CARESS Working Paper #95-08
Necessary and Su±cient Conditions for Convergence to Nash Equilibrium: The Almost Absolute Continuity Hypothesis.

Alvaro Sandroni[¤]
University of Pennsylvania
Department of Economics
3718 Locust Walk
Philadelphia PA 19104

April 25, 2000

Abstract

Kalai and Lehrer (93a, b) have shown that if players' beliefs about the future evolution of play is absolutely continuous with respect to play induced by optimal strategies then Bayesian updating eventually leads to Nash equilibrium. In this paper, we present the ¯rst set of necessary and su±cient conditions that ensure that Bayesian updating eventually leads to Nash equilibrium. More important, we show that absolute continuity does not rule out any observable behavior that is asymptotically consistent with Nash equilibrium.

Key words: Repeated Games, Bayesian Learning, Almost Absolute Continuity.

¤ I would like to express my gratitude to George Mailath for his constant encouraging support and for many valuable suggestions. I would also like to thank Yaw Nyarko and Andrew Postlewaite for useful comments. All errors are mine.

1. Introduction

Kalai and Lehrer (93a, b) have shown that if players' beliefs about the future evolution of play is absolutely continuous with respect to play induced by optimal strategies then players' predictions over play paths are eventually \accurate." Furthermore, outcomes induced by players' strategies are \close" to an almost Nash equilibrium play.[1] It is well known, however, that while absolute continuity is a su±cient condition for convergence to Nash equilibrium, it is not a necessary condition.

Let us ¯rst consider a simple example motivated by statistical inference. Suppose a coin is tossed repeatedly and independently. Nature selects the probability μ 2 [0; 1] of heads which is then ¯xed for each toss. If the player has a prior over μ that assigns positive probability to every neighborhood of μ then Bayesian updating eventually leads to precise predictions over outcomes. This is true even in the case that the player's prior assigns zero probability to μ in which case the player assigns probability zero to the event that the asymptotic frequency of heads is exactly μ: Note that, in fact, the true probability of this event is one. This is a simple example where Bayesian updating eventually leads to accurate predictions and absolute continuity does not hold.

Lehrer and Smorodinsky (94) have studied coordination assumptions over beliefs and best responses that are weaker than absolute continuity, but yet suf¯cient for convergence to Nash equilibrium. They consider a case where players may not assign positive probability to \the truth," but every player assign positive probability to \neighborhoods of the truth."[2] Their assumption is, of course, inspired by the example above. I return to their paper in the concluding section. In this paper, I consider the following questions: Is it possible to ¯nd coordination assumptions over beliefs and best responses that are necessary and su±cient for convergence to Nash equilibrium? To what extent is absolute continuity an unnecessarily strong assumption? Does absolute continuity rule out observable behavior that is asymptotically consistent with Nash equilibrium? When and why is it the case that absolute continuity can be relaxed but yet convergence to Nash equilibrium obtains?

In the example above, if players assign positive probability to every neighborhood of μ then a prior that assigns arbitrarily small but strictly positive probability to μ induces predictions over outcomes that are almost always (not only

[1] The absolute continuity assumption requires that if an event occurs with positive probability then all players also assign positive probability to this event.

[2] In this case, beliefs accommodate the truth.

in the limit) \arbitrarily close", to predictions under the original prior.[3] Clearly, the modi¯ed prior satis¯es absolute continuity. On the other hand, if there exists a modi¯ed prior that always induces predictions over outcomes that are similar to original prior's predictions and satis¯es absolute continuity, then Bayesian updating will eventually lead to accurate predictions because absolute continuity implies that modi¯ed beliefs' predictions will be eventually accurate and by assumption modi¯ed beliefs' predictions are always similar to original predictions.

Bayesian updating leads to accurate predictions if there exist prior beliefs that are absolutely continuous with respect to the truth and always induce predictions over outcomes that are similar to predictions under the original prior. In particular, priors can assign zero probability to μ and still lead to accurate predictions because these priors assign positive probability to neighborhoods of μ and therefore these priors always induces similar predictions to some prior that assigns strictly positive probability to μ: This suggests that predictions are eventually accurate if and only if there exists beliefs that satis¯es absolute continuity and induce predictions that are always similar to original beliefs' predictions. However, it is not clear, a priori, if this intuition is correct in more complex environments where both players' beliefs and the true play may follow a non-stationary stochastic process.

In the next section, I analyze the following example: Two players are engaged in an in¯nitely repeated coordination game. There is a sequence of beliefs and optimal strategies such that:

1. For each term in the sequence, outcomes induced by players' strategies are a Nash equilibrium play.

2. For each term in the sequence, players' beliefs over outcomes eventually converge, in the weak topology, to the true probability distribution induced by players' strategies.

3. For terms in the tail of the sequence, in almost all subgames players' beliefs over outcomes are arbitrarily close, in the weak topology, to the true probability distribution induced by players' strategies.

4. For each term in the sequence, both players assign zero probability to a set that has, in fact, full measure.

[3] This is not exactly correct. Later in the paper, we explain the precise meaning of \almost always."

This example also shows that although absolute continuity is a less demanding coordination assumption than Nash equilibrium, after arbitrarily small perturbations on beliefs, absolute continuity may no longer hold even if beliefs and optimal strategies were originally a Nash equilibrium. In particular, absolute continuity may not hold even if outcomes induced by players' strategies are a Nash equilibrium play and players' beliefs are \accurate."

Kalai and Lehrer (93a, c) observed that absolute continuity implies that players' predictions over outcomes are eventually accurate, even with respect to events that occurs in the distant future. That is, under absolute continuity, convergence occurs in the strong topology. If players' predictions over outcomes are eventually accurate, except with respect to events that occurs in the distant future, then players' predictions converge in the weak topology to the true probability distribution of outcomes induced by optimal strategies. In this case, I show that outcomes induced by players' optimal strategies are eventually close, in the weak topology, to an exact Nash equilibrium play. So, absolute continuity implies convergence in the strong topology, but convergence in the weak topology su±ces.

Absolute continuity is a necessary condition for convergence in the strong topology (see Kalai and Lehrer (93c) and below). That is, players' predictions over outcomes are eventually accurate, even with respect to events that occurs in the distant future, if and only if absolute continuity holds. However, given a player' belief it is possible to ¯nd modi¯ed belief that almost always induces predictions that are similar to original prediction in the short run and moreover, this modi¯ed belief assigns zero probability to events that are \very di®erent" from the original null sets. This is possible because some of the \nulls sets" may contain in¯nite play paths. So, absolute continuity may not hold with some belief, but may hold for some modi¯ed belief that induces short run predictions that are almost always similar to short run prediction under original belief. In this paper, I show that convergence to Nash equilibrium obtains if and only if there exists modi¯ed beliefs that satisfy absolute continuity and moreover, in \almost every subgame" this modi¯ed beliefs induce predictions over outcomes that are close, in the weak topology, to original players' predictions. Furthermore, in \almost every subgame" behavior strategies are an almost best response to modi¯ed beliefs.

The central concept introduced in this paper is almost absolute continuity. Players' beliefs are almost absolutely continuous with respect to optimal strategies if there exist modi¯ed beliefs such that:

1. In \almost every subgame," modi¯ed beliefs induce predictions over outcomes of play that are similar to original beliefs' predictions.

2. In “almost every subgame,” behavior strategies are an almost best responses with respect to modified beliefs.

3. Modified beliefs about the future evolution of play are absolutely continuous with respect to play induced by behavior strategies.

