E C O N O M I C S   B U L L E T I N

Conditional adaptive strategies

Yuichi Noguchi
Department of Economics, Harvard University

Abstract

A general class of adaptive strategies in Hart and Mas−Colell (2001) may be extended to conditional strategies in the same way as smooth fictitious play in Fudenberg and Levine (1999). We show that a generalized version of universal conditional consistency (UCC) obtains for conditional adaptive strategies under some assumption.

I am very grateful to Drew Fudenberg for his helpful comment.
Citation: Noguchi, Yuichi, (2003) "Conditional adaptive strategies." *Economics Bulletin,* Vol. 3, No. 26 pp. 1−10
Submitted: September 8, 2003. **Accepted:** November 6, 2003.
URL: http://www.economicsbulletin.com/2003/volume3/EB−03C70021A.pdf

1 Introduction

We show that a general class of adaptive strategies in Hart and Mas-Colell (2001) may be extended to *conditional* strategies in the same way as smooth fictitious play in Fudenberg and Levine (1999), and a "classwise generalization" of universal conditional consistency (abbreviated to UCC) obtains for conditional adaptive strategies under some assumption, as it does for conditional smooth fictitious play (Fudenberg and Levine (1999) and Noguchi (2000) and (2002)).

Hart and Mas-Colell (2001) shows a general class of adaptive strategies which has the property of universal consistency by generalizing Blackwell's approachability theorem; we say *universal* consistency because the consistency criterion is passed against *all* opposing strategies. On the other hand, Fudenberg and Levine (1999) extend smooth fictitious play to *conditional* one by introducing learning rules called "classification rules," taking into consideration the case in which a player has a sophisticated ability of learning *regularities* of opponent strategies. Then, they generalize universal consistency in Fudenberg and Levine (1995) to "universal conditional consistency" for conditional smooth fictitious play. Furthermore, Noguchi (2000) and (2002) generalizes their UCC theorem to a *classwise* version for smooth fictitious play. (Generalized) UCC is quite useful to show that a player's strategy is a sophisticated learning procedure. Noguchi (2000) makes use of (generalized) UCC to show that conditional smooth fictitious play passes strong time-average optimality criteria for many opposing strategies. Moreover, Noguchi (2002) shows that generalized UCC implies the same wide range no-regret property as Lehrer (2003) obtains.

Smooth fictitious play is not included in the Hart and Mas-Colell's general class of adaptive strategies, but it may be arbitrarily approximated by adaptive strategies, as Hart and Mas-Colell point out. This fact leads us to infer that adaptive strategies in the Hart and Mas-Colell's sense may also be extended to conditional ones, and that those have the generalized property of universal conditional consistency. The purpose of this paper is to show it is correct at least under some assumption.

The paper is organized as follows. In Section 2 we shall give the basic model and define conditional adaptive strategies. In Section 3 we shall show a classwise generalization of universal conditional consistency for conditional adaptive strategies under some assumption. Section 4 concludes.

2 The Model

2.1 The basic model and notations

We focus on one player who plays an infinitely repeated game against one opponent. The player's payoff at a stage game is denoted by $u(a, y)$, where a is a player's action in a finite set A and y is an opponent's action in a finite set Y. Let $\Delta(S)$ denote the set of all mixed actions over S. Let $u(\lambda, \pi)$ denote the player's expected payoff obtained by playing mixed actions $\lambda \in \Delta(A)$ and $\pi \in \Delta(Y)$. A finite history of actions (up to time T) is denoted by $h_T := (a_1, y_1 \cdots, a_T, y_T)$ and an infinite history of actions is denoted by $h_\infty := (a_1, y_1, a_2, y_2, \cdots)$. The set of all finite histories, including the null history $h_0 := \emptyset$, is denoted by H, and H^∞ is the set of all infinite histories. We denote a behavior strategy of a player by $\sigma : H \to \Delta(A)$, and a behavior strategy of an opponent by $\rho : H \to \Delta(Y)$. We write $\mu_{(\sigma,\rho)}$ for the stochastic process on H^∞ induced by playing σ and ρ.

