) that during (t_{2k}, t_{2k+3}) for any $k \geq 0$,

$$\frac{1}{t_{2k+3} - t_{2k}} \int_{2k}^{2k+3} \mathbb{1}_{\{t : z_t \notin NE[\epsilon]\}} \mathrm{d}t < \frac{3\hat{T}\epsilon}{t_{2k+2} - t_{2k+1}} < 3\epsilon.$$

So in the original $\mu^4/4$-dynamic $(x_t)_t$,

$$\frac{1}{T} \int_{\bar{t}}^{\bar{t}+T} \mathbb{1}_{\{t : x_t \in NE[\epsilon]\}} \mathrm{d}t \geq 1 - 3\epsilon$$

for all $T \geq 2\hat{T}$. We finally let $\mu(4\epsilon) := \mu^2/4$ and $T(4\epsilon) := \bar{t}/\epsilon + 2\hat{T}$. It follows that for any $\mu(4\epsilon)$-dynamic on Γ with any initial state

$$\frac{1}{T}\int_0^T \mathbb{1}_{\{t:x_t \in NE[4\epsilon]\}} \mathrm{d}t \geq 1 - 4\epsilon$$

for all $T \geq T(4\epsilon)$. We have thus shown that the induction hypothesis can also be applied to game Γ, and in fact to all generic extensive-form games of perfect information with $\eta+1$ decision nodes (without chance nodes). Q.E.D.

A.6 Proofs in Section 5

Proof of Theorem 5.1: We prove it by induction on the number of non-terminal nodes, with the following induction hypothesis. Denote $|N_d(G)|$ as the number of non-terminal nodes in the game G.

Induction hypothesis: Given a natural number $\eta \geq 1$ and any finite generic extensive-form game G of perfect information with $|N_d(G)| = \eta$, for every $\epsilon > 0$ there exists a number $\mu_G(\epsilon)$ with $0 < \mu_G(\epsilon) < \epsilon$ and a number $T_G(\epsilon) > 0$ such that for every interior $\mu_G(\epsilon)$-dynamic $(x_t)_t$ on G with any initial state x_0

$$x_t \in BIE[\epsilon] \tag{A.40}$$

for all $t \geq T_G(\epsilon)$. Note that this assumption trivially holds when $|N_d(G)| = 1$.

We prove Theorem 5.1 for the game Γ with $|N_d(\Gamma)| = \eta + 1 > 1$ under the assumption that the induction hypothesis above is true for all games G with $|N_d(G)| = \eta$. We pick a node $\tilde{n}$ in $L(\Gamma)$. (Recall the definition of $L(\Gamma)$ in (A.18).) If $\tilde{n}$ is a chance node, then by the generic assumption, the desired result trivially holds. Given a $\mu > 0$, we can see that there exists a time t_0 such that, for any interior μ-dynamic $(x_t)_t$, both (A.28) and (A.29) hold for all time $t \geq t_0$.

We obtain the game $\bar{\Gamma}$ by Tree Operation 2 in Section A.4, and a $\mu_{\bar{\Gamma}}(\epsilon/2) < \epsilon/2$. Take $\bar{\mu} := (2\mu_{\bar{\Gamma}}(\epsilon/2))^{1/2}$, and consider any $\mu < \bar{\mu}$. Note that the backward induction equilibrium path in $\bar{\Gamma}$ is included in the backward induction equilibrium path in Γ. By (A.40) in the induction hypothesis with respect to a $\mu^2/2$-dynamic in $\bar{\Gamma}$ and Lemma A.13, we can find $T_{\bar{\Gamma}}(\epsilon/2)$ such that for any $\mu^4/4$-dynamic $(x_t)_t$ in Γ

$$x_t \in BIE[\frac{\epsilon}{2} + \frac{\mu^2}{4}]$$

for all $t > T_{\bar{\Gamma}}(\epsilon/2) + t_0$. Note $\mu^2/4 < \mu_{\bar{\Gamma}}(\epsilon/2)/2 < \epsilon/4$, and we have completed the proof. Q. E. D.

Proof of Theorem 5.2: We prove it by induction on the number of non-terminal nodes, with the following induction hypothesis. Denote $|N_d(G)|$ as the number of non-terminal nodes in the game G.

Induction hypothesis: Given a natural number $\eta \geq 1$ and any finite generic extensive-form game G of perfect information with $|N_d(G)| = \eta$ and an $\epsilon' > 0$, there exists a number $T_G(\epsilon') > 0$ such that, for any initial state x_0 and any μ with $0 < \mu < \epsilon'$, there exists a μ-dynamic $(x_t)_t$ on G with the property

$$x_t \in BIE[\epsilon'] \tag{A.41}$$

for all $t > T_G(\epsilon')$. Note that the above assumption trivially holds when $|N_d(G)| = 1$.