Beliefs and optimal strategies play eventually weakly like a Nash equilibrium if beliefs are eventually arbitrarily accurate, in the weak topology, and best responses eventually induce outcomes that are close, in the weak topology, to a Nash equilibrium play.

Players' beliefs and optimal strategies plays eventually weakly like a Nash equilibrium if and only if optimal strategies are almost absolutely continuous with respect to players' beliefs.[4] This result shows that any outcome path obtained where there is convergence to Nash equilibrium is also the outcome path of behavior that is almost optimal with respect to beliefs that satisfies absolute continuity; and moreover, these modified beliefs induces predictions over outcomes that are arbitrarily close to original players' predictions. Thus, absolute continuity does not rule out any observable behavior that is asymptotically consistent with Nash equilibrium.

2. The Model.

2.1. The Stage Game

The stage game is described by :

1. There exists n players.

2. Each player $i \in \{1, 2, ..., n\}$ has a finite set Σ_i of possible actions with $\Sigma = \prod_{i=1} \Sigma_i$ denoting the set of action combinations. $\Delta(\Sigma_i)$ denote the set of probability distributions on Σ_i.

3. Each player $i \in \{1, 2, ..., n\}$ has a payoff function $u_i : \Sigma \to \Re$.

2.2. The Infinitely Repeated Game

The infinite horizon game is described by :

[4] This result is obtained under an assumption on players' behavior. This assumption is described later in the paper.

1. For every natural number t, let Σ^t be the set of all histories of length $t \leq 1$. Let $H = \bigcup_t H_t$ be the set of all finite histories. For every finite history $h \in H_t$, a cylinder with base on h is the set $C(h) = \{w \in \Sigma^\infty / w = (h, \ldots)\}$ of all infinite histories such that the t initial elements coincides with h.
2. Let $\Im_t$ be the $\sigma-$algebra on Σ^∞ whose elements are all finite unions of cylinders with base on H_t. The $\sigma-$algebras $\Im_t$ define a filtration

$$\Im_0 \subset \ldots \Im_t \subset \ldots \subset \Im,$$

where $\Im_0$ is the trivial σ-algebra and $\Im$ is the σ-algebra generated by the algebra of finite histories $\Im^0 \equiv \bigcup_{t \geq 0} \Im_t$.
3. Each player $i \in \{1, 2, \ldots, n\}$ has a behavior strategy $f_i : H \to \Delta(\Sigma_i)$ that describes how player i randomizes among his possible actions conditional to every possible history. We also denote by $(f_i(h))(a^i)$ the probability that f_i prescribes for the action $a^i \in \Sigma_i$, after the finite history $h \in H$.
4. Given any strategy profile $g = (g_1, \ldots, g_n)$, there exists a probability measure μ_g (see Kalai and Lehrer (93a) for details) that represents the probability distribution over play paths generated by the strategy profile g.
5. Given a strategy profile g and a finite history $h \in H$, the induced strategy profile g_h is defined by $g_h(\bar{h}) = g(h, \bar{h})$ for any $\bar{h} \in H$. Analogously, given $w \in \Sigma^\infty$, let $w(t) \in \Sigma^t$ be the t initial element of w, and let $\left\{g_{w(t)}; t \in N\right\}$ be the sequence of behavior strategy profiles induced by w.
6. Each player $i \in \{1, 2, \ldots, n\}$ believes that his opponents will play strategies $f^i = \left(f_1^i, \ldots, f_n^i\right)$. We assume that each player knows his own strategy, i.e., $f_i = f_i^i$.
7. Let λ_i, $0 < \lambda_i < 1$, be player i's discount factor. Given player i's beliefs $f^i = \left(f_1^i, \ldots, f_n^i\right)$ player i's discounted expected payoff function is given by

$$V_i(f^i) = E_{\mu_{f^i}} \left\{ \sum_{t=1}^{\infty} (\lambda_i)^t . u_i \right\}.$$

We say that f_i is a best response to f^i_{-i} if for every player i's strategy l_i

$$V_i(f^i) - V_i(g^i) \geq 0$$

where $g^i = \left(f^i_1, \ldots, f^i_{i-1}, l_i, f^i_{i+1}, \ldots, f^i_n\right)$. Analogously, we say that f_i is an ε-best response to f^i_{-i} if the inequality above holds replacing 0 by $-\varepsilon$. We also say that a strategy profile $f = (f_1, \ldots, f_n)$ is optimal if f_i is a best response to f^i_{-i}, for every player i.

3. The Motivating Example.

The coordination game is described by the matrix

$$\begin{bmatrix} (I;II) & L & R \\ T & (2;2) & (0;0) \\ B & (0;0) & (1;1) \end{bmatrix}$$

Player I believes that player II will play left with probability θ and Player II believes that player I will play top with probability λ. Players' priors over these parameters have densities

$$v_I(\theta) = \frac{\theta^m}{\int_0^1 \theta^m \cdot d\theta} = (m+1)\cdot\theta^m;$$

and

$$v_{II}(\lambda) = \frac{\lambda^m}{\int_0^1 \lambda^m \cdot d\lambda} = (m+1)\cdot\lambda^m.$$

So,

$$\mu_{f_I}(\text{"left at period 1"}) = \int_0^1 \theta \cdot v_I(\theta) \cdot d\theta = \frac{m+1}{m+2}.$$

After observing left for $t-1$ periods, player I's posterior density over θ become,

$$\frac{\theta^m \cdot \theta^{t-1}}{\int_0^1 \theta^m \cdot \theta^{t-1} d\theta}.$$

So,

$$\mu_{f_I}(\text{"left at period } t\text{"} \mid \text{"left until period } t-1\text{"}) = \frac{\int_0^1 \theta \cdot \theta^m \cdot \theta^{t-1} d\theta}{\int_0^1 \theta^m \cdot \theta^{t-1} d\theta} = \frac{m+t}{m+t+1}.$$

Let $A \in \Im$ be the event "(top; left) forever". Then,

$$\mu_{f_I}(A) \leq \mu_{f_I}(\text{"left forever"}) =$$

$$\prod_{t=1}^{\infty} \mu_{f_I}(\text{“left at period } t\text{”} \mid \text{“left until period } t-1\text{”}) =$$

$$\prod_{t=1}^{\infty} \frac{m+t}{m+t+1} = \lim_{j \to \infty} \frac{m+1}{m+j} = 0.$$

So, for every $m \in \mathbb{N}$, $\mu_{f_I}(A) = 0$. Analogously, for every $m \in \mathbb{N}$, $\mu_{f_{II}}(A) = 0$.

I now show that, for every $m \in \mathbb{N}$, $\mu_f(A) = 1$.

In the first period, player I believes that player II will play left with probability $\frac{m+1}{m+2}$. Therefore, player I believes that player II will play left with probability greater than $\frac{1}{3}$ and so, player I optimally plays top, with probability one, in the first period. Analogously, player II optimally plays left, with probability one, in the first period. So, (top, left) is played, with probability one, in the first period.

Player I observes that player II played left in the first period. So, in the second period, player I believes that player II will play left with probability even greater than in the first period. So, player I optimally plays top, with probability one, in the second period. Analogously, player II optimally plays left, with probability one, in the second period. By induction (top, left) is played, with probability one, in all periods.

For all $m \geq 0$, the absolute continuity hypothesis does not hold true. In fact, μ_f and μ_{f_i} are disjoint. That is, both players assign zero probability to an event that has, in fact, full measure.

I now show that properties $1-3$ hold as claimed in the introduction.

Let $g = (g_I, g_{II})$ be a strategy profile such that both players play (top, left) regardless of past plays. g is a Nash equilibrium and so, the play path induced by players' behavior strategies, “(top, left) forever”, is a Nash equilibrium play.