2.2 Conditional adaptive strategies and classification rules

We shall extend adaptive strategies in Hart and Mas-Colell (2001) to *conditional* strategies. In this paper we focus on a *stationary regret-based action* $\Phi : \mathsf{R}^A \to \Delta(A)$:[1] (i) there exists a continuously differentiable function $P : \mathsf{R}^A \to \mathsf{R}$ such that $\Phi(x)$ is positively proportional to the derivative $\partial P(x)$ for all $x \notin \mathsf{R}^A_-$ and (ii) $\Phi(x) \cdot x > 0$ for all $x \notin \mathsf{R}^A_-$.[2] P is called a *potential* of Φ.

First of all, we shall define an important concept: *classification rules.* Classification rules, introduced as learning rules by Fudenberg and Levine (1999), classify observed samples into *categories.* Formally, a classification rule $\mathcal{R}$ is defined as a partition of $H \times A$,[3] and an element in $\mathcal{R}$ is called a category, denoted by γ; thus, a category γ may be considered as a subset of $H \times A$. If a realized history $(h_{t-1}, a_t) \in \gamma$, we say that time t is a $\gamma-$*effective period* or γ is *effective at time* t; given an infinite history, each period has exactly one effective category because a classification rule is a partition of $H \times A$. Given a history h_T, let n^γ_T denote the number of times that γ has been effective up to time T, and D^γ_T denote the empirical distribution of opponent actions observed in $\gamma-$effective periods up to time T.

[1] Hart and Mas-Colell (2001) call it a stationary regret-based strategy.

[2] A also denotes the cardinality of itself. R^A is an $A-$dimensional Euclidean space. $\mathsf{R}^A_- := \{x \in \mathsf{R}^A \mid x[a] \leq 0 \text{ for all } a \in A\}$.

[3] Fudenberg and Levine (1999) define a classification rule as a function from $H \times A$ to a countable set of categories. But their definition is equivalent to ours.

If a player knew in advance that the current period, say time T, was $\gamma-$effective, then he would pick up observed *regret vectors* $g_t := (u(b, y_t) - u(a_t, y_t))_{b \in A}$ in past $\gamma-$effective periods, and obtain the *conditional average* $\bar{g}^{\gamma}_{T-1}$ *of regret vectors on* γ *up to time* $T-1$: $\bar{g}^{\gamma}_{T-1} := \sum_{\substack{(h_{t-1},a_t)\in\gamma \\ 1\le t\le T-1}} g_t \Big/ n^{\gamma}_{T-1}$. Thus, he would take its stationary regret-based action $\Phi(\bar{g}^{\gamma}_{T-1})$ at time T. However, an effective category may be endogenous in the sense that which category is effective in the current period may depend on which player's action is realized in the current period. Then, we define conditional adaptive strategy σ on $\mathcal{R}$ by extending the fixed point argument in Fudenberg and Levine (1999). We first assume a weight function $w : H \to \mathbb{R}^A_+$; we will precisely define it in the Appendix.[4] Then, let h_{T-1} is a realized past history up to the last period. Let γ_a be the category that is effective at time T if a is realized at time T: $(h_{T-1}, a) \in \gamma_a$. Then, for each $a \in A$, a player obtains $\bar{g}^{\gamma_a}_{T-1}$ and $\Phi(\bar{g}^{\gamma_a}_{T-1})$. Let Z be the matrix in which each column consists of a *weighted* action $w(h_{T-1})[a] \cdot \Phi(\bar{g}^{\gamma_a}_{T-1})$: $Z := [w(h_{T-1})[a] \cdot \Phi(\bar{g}^{\gamma_a}_{T-1})]_{a\in A}$. Note that the magnitudes of $w(h_{T-1})[a]$'s may be different. Thus, let J be the matrix whose diagonal elements are $w(h_{T-1})[a]$'s, and whose off-diagonal elements are all zero. Then, we always find out a mixed action $\lambda^* \in \Delta(A)$ such that $Z\lambda^* = J\lambda^*$.[5] Finally, the player takes $\sigma(h_{T-1}) := \lambda^*$ at time T. We call the procedure *conditional (weighted) adaptive strategy on* $\mathcal{R}$.