We prove Theorem 5.2 for the game Γ with $|N_d(\Gamma)| = \eta + 1 > 1$ under the assumption that the induction hypothesis above is true for all games G with $|N_d(G)| = \eta$. We take the last non-terminal node in a backward induction equilibrium path as $\tilde{n}$, so $\tilde{n} \in L(\Gamma) \cap h_{BC}$, where h_{BC} denotes a backward induction equilibrium path. If $\tilde{n}$ is a chance node, then by the generic assumption, the desired result trivially holds.

Given a $\mu > 0$, we can see that there exists a time t_0 such that, for any $\mu/2$-dynamic $(x_t)_t$, (A.28) holds for all $t \geq t_0$.

We apply the function $\hat{q}$ defined at the start of Step 2 in the proof of Theorem 4.2, and let $\bar{x}_0 = q(x_{t_0})$. We obtain the game $\bar{\Gamma}$ by Tree Operation 2 with respect to $\tilde{n}$. We apply the induction hypothesis on $\bar{\Gamma}$ with respect to $\epsilon' = \epsilon/2$ and obtain $T_{\bar{\Gamma}}(\epsilon/2)$ and a $\mu/2$-dynamic $(\bar{x}_t)_t$ with initial state $\bar{x}_0$ and the property (A.41). For all players i and all $t \geq t_0$, we can simply let $g^i_{t+t_0} = \bar{g}^i_t$, and we find $x_{t+t_0} \in \bar{x}_t[\mu^4/4]$ and $g^i_{t+t_0} \in BR^i(x[\mu])$. So $(x_t)_{t \geq t_0}$ is a μ-dynamic. From (A.41), we then infer that

$$x_t \in BIE[\epsilon]$$

for all $t > t_0 + T_{\bar{\Gamma}}(\epsilon/2)$. Q. E. D.

References

J. Aubin and A. Cellina, Differential inclusions, Springer, Berlin, 1984.

M. Benaim and J. W. Weibull, Deterministic approximation of stochastic evolution in games, Econometrica, 71 (2003) 873–903.

R. Cressman, Evolutionary Dynamics and Extensive Form Games, MIT Press, Cambridge, 2003.

R. Cressman and K.H. Schlag, The dynamic (in)stability of backward induction, J. Econ. Theory 83 (1998) 260–285.

Y. Gilboa and A. Matsui, Social stability and equilibrium, Econometrica, 59 (1991) 859–867.

Z. Gorodeisky, Evolutionary stability for large populations and backward induction, Mathematics of Operations Research 31 (2006) 369–380.

S. Hart, Games in Extensive and Strategic Forms. In R.J. Aumann and S. Hart, editors, Handbook of Game Theory with Economic Applications, Elsevier, edition 1, volume 1, North Holland, 1992.

S. Hart, Evolutionary dynamics and backward induction, Games Econ. Behav. 41 (2002) 227–264.

S. Hart, Adaptive Heuristics, Econometrica, Vol. 73, No. 5, (2005) 1401-1430.

S. Hart and A. Mas-Colell, Regret-based continuous-time Dynamics, Games and Economic Behavior, 45 (2003), 375-394.

S. Hart and A. Mas-Colell, Uncoupled dynamics do not lead to Nash equilibrium, American Economic Review, 93 (2005) 1830–1836.

J. Hofbauer, Stability for the best response dynamics, mimeo, University of Vienna, 1995.

J. Hofbauer and W. H. Sandholm, On the global convergence of stochastic fictitious play, Econometrica, 70 (2002), 2265-2294.

J. Hofbauer and K. Sigmund, Evolutionary games and population dynamics, Cambridge University Press, 1998.

M. Kandori, G. Mailath and R. Rob, Learning, mutation, and long-run equilibrium in games, Econometrica, 61 (1993) 29–56.

M. Kandori, R. Rob, Evolution of Equilibria in the Long Run: A General Theory and Applications, Journal of Economics Theory, 65 (1995) 383–414.

D. Kreps and R. Wilson, Sequential Equilibria, Econometrica, 50 (1982) 863–894.

H. W. Kuhn, Extensive games and the problem of information. In H. W. Kuhn and A. W. Tucker, editors, Contributions to the Theory of Games II. Annals of Mathematics Studies, Vol. 28, Princeton University Press, 1953.

A. Matsui, Social stability and equilibrium, CMS-DMS No. 819 Northwestern University, 1989.

J. F. Nash, Non-Cooperative Games, PhD Dissertation at Princeton University, 1950.

G. Noldeke and L. Samuelson, An evolutionary analysis of backward and forward induction, Games and Economic Behavior, 5 (1993) 425–454.

K. Ritzberger, Foundations of Non-Cooperative Game Theory, Oxford University Press, 2002.

K. Ritzberger and J. Weibull, Evolutionary Selection in Normal-form Games, Econometrica, 63 (1995) 1371–1399.

W. H. Sandholm, Population Games and Evolutionary Dynamics, MIT Press, 2010.

Z. Xu, Stochastic stability in finite extensive-form games of perfect information, SSE/EFI Working Paper Series in Economics and Finance, No 743, March 2013.

H.P. Young, Learning by trial and error, Games and Economic Behavior, 65 (2009) 626–643.