However, for every $l \in \mathbb{N}$,

$$\mu_{f_i}(\text{“(top, left) from period } t \text{ to } t+l\text{”} \mid \text{“(top, left) until period } t-1\text{”}) =$$

$$\prod_{j=t}^{t+l} \frac{m+j}{m+j+1} = \frac{m+t}{m+t+l+1} \underset{t \to \infty}{\longrightarrow} 1 \quad i = I, II$$

and

$$\mu_f(\text{“(top, left) from period } t \text{ to } t+l\text{”} \mid \text{“(top, left) until period } t-1\text{”}) = 1 \ \forall t \in \mathbb{N}.$$

So, for all $m \in \mathbb{N}$, players' beliefs over outcomes eventually converge, in the weak topology, to the true probability distribution induced by behavior players' strategies.

Furthermore, for every $l \in N$; for every $t \in N$;

μ_{f^i}(“(top; left) from period t to $t+l$” | “(top; left) until period $t-1$”) =

$$\left(\frac{m+t}{m+t+l+1}\right) \to_{m \to \infty} 1 \quad i = I; II \text{ and}$$

μ_f(“(top; left) from period t to $t+l$” | “(top; left) until period $t-1$”) = 1.

Therefore, as m goes to infinity, in all subgames reached by (top; left) in every previous period; players' beliefs over outcomes are arbitrarily close, in the weak topology, to the true probability distribution induced by players' behavior strategies. Furthermore, “(top; left) forever” is a full measure event: So, with probability one, in all subgames players' beliefs over outcomes are arbitrarily close, in the weak topology, to the true probability distribution induced by players' behavior strategies, provided that m is large enough.

I now show that, for every $\varepsilon > 0$; it is possible to find an $\varepsilon-$perturbation of players' beliefs such that behavior strategies are absolutely continuous with respect to these modified beliefs.

For every $\varepsilon > 0$; let $l \in N$ be the period such that after observing (top; left) for l periods, players' beliefs over outcomes are $\varepsilon-$close, in the weak topology, to the true probability distribution induced by players' strategies. For $i = I; II$; let k^i be such that k^i coincide with f^i until period l, and k^i coincide with f after period l.

Consider the full measure event “(top; left) forever”. Let $h \in H$ be any finite history of the form “(top; left) until period $t-1$”. Then, by definition, k^i_h plays weakly $\varepsilon-$like f^i_h; $i = I; II$. Furthermore, $(f_i)_h$ is also a best response to $\left(k^i_{-i}\right)_h$ $i = I; II$; because $\left(k^I_{II}\right)_h$ still prescribes a probability greater than $\frac{1}{3}$ to “left”. So, “top” with probability one is still a best response in this case. Analogously, $(f_{II})_h$(“left”) = 1 and $\left(k^{II}_I\right)_h$(“top”) $\geq \frac{1}{3}$: So, $\left(k^I; k^{II}\right)$ are an $\varepsilon-$perturbation of $\left(f^I; f^{II}\right)$:

On the other hand, μ_{k^i}(“(top; left) forever”) =

μ_{k^i}(“(top; left) from period 1 to l”) =

$$\prod_{t=1}^{l} \mu_{k^i}(\text{“(top; left) at period } t \text{” | “(top; left) until period } t-1\text{”}) =$$

$$\prod_{t=1}^{l} \frac{m+t}{m+t+1} > 0 \quad i = I; II.$$

So, μ_f is absolutely continuous with respect to μ_{k^I} and to $\mu_{k^{II}}$:

4. Main Concepts and Results.

Definition 4.1 Let $\varepsilon > 0$ and let μ and $\tilde{\mu}$ be two probability measures defined on $\mathcal{F}$. The probability measure μ is $\varepsilon-$close to $\tilde{\mu}$ if

$$\|\mu - \tilde{\mu}\| = \sup_{A \in \mathcal{F}} |\mu(A) - \tilde{\mu}(A)| \leq \varepsilon.$$

The probability measure μ is weakly $\varepsilon-$close to $\tilde{\mu}$ if

$$d(\mu, \tilde{\mu}) = \sum_{k=1}^{\infty} 2^{-k} \cdot \left(\sup_{A \in \mathcal{F}_k} |\mu(A) - \tilde{\mu}(A)| \right) \leq \varepsilon.$$

The norm $\|\ \|$ induces the strong topology on the set of probability measures on Σ^{∞} while d is the metric of the weak topology.

Definition 4.2 Given two strategy profiles $f = (f_1, ..., f_n)$ and $g = (g_1, ..., g_n)$, we say that f plays (weakly) $\varepsilon-$like g if μ_f is (weakly) $\varepsilon-$close to μ_g.

If f plays $\varepsilon-$like g, then these two strategy profiles induce two probability measures on play paths that assign similar probabilities for all measurable events. However, if f plays weakly $\varepsilon-$like g, then these strategy profiles generate two probability measures that assign similar probabilities for all measurable events, except possibly the ones that may only be observed in the distant future.

Definition 4.3 A strategy profile $g = (g_1, ...g_n)$ is a (weak) subjective $\varepsilon-$equilibrium if there exists a matrix of strategies $(g^i_j)_{1 \leq i \leq n, 1 \leq j \leq n}$, with $g^i_i = g_i$ such that

(i) g_i is a best response to g^i_{-i}, $i = 1, ...n$, and

(ii) g plays (weakly) $\varepsilon-$like $g^i = (g^i_1, ...g^i_n)$, $i = 1, ...n$.

A strategy profile g is a (weak) subjective $\varepsilon-$equilibrium if players' predictions over outcomes are $\varepsilon-$close, in the (weak) strong topology, to the true probability distribution of play paths, induced by players' optimal strategies. Clearly, if $\varepsilon > 0$, then a subjective $\varepsilon-$equilibrium is a weak subjective $\varepsilon-$equilibrium, but not conversely. However, there is no difference between a subjective 0 - equilibrium and a weak subjective 0 - equilibrium.

Any Nash equilibrium is a subjective 0 - equilibrium but not conversely. The difference is that a subjective $0-$equilibrium does not require that players' beliefs and strategies coincide off the play path.

Definition 4.4 Beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies $f = (f_1, \ldots, f_n)$ play eventually (weakly) $\varepsilon-$like a Nash equilibrium if there exists a set $\Omega \in \mathcal{F}$ such that

1. $\mu_f(\Omega) = 1$
2. For every $w \in \Omega$, for every $\varepsilon > 0$, there exists a period $t(w, \varepsilon)$ such that for all $t \geq t(w, \varepsilon)$, $f_{w(t)}$ and $f^i_{w(t)}$, $i = 1, \ldots, n$, plays (weakly) $\varepsilon-$like the same Nash equilibrium.

Beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies $f = (f_1, \ldots, f_n)$ plays eventually (weakly) like a Nash equilibrium if, for every $\varepsilon > 0$, f and f^i, $i = 1, \ldots, n$, plays eventually (weakly) $\varepsilon-$like the same Nash equilibrium.

That is, beliefs and optimal strategies plays eventually (weakly) $\varepsilon-$like a Nash equilibrium if, in finite time, beliefs and best responses play (weakly) $\varepsilon-$like the same Nash equilibrium. The same definition apply if "Nash equilibrium" is replaced by "$\varepsilon-$Nash equilibrium" or "(weak) subjective $\varepsilon-$equilibrium".

Definition 4.5 Given two strategy profiles f and g, f is absolutely continuous with respect to g if μ_f is absolutely continuous with respect to μ_g, i.e., for every $A \in \mathcal{F}$, $\mu_g(A) = 0$ imply $\mu_f(A) = 0$.