2.3 Class

A subset of $H \times A$ will also be called a *class*, denoted by β. When a realized history $(h_{t-1}, a_t) \in \beta$, we say that time t is a $\beta-$*active period*, or that β is *active at time* t. A class indicates periods when payoffs are evaluated. Given a history h_T, let n^{β}_T denote the number of times that β has been active up to time T, and D^{β}_T denote the empirical distribution of opposing actions observed in $\beta-$active periods up to time T.

[4] $\mathbb{R}^A_+ := \{x \in \mathbb{R}^A \mid x[a] \ge 0 \text{ for all } a \in A\}$.

[5] Let $\varphi(\lambda) := \frac{J^{-1}Z\lambda}{\sum_{b\in A}(J^{-1}Z\lambda)[b]}$. It is a continuous function from $\Delta(A)$ to $\Delta(A)$. By the fixed point theorem, there exists a fixed point λ^* of φ. Then, $Z\lambda^* = \alpha J\lambda^*$, where $\alpha = \sum_{b\in A}(J^{-1}Z\lambda^*)[b]$. Note that $\sum_{b\in A}(Z\lambda^*)[b] = \sum_{a\in A} w(h_{T-1})[a]\lambda^*[a] = \sum_{b\in A}(J\lambda^*)[b]$. Thus, $\alpha = 1$. Therefore, $Z\lambda^* = J\lambda^*$. Hart and Mas-Colell (2000) call this type of procedure eigenvector procedures.

3 Universal classwise conditional consistency

3.1 Definition and main result

We shall generalize universal conditional consistency in Fudenberg and Levine (1999) to a *classwise* version for adaptive strategies. Let us first define *classwise* conditional consistency for a *countable* set Ω of classes; Ω is always countable in the following. The criterion requires that conditional consistency hold in active periods of any class in Ω (if that class is active infinitely many times). To define it precisely, when a realized history $(h_{t-1}, a_t) \in \beta \bigcap \gamma$, i.e., both β is active and γ is effective at time t, we say that time t is $\beta\gamma-$*effective*. Given a history h_T, let $n_T^{\beta\gamma}$ denote the number of $\beta\gamma-$effective periods up to time T, and $D_T^{\beta\gamma}$ denote the empirical distribution of opponent actions observed in $\beta\gamma-$effective periods up to time T. Let $\bar{U}_T^\beta$ designate time-average payoff in $\beta-$active periods (up to time T): $\bar{U}_T^\beta := \sum_{\substack{(h_{t-1},a_t)\in\beta \\ 1\le t\le T}} u(a_t, y_t) / n_T^\beta$. The maximum payoff against π is given by $V(\pi) := \max_a u(a, \pi)$.

Definition 1 *We say that conditional (weighted) adaptive strategy σ on $\mathcal{R}$ passes classwise conditional consistency for Ω against ρ, if for all $\beta \in \Omega$, if $n_T^\beta \to \infty$ as $T \to \infty$, then*

$$\limsup_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^\beta} V(D_T^{\beta\gamma}) - \bar{U}_T^\beta \le 0, \ \mu_{(\sigma,\rho)} - a.s.$$

When $\Omega = \{H \times A\}$, classwise conditional consistency is reduced to conditional consistency. To obtain *universal* classwise conditional consistency, we impose two assumptions on $\mathcal{R}$ and Ω. The first assumption requires that a classification rule be eventually finer than any class.

Assumption (A1) *For all $h_\infty \in H^\infty$ and all $\beta \in \Omega$, there exists T_0 such that for all $\gamma \in \mathcal{R}$, either*

$$\begin{aligned} \textit{for all } T &\ge T_0, \textit{ if } (h_{T-1}, a_T) \in \gamma, \textit{ then } (h_{T-1}, a_T) \in \beta, \\ \textit{or for all } T &\ge T_0, \textit{ if } (h_{T-1}, a_T) \in \gamma, \textit{ then } (h_{T-1}, a_T) \notin \beta. \end{aligned}$$

The second one requires that the number of effective categories grow quite slowly in active periods of any class. Given a history h_T, let K_T^β denote the number of categories that have been effective in $\beta-$active periods (up to time T).