In particular, players' best responses are absolutely continuous with respect to players' beliefs if any event in the $\sigma-$algebra $\mathcal{F}$ that occurs with strictly positive probability is assigned strictly positive probability by all players.

I assume that the optimal strategies f are absolutely continuous with respect to the players' beliefs f^i on the algebra $\mathcal{F}^0$. That is,

$$\mu_{f^i}(A) = 0 \Rightarrow \mu_f(A) = 0 \quad \forall A \in \mathcal{F}^0, \; i = 1, \ldots, n.$$

This assumption is a necessary condition for players to update their beliefs by Bayes' rule. Notice, however, that absolute continuity on the algebra $\mathcal{F}^0$ is a weaker condition than the absolute continuity on the $\sigma-$algebra $\mathcal{F}$. Moreover, the mere fact that players are able to revise their beliefs by Bayes' rule does not necessarily imply convergence to Nash equilibrium. I now restate the main results of Kalai and Lehrer (93a, b, c).

Proposition 4.1 Beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies f plays eventually $\varepsilon-$like a subjective $\varepsilon-$equilibrium, for all $\varepsilon > 0$, if and only if f is absolutely continuous with respect to f^i $i = 1, \ldots, n$.

Proof - The "if" part is a direct consequence of the Blackwell-Dubins theorem. For the converse, see Kalai and Lehrer (93c). In the appendix, we give an alternative proof for the converse.

Kalai and Lehrer (93b) have also shown that for every $\varepsilon > 0$, there is $\eta > 0$ such that if g is a subjective $\eta-$equilibrium, then g plays $\varepsilon-$like an $\varepsilon-$Nash equilibrium. These propositions imply the main result of Kalai and Lehrer (93a, c).

Proposition 4.2 Beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies f plays eventually $\varepsilon-$like a $\varepsilon-$Nash equilibrium, for all $\varepsilon > 0$, if and only if f is absolutely continuous with respect to f^i, $i = 1, \ldots, n$.

That is, proposition 4.2 shows that absolute continuity is a sufficient and necessary condition for convergence to an almost Nash equilibrium play, in the strong topology.

I now show that absolute continuity is "robust to perturbations" in the strong topology.

Proposition 4.3 Let f and f^i, $i = 1, \ldots, n$, be optimal strategies and beliefs of the players. Consider a sequence of strategy profiles, $g^i(m)$, such that $g^i(m)$ plays $\eta^i(m)-$like f^i, and $\eta^i(m) \xrightarrow[m \to \infty]{} 0$, $i = 1, \ldots, n$.

If f is absolutely continuous with respect to f^i, then there exists $m(\varepsilon)$ such that for all $m \geq m(\varepsilon)$, there exists a set $B(m) \in \mathcal{F}$ such that

1. $\mu_f(B(m)) \geq 1 - \varepsilon$;
2. if $A \subset B(m)$, $A \in \mathcal{F}$, then $\mu_{g^i(m)}(A) = 0 \Rightarrow \mu_f(A) = 0$; and
3. $\forall w \in B(m)$, $\left\| \left(\mu_{g^i(m)}\right)_{w(t)} - \left(\mu_f\right)_{w(t)} \right\| \xrightarrow[t \to \infty]{} 0$.

Proof - See Appendix.

In particular, proposition 4.3 shows that if players' optimal strategies f are absolutely continuous with respect to players' beliefs, f^i; then any "modified" beliefs, $g^i(m)$; that plays almost like f^i; eventually plays almost like f. Proposition 3.3 is false if we just assume that $g^i(m)$ plays almost weakly like f^i. For example, in the coordination game given before, players' beliefs and optimal strategies plays almost weak like the same Nash equilibrium. However, players assign zero probability to a set that has, in fact, full measure.

This makes clear the differences between "modifying a player belief" in the weak and strong topology. If a player's belief is modified, but predictions over outcomes remain close in the strong topology to original predictions, then these two probability distributions assign zero measure to the same sets that lie inside a "large" set (a set that has high probability with respect to both probability distributions). In this case, absolute continuity can be "preserved." However, if only short run predictions over outcomes remain close to original short run predictions then absolute continuity may not be "preserved." Therefore, absolute continuity cannot be a necessary condition for convergence to Nash equilibrium, because whenever absolute continuity holds players' beliefs can be modified in such a way that absolute continuity is no longer satisfied but convergence to Nash equilibrium still holds.

Definition 4.6 Let f and f^i $i = 1, ..., n$ be the optimal strategies and the beliefs of the players. The strategy profile $k^i = (k^i_1, ..., k^i_{i-1}, f_i, k^i_{i+1}, ..., k^i_n)$ is an $\varepsilon -$ perturbation of player i's beliefs if there exists a set $A \in \Im$ such that

1. $\mu_f(A) \geq 1 - \varepsilon$;
2. $\forall w \in A, \forall t \in N, (f_i)_{w(t)}$ is an $\varepsilon -$best response to $\left(k^i_{-i}\right)_{w(t)}$; and
3. $\forall w \in A, \forall t \in N, k^i_{w(t)}$ plays weakly $\varepsilon -$like $f^i_{w(t)}$.

The strategy profile k^i is an $\varepsilon -$perturbation of player i's beliefs if, in all subgames, except possibly in a set of play paths that has probability less than ε; k^i always plays weakly $\varepsilon -$like player i's beliefs, and behavior strategy f is always an $\varepsilon -$best response to k^i.

The notion of "perturbation" is not an asymptotic notion.[5] A strategy profile is an $\varepsilon -$ perturbation of player i's beliefs if with probability $1 - \varepsilon$; "modified beliefs" always induce similar predictions over the future evolution of the play and

[5] This is in contrast with the notion of neighborhood in Lehrer and Somorodinsky (94).

behavior strategies are always an $\varepsilon-$best response to "modified beliefs". This is an important restriction because if observed outcomes induced by behavior strategies can be justified by players' beliefs, then observed outcomes can also be justified by modified beliefs.

Of course, perturbations on beliefs should not perturb beliefs that a player has over his or her own behavior strategy.

I now define the central notion of almost absolute continuity.

Definition 4.7 Let f and $f^1, \ldots, f^n$ be the optimal strategies and the beliefs of the players. The profile f is $\varepsilon-$absolutely continuous with respect to f^i if there exists an $\varepsilon-$ perturbation of players' beliefs, k^i, such that f is absolute continuous with respect to k^i.

The profile f is almost absolutely continuous with respect to f^i if, for every $\varepsilon > 0$, f is $\varepsilon-$absolutely continuous with respect to f^i.

That is, f is almost absolutely continuous with respect to f^i if, for every $\varepsilon > 0$, there exists an $\varepsilon-$ perturbation of f^i, k^i, such that f is absolutely continuous with respect to k^i. Clearly, if f is absolutely continuous with respect to f^i then f is almost absolutely continuous with respect to f^i.

I assume

R) There exists some $\sigma > 0$ such that

$$\forall h \in H, \forall a^i \in \Sigma_i \text{ If } (f_i(h))(a^i) \neq \{0,1\} \text{ then } (f_i(h))(a^i) \geq \sigma.$$

R) is an assumption on players' behavior. It requires that if, at a certain period, a player decides to randomize over some pure strategies, then he or she will not assign an arbitrarily small probability to any of the pure strategies choices. I do not know if this assumption can be dispensed.

Proposition 4.4 Let f and $\left(f^1, \ldots, f^n\right)$ be the optimal strategies and the beliefs of the players. For every $\varepsilon > 0$, there is $\delta > 0$ such that for every $\eta \leq \delta$ if f is $\eta-$absolutely continuous with respect to f^i $i = 1, \ldots, n$ then f and f^i $i = 1, \ldots, n$ plays eventually weakly $\varepsilon-$like a weak subjective $\varepsilon-$equilibrium.

On the other hand, under assumption R), for every $\varepsilon > 0$, there is $\delta > 0$ such that for every $\eta \leq \delta$ if f^i $i = 1, \ldots, n$ and f plays eventually weakly $\eta-$like a weak subjective $\eta-$equilibrium, then f is $\varepsilon-$absolutely continuous with respect to f^i $i = 1, \ldots, n$.

Proof - See Appendix.

Proposition 4.4 shows that if almost absolute continuity holds then players eventually make accurate predictions over play paths. On the other hand, if players' prediction over play paths are eventually accurate then almost absolute continuity holds.

If absolute continuity holds then players' beliefs are eventually accurate. Therefore, all modified beliefs that induce predictions over play paths that are similar to original beliefs are also eventually accurate. However, absolute continuity does not hold for all these beliefs. On the other hand, by definition, if almost absolute continuity holds for some beliefs, then it holds for all modified beliefs that induces predictions over play paths that are similar to original beliefs.

Proposition 4.4 characterizes the distinction between the case where players' predictions over play paths are eventually accurate and the case where players' predictions over play paths are not eventually accurate. In the first case there exists modified beliefs that satisfies absolute continuity; and moreover, with the possible exception of play paths that have small probability, these modified beliefs always induces similar predictions (on future evolution of play) to original beliefs. Furthermore, if modified beliefs induces similar predictions (over future evolution of play) to original beliefs, then behavior strategies must be almost optimal with respect to modified beliefs. So, with the possible exception of play paths that have small probability, behavior strategies must be almost optimal with respect to modified beliefs.

Proposition 4.5 For every $\varepsilon > 0$; there is $\hat{\eta} > 0$ such that for every $\eta \leq \hat{\eta}$; if g is a weak subjective $\eta-$equilibrium, then g plays weakly $\varepsilon-$like an Nash equilibrium.

Proof - See Appendix.

Proposition 4.5 shows that if players make accurate predictions over play paths, then outcomes induced by players' optimal strategies are close, in the weak topology, to an exact Nash equilibrium play.

These two propositions imply our main result.

Proposition 4.6 Under assumption R), beliefs $\left(f^1; \ldots; f^n\right)$ and optimal strategies $f = (f_1; \ldots; f_n)$ plays eventually weakly like a Nash equilibrium if and only if f is almost absolutely continuous with respect to f^i; $i = 1; \ldots; n$:

Proof - See Appendix.

Proposition 4.6 shows that almost absolute continuity is necessary and su±cient for convergence to Nash equilibrium. More important, consider the case that players' beliefs and optimal strategies eventually induces probability distributions over outcomes that resembles a Nash equilibrium play. If absolute continuity is not satis¯ed then behavior strategies can also be \justi¯ed" by some modi¯ed beliefs that satis¯es absolute continuity and with the possibly exception of play paths that have small probability, these modi¯ed beliefs always induces predictions over outcomes that are arbitrarily close, in the weak topology, to players' original predictions. On the other hand, if players' beliefs and optimal strategies satis¯es absolute continuity or almost absolute continuity then, by de¯nition, any modi¯ed beliefs over play paths that induces similar prediction to players' original predictions also satis¯es almost absolute continuity. So, convergence to Nash equilibrium obtains with respect to any of these beliefs.

Remark: Outcomes induced by exact best responses to \modi¯ed beliefs", are not necessarily \close" to outcomes induced by best responses to \original beliefs".

Take, for example, two players playing an in¯nitely repeated \matching pennies". Assume that both players believe that his or her opponent is randomizing among all pure strategies with equal probability, regardless of past outcomes. In this case, any strategy is a best response. Assume that both players adopts a behavior strategy such that posteriors of the probability measures induced by players' beliefs and optimal strategies merge in the weak, but not in the strong, topology. In this case, by proposition 3.1, the absolutely continuity assumption does not hold. Take any perturbation of original beliefs that are absolutely continuous with respect to behavior strategies and consider an exact best response to these modi¯ed beliefs. According to these modi¯ed beliefs, at some period, a player believes that his or her opponent is not randomizing with equal probability. At those periods, this player will choose a pure strategy, with probability one. Therefore, outcomes induced by such strategies can not be close, in the weak topology, to outcomes induced by best responses to \original beliefs".

5. Conclusion

Lehrer and Smorodinsky (94), hereafter LS, obtained coordination assumptions on beliefs and best responses that are weaker than absolute continuity but ensure convergence to Nash equilibrium. Although independently obtained, the results in this paper and the results in Lehrer and Smorodinsky's (94) paper are complementary. LS focus on behavior strategies' perturbations while I only consider perturbations on beliefs. Therefore, there are substantial di®erences in the assumptions made in this paper and the conditions used by LS. Moreover, thecniques used by LS and the ones in this paper are completely di®erent. LS show that if beliefs accommodate the truth, i.e., if beliefs assign positive probability to \neighborhoods of the truth" then convergence to Nash equilibrium obtains. A direct consequence of the results in this paper and the ones in LS is that if beliefs accommodate the truth then convergence to Nash equilibrium obtains because there exists modi¯ed belief that satis¯es absolute continuity and moreover, in \almost every subgame" this modi¯ed belief induces similar short run predictions to original predictions.

I also show that any outcome path obtained where there is convergence to Nash equilibrium is also the outcome path of behavior that is almost optimal with respect to beliefs that satis¯es absolute continuity; and moreover, these beliefs are arbitrarily close to original beliefs. Thus, absolute continuity does not rule out any observable behavior that is asymptotically consistent with Nash equilibrium.

Appendix

Proof of Proposition 4.1's Converse Let -(") 2 = be the set such that beliefs play " ¡ like optimal strategies in ¯nite time. Let - be $\bigcap_{n \geq 1} -(\frac{1}{n})$: Clearly, ¹f(-) = 1; and the posteriors of ¹f and ¹fi i = 1; :::; n; converge in the strong topology on -:

Let A 2 = be any set such that ¹f(A) > 0: Let ¸ be a probability measure de¯ned by

$$¸(B) = \frac{¹_f(A \setminus B)}{¹_f(A)} \quad 8B\ 2 =$$

Clearly ¸ is absolutely continuous with respect to ¹f: By the Blackwell-Dubins theorem, there exists a set C 2 = such that ¸(C) = 1; and the posteriors of ¹f and ¸ converge in the strong topology on C:

By de¯nition, ¹f(A \ C) = ¹f(A) > 0: So, ¹f(C) > 0 and C \ - 6= ;: The posteriors of ¹fi; ¹f; and ¸ converge, in the strong topology, on C \ -:

Also by de¯nition, ¸(A) = 1. So, for every observation h 2 §t; the posteriors of ¸ are such that ¸h(A) = 1: Assume by contradiction that ¹fi(A) = 0: Then, for every observation $h \in \Sigma^t$; $\left(\mu_{f^i}\right)_h(A) = 0$. Therefore, the posteriors of μ_{f^i} and λ can not converge in the strong topology on $C \cap \Omega$. A contradiction.

q.e.d.