Assumption (A2) *For all $h_\infty \in H^\infty$ and all $\beta \in \Omega$, if $n_T^\beta \to \infty$ as $T \to \infty$, then*

$$\lim_{T\to\infty} \frac{K_T^\beta}{n_T^\beta} = 0.$$

When $\Omega = \{H \times A\}$, Assumption (A1) is automatically satisfied and Assumption (A2) is reduced to Assumption 1 in Fudenberg and Levine (1999). Assumptions (A1) and (A2) may be natural requirements about a classification rule and a countable set of classes, but the following last assumption about a stationary regret-based action Φ is rather restrictive.

Assumption (B1) A potential P of Φ satisfies that $\partial P(x)\cdot x \geq P(x)$ for all $x \notin \mathbb{R}_-^A$.

l_p−potentials ($1 < p < \infty$) and separable potentials with a monotone property are typical examples that satisfy Assumption (B1).[6] We shall show that under the assumptions above, conditional (weighted) adaptive strategy on $\mathcal{R}$ has universal classwise conditional consistency for Ω.

Theorem *Suppose that a classification rule $\mathcal{R}$ and a countable set Ω of classes satisfy Assumptions (A1) and (A2) and a stationary regret-based action Φ satisfies Assumption (B1). Then, conditional weighted adaptive strategy σ on $\mathcal{R}$ has universal classwise conditional consistency for Ω: σ passes classwise conditional consistency for Ω against all opposing strategies.*

3.2 Proof of Theorem

According to Lemma 2.3 in Hart and Mas-Colell (2001), any potential P of a stationary regret-based action Φ has the following property: there exists a constant c_0 such that $P(x) > c_0$ for all $x \notin \mathbb{R}_-^A$ and $P(x) = c_0$ for all $x \in bd(\mathbb{R}_-^A)$.[7] Thus, without loss of generality, we may assume that any potential P of Φ satisfies (P1) $\partial P(x)$ is positively proportional to $\Phi(x)$ for all $x \notin \mathbb{R}_-^A$, (P2) $P(x) \geq 0$ for all x, and (P3) $P(x) = 0$ if and only if $x \in \mathbb{R}_-^A$.[8] In order to obtain Theorem, it suffices

[6] $l_p(x) := (\sum_a x_+[a]^p)^{\frac{1}{p}}$ where $x_+[a] := \max\{0, x[a]\}$. Let $\{\psi_a\}_{a\in A}$ be continuous functions from $\mathbb{R}$ to $\mathbb{R}$ such that $\psi_a(z) = 0$ for all $z \leq 0$ and $\psi_a(z) > 0$ for all $z > 0$. Then, a separable potential is defined as $P(x) := \sum_{a\in A} \Psi_a(x[a])$, where $\Psi_a(x[a]) := \int_{-\infty}^{x[a]} \psi_a(z)dz$ for all $a \in A$. For example, when $\{\psi_a\}_a$ are non-decreasing, $\partial P(x) \cdot x \geq P(x)$ for all $x \notin \mathbb{R}_-^A$. See Hart and Mas-Colell (2001) for more examples.

[7] $bd(\mathbb{R}_-^A)$ is the boundary of $\mathbb{R}_-^A$: $bd(\mathbb{R}_-^A) = \{x \in \mathbb{R}^A \mid x[a] \leq 0$ for all $a \in A$, and $x[a] = 0$ for some $a \in A\}$.

[8] Let $\tilde{P}(x) := (P(x) - c_0)^2$ if $x \notin \mathbb{R}_-^A$ and $\tilde{P}(x) := 0$ if $x \in \mathbb{R}_-^A$. Then, $\tilde{P}$ is a potential of Φ and satisfies (P1), (P2) and (P3).

to show two lemmas. We first extend an important result in Hart and Mas-Colell (2001) to a classwise conditional version: conditional average of P in active periods of any class converges to zero.