Proof of Proposition 4.3 Sandroni (94), proposition 3.4 pg 13, has shown that

$$\forall \text{ sequences } \{B_n\} \in \mathcal{F}^0;\ \mu_{f^i}(B_n) \underset{n\to\infty}{\to} 0 \Rightarrow \mu_f(B_n) \underset{n\to\infty}{\to} 0$$

Assume that for some sequence $\{A_n\} \in \mathcal{F}$;

$$\mu_{f^i}(A_n) \underset{n\to\infty}{\to} 0.$$

By the Caratheodory extension theorem, there exists a sequence of sets $\{B_n\} \in \mathcal{F}^0$ such that

$$\left|\mu_{f^i}(A_n) - \mu_{f^i}(B_n)\right| \leq \frac{1}{m} \text{ and } \left|\mu_f(A_n) - \mu_f(B_n)\right| \leq \frac{1}{m}$$

So, $\mu_{f^i}(B_n) \underset{n\to\infty}{\to} 0$ and therefore, $\mu_f(B_n) \underset{n\to\infty}{\to} 0$. So,

$$\mu_f(A_n) \underset{n\to\infty}{\to} 0.$$

Therefore,

$$\forall \text{ sequences } \{A_n\} \in \mathcal{F};\ \mu_{f^i}(A_n) \underset{n\to\infty}{\to} 0 \Rightarrow \mu_f(A_n) \underset{n\to\infty}{\to} 0$$

Let $\delta > 0$, $\delta(m)$ be any sequence of strictly positive numbers such that

$$\sum_{m=1}^{\infty} \delta(m) + \delta = 1.$$

Let $\lambda = \sum_{m=1}^{\infty} \delta(m).\mu_{g^i(m)} + \delta.\mu_{f^i}$ be a probability measure. For every m; $\mu_{g^i(m)}$ and μ_{f^i} are absolutely continuous with respect to λ. Therefore, by the Radon-Nykodym theorem, there exists functions $\varphi(m)$ and φ such that

$$\mu_{g(m)} = \varphi(m).\lambda,\ \mu_{f^i} = \varphi.\lambda$$

Let $C(m) \in \mathcal{F}$ be the set $\{\varphi(m) > 0; \varphi > 0\}$.

Clearly, if $A \in \mathcal{F}$ is a set such that $A \subset C(m)$ then

$$\mu_{g^i(m)}(A) = 0 \;\Leftrightarrow\; \lambda(A) = 0 \;\Leftrightarrow\; \mu_{f^i}(A) = 0 \;\Rightarrow\; \mu_f(A) = 0$$

Let $\eta(m)$ be the function $T.\min\{\phi(m), \phi\}$ and let $\nu(m) = \eta(m).\lambda$ be a positive measure, where $T \in \Re_+$ is such that $\nu(m)$ is also a probability measure. For every m; $\nu(m)$ is absolutely continuous with respect to $\mu_{g^i(m)}$ and μ_{f^i}; and μ_{f^i} is absolutely continuous with respect to μ_f. Therefore, by the Blackwell-Dubins theorem, there exists a sets $D(m) \in \mathcal{F}$ and $\Omega \in \mathcal{F}$ such that $\nu(m)(D(m)) = 1$ and $\mu_f(\Omega) = 1$ such that

$$\forall w \in D(m) \cap \Omega;\; \left\| \left(\mu_f\right)_{w(t)} - \left(\mu_{g^i(m)}\right)_{w(t)} \right\| \xrightarrow[t \to \infty]{} 0.$$

$C(m)^c = \{\phi(m) = 0\} \cup \{\phi = 0\}$. So, $\mu_{f^i}(C(m)^c) \xrightarrow[m \to \infty]{} 0$.

However, $\nu(m)(D(m)^c) = 0$. So, $D(m)^c \subset \{\eta(m) = 0\} \cup F = C(m)^c \cup F$, where $F \in \mathcal{F}$ is a set such that $\lambda(F) = 0$.

Clearly, $\mu_{f^i}(F) = 0$. Therefore, $\mu_{f^i}(D(m)^c) \leq \mu_{f^i}(C(m)^c)$. So,

$$\mu_{f^i}(C(m) \cap D(m)) \xrightarrow[m \to \infty]{} 1.$$

Let $B(m)$ be $C(m) \cap D(m) \cap \Omega$. Then,

$$\mu_f(B(m)) \xrightarrow[m \to \infty]{} 1;$$

and

$$\forall A \subset B(m);\; \mu_{g^i(m)}(A) = 0 \Rightarrow \mu_f(A) = 0;$$

and the posteriors of μ_f and $\mu_{g^i(m)}$ converge, in the sup norm, on $B(m)$.

q.e.d.

Lemma A.1. Let f and $f^1, \ldots, f^n$ be the optimal strategies and the beliefs of the players. Let $k^i = (k^i_1, \ldots, f_i, \ldots, k^i_n)$ be a strategy profile such that

1. k^i plays weakly ε-like f^i
2. $\left(k^i\right)_h = \left(f^i\right)_h$ if $\mu_f(C(h)) = 0$; $h \in H$.
3. $\left(k^i_j\right)_h (a^j) = \left(f^i_j\right)_h (a^j)$ if $\mu_f(C(h; a)) = 0$; $h \in H$; $a = (a^1, \ldots, a^n) \in \Sigma$.

Then, given assumption R), for every $\varepsilon > 0$ there exists $\delta > 0$, such that if $\eta \leq \delta$, then f_i is an $\varepsilon-$best response to k^i_{-i}.

Proof- Assume by contradiction that there exists an $\varepsilon_0 > 0$, and a sequence of strategy profiles $k^i(m) = (k^i_1(m), \ldots, f_i(m), \ldots, k^i_n(m))$ such that

1. $f_i(m)$ is a best response to $f^i_{-i}(m)$;
2. $d\left(\mu_{k^i(m)}, \mu_{f^i(m)}\right) \underset{m\to\infty}{\longrightarrow} 0$;
3. $\left[k^i(m)\right]_h = \left[f^i(m)\right]_h$ if $\mu_{f(m)}(C(h)) = 0$, $h \in H$;
4. $\left(k^i_j(m)\right)_h(a^j) = \left(f^i_j(m)\right)_h(a^j)$ if $\mu_{f(m)}(C(h,a)) = 0$, $h \in H$, $a = (a^1, \ldots, a^n) \in \Sigma$;
5. $f_i(m)$ is not an ε_0-best response to $k^i_{-i}(m)$.

By definition, there exists a behavior strategy $l(m)$ such that

$$V_i(b(m)) = E_{\mu_{b(m)}}\left\{\sum_{t=1}^{\infty}(\lambda_i)^t . u_i\right\} \geq V_i(k^i(m)) + \varepsilon_0$$

where $b(m) = (k^i_1(m), \ldots, l(m), \ldots, k^i_n(m))$.

Also by definition,

$$V_i(f^i(m)) \geq V_i(c(m))$$

where $c(m) = (f^i_1(m), \ldots, l(m), \ldots, f^i_n(m))$.

By the Banach-Alaoglu theorem, there exists probability measures $\mu_b, \mu_c, \mu_{f^i}, \mu_{k^i}$ and a subsequence, also indexed by m, such that

$$d\left(\mu_{z(m)}, \mu_z\right) \underset{m\to\infty}{\longrightarrow} 0 \text{ where } z = b, c, f^i, k^i.$$

Clearly, $\mu_{k^i} = \mu_{f^i}$ because $\mu_{k^i(m)}$ is arbitrarily close to $\mu_{f^i(m)}$, in the weak topology.

We want to show that $\mu_b = \mu_c$.