Lemma 1 *Suppose that a classification rule $\mathcal{R}$ and a countable set Ω of classes satisfy Assumptions (A1) and (A2), and a stationary regret-based action Φ satisfies Assumption (B1). Then, conditional weighted adaptive strategy σ on $\mathcal{R}$ has the following property: for all ρ and all $\beta \in \Omega$, if $n_T^{\beta} \to \infty$ as $T \to \infty$, then*

$$\lim_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} P(\bar{g}_T^{\beta\gamma}) = 0,\ \mu_{(\sigma,\rho)} - a.s.$$

where $\bar{g}_T^{\beta\gamma}$ is the conditional average of regret vectors on $\beta\gamma$ up to time T: $\bar{g}_T^{\beta\gamma} := \sum_{\substack{(h_{t-1},a_t)\in\beta\cap\gamma \\ 1\le t\le T}} g_t / n_T^{\beta\gamma}$.

Proof. See the Appendix. ■

The key is that the above lemma induces classwise conditional consistency.

Lemma 2 *If $\lim_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} P(\bar{g}_T^{\beta\gamma}) = 0$, then classwise conditional consistency obtains:*

$$\limsup_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} V(D_T^{\beta\gamma}) - \frac{1}{n_T^{\beta}} \sum_{\substack{(h_{t-1},a_t)\in\beta \\ 1\le t\le T}} u(a_t, y_t) \le 0.$$

Proof. We may assume that the domain of P is bounded because the range of regret vectors are bounded; thus P is uniformly continuous. Then, (P2) and (P3) imply that for all $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that $P(x) < \delta_\varepsilon \Rightarrow \max_a x[a] < \varepsilon$. Given a history h_T, let $\mathcal{R}_T^{\beta}(P;\delta) := \{\gamma \in \mathcal{R} \mid P(\bar{g}_T^{\beta\gamma}) < \delta\}$. Then, it follows from the assumption that for all δ, $\eta > 0$ there exists $T_{\delta\eta}$ such that $\sum_{\gamma\in\mathcal{R}_T^{\beta}(P;\delta)} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} \ge 1 - \eta$ for all $T \ge T_{\delta\eta}$.

Take any $\varepsilon > 0$. Let $\delta^* := \delta_{\frac{\varepsilon}{2}}$ and $\eta^* := \varepsilon / 4\bar{u}$, where $\bar{u} := \max_{a,y} \mid u(a,y) \mid$. Then, for all $T \ge T_{\delta^*\eta^*}$, $\sum_{\gamma\in\mathcal{R}_T^{\beta}(P;\delta^*)} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} \ge 1 - \eta^*$. Note that for all $T \ge T_{\delta^*\eta^*}$, all $\gamma \in \mathcal{R}_T^{\beta}(P;\delta^*)$, and all $a \in A$, $\bar{g}_T^{\beta\gamma}[a] = u(a, D_T^{\beta\gamma}) - \bar{U}_T^{\beta\gamma} < \frac{\varepsilon}{2}$, where $\bar{U}_T^{\beta\gamma} := \sum_{\substack{(h_{t-1},a_t)\in\beta\cap\gamma \\ 1\le t\le T}} u(a_t, y_t) / n_T^{\beta\gamma}$. Thus, for all $T \ge T_{\delta^*\eta^*}$ and all $\gamma \in \mathcal{R}_T^{\beta}(P;\delta^*)$, $V(D_T^{\beta\gamma})-$

$\bar{U}_T^{\beta\gamma} < \frac{\varepsilon}{2}$. Therefore, for all $T \geq T_{\delta^*\eta^*}$

$$
\begin{aligned}
& \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} V(D_T^{\beta\gamma}) - \frac{1}{n_T^{\beta}} \sum_{\substack{(h_{t-1},a_t)\in\beta \\ 1\leq t\leq T}} u(a_t, y_t) \\
= & \sum_{\gamma\in\mathcal{R}_T^{\beta}(P;\delta^*)} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} [V(D_T^{\beta\gamma}) - \bar{U}_T^{\beta\gamma}] + \sum_{\gamma\notin\mathcal{R}_T^{\beta}(P;\delta^*)} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} [V(D_T^{\beta\gamma}) - \bar{U}_T^{\beta\gamma}] \\
\leq & \frac{\varepsilon}{2} + \frac{\varepsilon}{4\bar{u}} 2\bar{u} \\
\leq & \varepsilon.
\end{aligned}
$$

■

Proof of Theorem. It is immediate from Lemmas 1 and 2. ¥

4 Concluding Remark

We conclude with giving a remark. We have shown universal classwise conditional consistency by imposing some assumption (i.e., Assumption (B1)) on a stationary regret-based action. However, we may conjecture that universal classwise conditional consistency obtains *without* Assumption (B1). Indeed, Noguchi (2003) shows generalized UCC obtains (without Assumption (B1)) in an uncalibrated case that a classification rule and classes do not depend at all on player's current actions.

Appendix

Proof of Lemma 1. We may assume that the domain of P is bounded because the range of regret vectors are bounded; thus, ∂P is uniformly continuous. Let $S_T^{\beta} := \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^{\beta}} P(\bar{g}_T^{\beta\gamma})$. We define a random variable $X_T[\beta](h_\infty)$ as

$$X_T[\beta](h_\infty) := n_T^{\beta} S_T^{\beta} - n_{T-1}^{\beta} S_{T-1}^{\beta} - c_{n_T^{\beta\gamma}}^{\beta\gamma} - o_{n_T^{\beta\gamma}}^{\beta\gamma}, \text{ if } (h_{T-1}, a_T) \in \beta,$$

$$X_T[\beta](h_\infty) := 0, \text{ otherwise}$$

where $c_{n_T^{\beta\gamma}}^{\beta\gamma} := (\partial P(\bar{g}_{T-1}^{\beta\gamma}) - \partial P(\bar{g}_{T-1}^{\gamma})) \cdot g_T$, $o_{n_T^{\beta\gamma}}^{\beta\gamma} := n_T^{\beta\gamma}(P(\bar{g}_T^{\beta\gamma}) - P(\bar{g}_{T-1}^{\beta\gamma})) - \partial P(\bar{g}_{T-1}^{\beta\gamma}) \cdot (g_T - \bar{g}_{T-1}^{\beta\gamma})$, and γ is the effective category at time T, i.e., $(h_{T-1}, a_T) \in \gamma$. Note that $o_m^{\beta\gamma}$ uniformly converges to 0 as $m \to \infty$.

Let $\bar{X}_T[\beta] := \frac{1}{n_T^\beta}\sum_{t=1}^{T} X_t[\beta]$ and $[\bar{X}_T]_+(h_\infty)[\beta] := \max\{0, \bar{X}_T(h_\infty)[\beta]\}$. Take any probability distribution $\mathbf{p} = (p_\beta)_\beta$ on Ω such that $p_\beta > 0$ for all $\beta \in \Omega$. Then, we define a weight function $w : H \to \mathbb{R}_+^A$ as follows:

$$w(h_T)[a] := \Big(\sum_{\beta \ni (h_T,a)} p_\beta \cdot [\bar{X}_T]_+[\beta] \cdot \frac{1}{n_{T+1}^\beta}\Big) \cdot \Big(\sum_{b\in A} \partial P(\bar{g}_T^{\gamma_a})[b]\Big), \text{ if } \exists \beta \in \Omega(\beta \ni (h_T, a)),$$

$$w(h_T)[a] := 0, \text{ otherwise}$$

where γ_a is the category that is effective at time $T+1$ if a is realized at time $T+1$: $(h_T, a) \in \gamma_a$. Let $\langle , \rangle$ denote an inner product on L^2.[9] The product measure of $\mu_{(\sigma,\rho)}$ and $\mathbf{p}$ is denoted by $\mu \times \mathbf{p}$.

Step 1: When $(h_{T-1}, a_T) \in \beta$, it follows from (P3) and Assumption (B1) that

$$\begin{aligned} X_T[\beta](h_\infty) &= \partial P(\bar{g}_{T-1}^{\beta\gamma}) \cdot (g_T - \bar{g}_{T-1}^{\beta\gamma}) + P(\bar{g}_{T-1}^{\beta\gamma}) + o_{n_T^{\beta\gamma}}^{\beta\gamma} - c_{n_T^{\beta\gamma}}^{\beta\gamma} - o_{n_T^{\beta\gamma}}^{\beta\gamma} \\ &\le \partial P(\bar{g}_{T-1}^{\gamma}) \cdot g_T + c_{n_T^{\beta\gamma}}^{\beta\gamma} - c_{n_T^{\beta\gamma}}^{\beta\gamma} \\ &\le \partial P(\bar{g}_{T-1}^{\gamma}) \cdot g_T. \end{aligned}$$