Assume, by contradiction, that there exists some $h \in H$, $h = (h_0, \ldots, h_r)$, such that

$$\mu_b(C(h)) \neq \mu_c(C(h)).$$

Then, for some $0 \leq s \leq r - 1$,

$$(\mu_b)_{(h_0,\ldots,h_s)}(C(h_0, \ldots, h_{s+1})) \neq (\mu_c)_{(h_0,\ldots,h_s)}(C(h_0, \ldots, h_{s+1})).$$

But, $h_{s+1} = (h^1_{s+1}, \ldots, h^n_{s+1})$, and

$$(\mu_b)_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})) = \lim_{m\to\infty} \prod_{j\neq i, j=1} \left(k^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \cdot (l(m))_{(h_0,\ldots,h_s)}(h^i_{s+1})$$

$$(\mu_c)_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})) = \lim_{m\to\infty} \prod_{j\neq i, j=1} \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \cdot (l(m))_{(h_0,\ldots,h_s)}(h^i_{s+1}).$$

So, there exists a subsequence, also indexed by m, such that

$$\lim_{m\to\infty} \prod_{j\neq i, j=1} \left(k^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \neq \lim_{m\to\infty} \prod_{j\neq i, j=1} \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}).$$

Furthermore, it has to be the case that for some $\bar{m}$

$$\mu_{f(m)}(C(h_0,\ldots,h_s)) \neq 0 \text{ and } (f_i(m))_{(h_0,\ldots,h_s)}(h^i_{s+1}) \neq 0 \text{ for all } m \geq \bar{m}.$$

Otherwise,

$$\left(k^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) = \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \text{ infinitely often.}$$

By assumption R), there exists $\sigma > 0$ such that

$$(f_i(m))_{(h_0,\ldots,h_s)}(h^i_{s+1}) > \sigma \text{ for all } m \geq \bar{m}.$$

However,

$$(\mu_{k^i})_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})) = \lim_{m\to\infty} \prod_{j\neq i, j=1} \left(k^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \cdot \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}).$$

$$\left(\mu_{f^i}\right)_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})) = \lim_{m\to\infty} \prod_{j\neq i, j=1} \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}) \cdot \left(f^i_j(m)\right)_{(h_0,\ldots,h_s)}(h^j_{s+1}).$$

So,

$$(\mu_{k^i})_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})) \neq \left(\mu_{f^i}\right)_{(h_0,\ldots,h_s)}(C(h_0,\ldots,h_{s+1})).$$

A contradiction. Therefore,

$$\mu_b = \mu_c.$$

But,

$$\lim_{m\to\infty} V_i(b(m)) = V_i(b) = V_i(c) = \lim_{m\to\infty} V_i(c^i(m)) \leq \lim_{m\to\infty} V_i(f^i(m)) = V_i(f^i)$$

and

$$V_i(f^i) = V_i(k^i) = \lim_{m\to\infty} V_i(k^i(m)) \leq \lim_{m\to\infty} V_i(b(m)) - \varepsilon_0 = V_i(b) - \varepsilon_0.$$

Therefore,

$$V_i(b) \leq V_i(b) - \varepsilon_0. \text{ A contradiction.}$$

q.e.d

Proof of Proposition 4.4 $\Rightarrow$) Assume that f is an $\frac{1}{m}-$absolutely continuous with respect to f^i, $m \in N$.

By definition, there exists sets $A(m) \in \mathcal{F}$ and $\Omega(m) \in \mathcal{F}$, and strategy profile $k^i(m)$ such that

$$\mu_f(A(m)) \geq 1 - \frac{1}{m}, \ \mu_f(\Omega(m)) = 1, \text{ and}$$

the posteriors of μ_f and $\mu_{k^i(m)}$ merge, in the strong topology on $\Omega(m)$; and

$$\forall w \in A(m), \ \forall t \in N, \ \left(k^i(m)\right)_{w(t)} \text{ plays weakly } \frac{1}{m} - \text{like } \left(f^i\right)_{w(t)}.$$

Let δ be $\frac{1}{\hat{m}} \leq \frac{\varepsilon}{2}$.

Let $A \in \mathcal{F}$ be the set $\bigcap_{j=1}^{\infty} \bigcup_{m \geq j} A(m)$, and let $\Omega \in \mathcal{F}$ be the set $\bigcap_{m=1}^{\infty} \Omega(m)$. Clearly,

$$\mu_f(A \cap \Omega) = 1.$$

If $w \in A \cap \Omega$ then, $w \in \bigcup_{m \geq \hat{m}} A(m)$. So, $w \in A(\hat{m}) \cap \Omega(\hat{m})$ for some $\hat{m} \geq \hat{m}$.

Therefore, for all $\eta \leq \delta$, if f is an $\eta-$absolutely continuous with respect to f^i and $w \in A \cap \Omega$, then there exists a period $\hat{t}$ such that

$$\left(k^i(m)\right)_{w(t)} \text{ plays } \frac{\varepsilon}{2} - \text{like } f_{w(t)}, \text{ for all } t \geq \hat{t}.$$

and

$$\left(k^i(m)\right)_{w(t)} \text{ plays weakly } \frac{\varepsilon}{2} - \text{like } f^i_{w(t)}, \text{ for all } t \geq \hat{t}.$$

Therefore, for every $w \in A \cap \Omega$, there exists a time $\bar{t}$ such that, if $t \geq \bar{t}$ then

$$\left(f^i(m)\right)_{w(t)} \text{ plays weakly } \varepsilon - \text{ like } f_{w(t)}.$$

( ) For every $\eta > 0$, let $T^j_\eta : \Sigma^\infty \to N \cup \{\infty\}$ be a function such that $T^j_\eta(w) = j$ if

$$d\left(\left(\mu_{f^i}\right)_{w(t)}, (\mu_f)_{w(t)}\right) \leq \eta \quad \forall t \geq j,$$

and $T^j_\eta(w) = \infty$ if

$$d\left(\left(\mu_{f^i}\right)_{w(t)}, (\mu_f)_{w(t)}\right) > \eta \text{ infinitely often.}$$

Let $0 < \hat{\eta} \leq \varepsilon$ be as in lemma A.1. By assumption, $\mu_f(T^j_{\hat{\eta}} < \infty) = 1$. So, there exists $\bar{t} \in N$ such that

$$\mu_f\left(\left\{T^j_{\hat{\eta}} < \bar{t}\right\}\right) \geq 1 - \varepsilon.$$

Let k^i be a strategy profile such that if $t \leq \bar{t}$, then

$$k^i(h) = f^i(h) \;\forall h \in \Sigma^t,$$

and if $t > \bar{t}$, then

$$\left(k^i\right)_h = \left(f^i\right)_h \text{ if } \mu_f(C(h)) = 0;\ h \in \Sigma^t;$$

$$\left(k^i_j\right)_h(a^j) = \left(f^i_j\right)_h(a^j) \text{ if } \mu_f(C(h,a)) = 0;\ j = 1,\ldots,n; h \in \Sigma^t;\ a = (a^1,\ldots,a^n) \in \Sigma;$$

$$k^i(h) = f(h) \text{ otherwise.}$$

Clearly, with probability one according to μ_f, the posteriors of μ_f and μ_{k^i} coincide after period $\bar{t}$. So, by the converse of proposition 4.1, f is absolutely continuous with respect to k^i.

Let $A \in \mathcal{F}$ be the set

$$\left\{T^j_{\hat{\eta}} < \bar{t}\right\} \cap \left\{\bigcup_{\mu_f(C(h))=0} C(h)\right\}^c.$$

Clearly,

$$\mu_f(A) \geq 1 - \varepsilon.$$

Furthermore, by definition, if $w \in A$ then,

$$\left(k^i\right)_{w(t)} \text{ plays weakly } \varepsilon - \text{ like } \left(f^i\right)_{w(t)} \quad \forall t \in N.$$

By lemma A.1, $\forall w \in A; \forall t \in N$; $(f_i)_{w(t)}$ is an $\varepsilon-$best response to $\left(k^i_{-i}\right)_{w(t)}$. So, k^i is an $\varepsilon-$perturbation of f^i.

q.e.d.

Proof of Proposition 4.5 Let $\mu(m) \xrightarrow[m\to\infty]{} \mu$ denote a sequence of probability measures such that

$$d(\mu(m); \mu) \xrightarrow[m\to\infty]{} 0.$$

Suppose by contradiction that there exists an $\varepsilon_0 > 0$ and a sequence of strategy profiles $g(m)$ such that

1. $g(m)$ is a $\frac{1}{m}-$weak subjective equilibrium;
2. A strategy profile p plays weakly ε_0- like $g(m)$; then p is not a Nash equilibrium.

By the definition of $\frac{1}{m}-$weak subjective equilibrium, there exists strategy profiles $g^i_j(m)$ $i = 1, ..., n$ $j = 1, ..., n$ such that

$$g_i(m) = g^i_i(m) \text{ is a best response to } g^i_{-i}(m) = \left(g^i_1(m), ..., g^i_{i-1}(m), g^i_{i+1}(m), ..., g^i_n(m)\right).$$

and

$$g^i(m) = \left(g^i_1(m), ..., g^n_n(m)\right) \text{ plays weakly } \frac{1}{m} - \text{ like } g(m) = \left(g^1_1(m), ..., g^n_n(m)\right)$$

$\Delta(\Sigma_j)$ is a compact set and $\forall h \in H$;

$$\left(g^i_j(m)\right)(h) \in \Delta(\Sigma_j).$$

Let h_0 be the null history.

Consider a first subsequence, also indexed by m, such that

$$\left(g^i_j(m)\right)(h_0) \xrightarrow[m\to\infty]{} \left(g^i_j\right)(h_0) \quad i = 1, ..., n; \ j = 1, ..., n.$$

The second subsequence, also indexed by m, is a subsequence of the first subsequence such that

$$\left(g_j^i(m)\right)(h) \xrightarrow[m\to\infty]{} \left(g_j^i\right)(h) \quad \forall h \in \Sigma^1 \quad i = 1, \ldots, n; \; j = 1, \ldots, n.$$

The $k-$th subsequence, also indexed by m, is a subsequence of the $(k-1)-$th subsequence such that

$$\left(g_j^i(m)\right)(h) \xrightarrow[m\to\infty]{} \left(g_j^i\right)(h) \quad \forall h \in \Sigma^t; \; t \le k; \; i = 1, \ldots, n; \; j = 1, \ldots, n.$$

Consider a final subsequence, also indexed by m. Take the first elements of the first subsequence, the second elements of the second subsequence and so on, ad infinitum.

Clearly,
$$\left(g^i(m)\right)(h) \xrightarrow[m\to\infty]{} \left(g^i\right)(h) \quad \forall h \in H; \; i = 1, \ldots, n$$

and

$$(g(m))(h) \xrightarrow[m\to\infty]{} g(h) \quad \forall h \in H.$$

Therefore,

$$\mu_{g^i(m)} \xrightarrow[m\to\infty]{} \mu_{g^i} \quad i = 1, \ldots, n \text{ and } \mu_{g(m)} \xrightarrow[m\to\infty]{} \mu_g.$$

But, $\mu_{g^i(m)}$ and $\mu_{g(m)}$ are arbitrarily close in the weak topology. So,

$$\mu_g = \mu_{g^i} \quad i = 1, \ldots, n$$

We want to show that g is a $0-$subjective equilibrium.

Suppose, by contradiction, that there exists a player, say player 1, such that

$$g_1 \text{ is not a best response to } g_{-1}^1 = \left(g_2^1, \ldots, g_n^1\right).$$

So, there exists a behavior strategy l such that

$$V_1(b) = E_{\mu_b}\left\{\sum_{t=1}^{\infty} (\delta_1)^t . u_1\right\} > V_1(g^1) = E_{\mu_{g^1}}\left\{\sum_{t=1}^{\infty} (\delta_1)^t . u_1\right\}$$

where $b = (l, g_2^1, \ldots, g_n^1)$.

Let $b(m)$ be $\left(l, g_2^1(m), \ldots, g_n^1(m)\right)$. By definition,

$$(b(m))(h) \xrightarrow[m\to\infty]{} b(h) \quad \forall h \in H.$$

So,

$$\mu_{b(m)} \underset{m\to\infty}{\longrightarrow} \mu_b.$$

Therefore,

$$V_1(b(m)) \underset{m\to\infty}{\longrightarrow} V_1(b) \text{ and } V_1(g^1(m)) \underset{m\to\infty}{\longrightarrow} V_1(g^1)$$

So, there exists m large enough such that

$$V_1(b(m)) > V_1(g^1(m))$$

which contradicts the fact that

$$g_1(m) \text{ is a best response to } g^1_{-1}(m)$$

So, g is a $0-$subjective equilibrium. There exists a Nash equilibrium $\hat{g}$ that plays $0-$like g (see Kalai and Lehrer (93b)). However,

$$\mu_{g(m)} \underset{m\to\infty}{\longrightarrow} \mu_g = \mu_{\hat{g}}$$

So, if m is large enough, $g(m)$ plays $\frac{\varepsilon}{2}$ like a Nash equilibrium $\hat{g}$.
A contradiction

q.e.d.

Proof of Proposition 4.6 If f is almost absolutely continuous with respect to f^i $i = 1, \ldots, n$ then, by proposition 3.4, $\left(f^1, \ldots, f^n\right)$ and f plays eventually weakly $\gamma-$like a weak subjective $\gamma-$equilibrium, for every $\gamma > 0$. Therefore, by proposition 3.5, $\left(f^1, \ldots, f^n\right)$ and f plays eventually weakly $\varepsilon-$like a Nash equilibrium, for every $\varepsilon > 0$. So, $\left(f^1, \ldots, f^n\right)$ and f plays eventually weakly like a Nash equilibrium.

On the other hand, if beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies $f = (f_1, \ldots, f_n)$ plays eventually weakly like a Nash equilibrium, then, beliefs $\left(f^1, \ldots, f^n\right)$ and optimal strategies $f = (f_1, \ldots, f_n)$ plays eventually weakly $\eta-$like a weak subjective $\eta-$equilibrium, for every $\eta > 0$. So, by proposition 3.4, f is $\varepsilon-$absolutely continuous with respect to f^i $i = 1, \ldots, n$, for every $\varepsilon > 0$. Therefore, f is almost absolutely continuous with respect to f^i $i = 1, \ldots, n$.

q.e.d.

References

[1] Blackwell, D., Dubins, L.: Merging of opinions with increasing information. Annals of Math. Statistics, 38, 882-886 (1962).

[2] Kalai, E., Lehrer, E.: Rational learning leads to Nash equilibriu. Econometrica, 61, 1021-1045 (1993a).

[3] Kalai, E., Lehrer, E.: Subjective equilibrium in repeated games. Econometrica, 61, 1231-1240 (1993b).

[4] Kalai, E., Lehrer, E.: Weak and strong merging of opinions. J. Math. Econ., 23, 73-100 (1993c).

[5] Lehrer, E., Smorodinsky, R.: Repeated large games with incomplete information. imeo, Dept. of Managerial Economics, Northwestern University (1994).

[6] Sandroni, A.: Does rational learning lead to Nash equilibrium in ¯nitely repeated games? University of Pennsylvania, CARESS working paper #94-20 (1994).