Let $\delta_T^\beta := n_T^\beta - n_{T-1}^\beta$. Note that $\delta_T^\beta = 1$ if β is active at time T, and $\delta_T^\beta = 0$ otherwise. Let $E_\mu[\cdot \mid h_T]$ be conditional expectation on h_T (with respect to $\mu_{(\sigma,\rho)}$). Then, it follows from the inequality above, (P1), and the definition of conditional weighted adaptive strategy that

$$\begin{aligned} &E_\mu[\langle [\bar{X}_T]_+, \frac{\delta_{T+1}^\beta}{n_{T+1}^\beta} X_{T+1}\rangle \mid h_T] \\ &= E_\mu[\sum_\beta p_\beta \cdot [\bar{X}_T]_+[\beta] \cdot \frac{1}{n_{T+1}^\beta} \cdot X_{T+1}[\beta] \mid h_T] \\ &\le \sum_a \sum_y \sigma(h_T)[a] \cdot \rho(h_T)[y] \cdot \Big(\sum_{\beta \ni (h_T,a)} p_\beta \cdot [\bar{X}_T]_+[\beta] \cdot \frac{1}{n_{T+1}^\beta} \cdot \partial P(\bar{g}_T^{\gamma_a}) \cdot g_{T+1}\Big) \\ &= u(Z\sigma(h_T), \rho(h_T)) - u(J\sigma(h_T), \rho(h_T)) \\ &= 0. \end{aligned}$$

Step 2: Define $[\bar{X}_T]_- := \bar{X}_T - [\bar{X}_T]_+$ and $proj_{L^2_-}(\bar{X}_T)(h_\infty) := \arg\min_{Y \in L^2_-} \|\bar{X}_T(h_\infty) - Y\|$.[10] Then, $proj_{L^2_-}(\bar{X}_T) = [\bar{X}_T]_-$ and $\langle [\bar{X}_T]_+(h_\infty), [\bar{X}_T]_-(h_\infty)\rangle = 0$. Thus, letting

[9] $\langle X, Y\rangle := \sum_\beta p_\beta \cdot X[\beta] \cdot Y[\beta]$ and $L^2 := \{Y \in \mathbb{R}^\Omega \mid \sum_\beta p_\beta \cdot (Y[\beta])^2 < \infty\}$.
[10] $L^2_- := \{Y \in L^2 \mid Y[\beta] \le 0 \text{ for all } \beta \in \Omega\}$.

$E_{\mu\times\mathbf{p}}[\cdot]$ be expectation with respect to $\mu \times \mathbf{p}$, the second inequality in Step 1 implies that

$$\begin{aligned}
&\sum_{t=1}^{\infty} E_{\mu\times\mathbf{p}}[(\bar{X}_t - proj_{L^2_-}(\bar{X}_t)) \cdot (\frac{\delta^{\beta}_{t+1}}{n^{\beta}_{t+1}}(X_{t+1} - proj_{L^2_-}(\bar{X}_t)))] \\
=& \sum_{t=1}^{\infty} E_{\mu}[\langle \bar{X}_t - proj_{L^2_-}(\bar{X}_t), \frac{\delta^{\beta}_{t+1}}{n^{\beta}_{t+1}}(X_{t+1} - proj_{L^2_-}(\bar{X}_t))\rangle] \\
=& \sum_{t=1}^{\infty} E_{\mu}[\langle [\bar{X}_t]_+, \frac{\delta^{\beta}_{t+1}}{n^{\beta}_{t+1}} X_{t+1}\rangle] \\
\leq& \ 0.
\end{aligned}$$

Therefore, we can apply a conditional version of the strong law of large numbers (see Theorem 4 (and Corollary 1) in Lehrer (2002)), so that for all $\beta \in \Omega$, if $n^{\beta}_T \to \infty$, $[\bar{X}_T]_+(\beta) \to 0$, $\mu_{(\sigma,\rho)} - a.s.$

Step 3: Given a history h_T, let $\mathcal{R}^{\beta}_T(\epsilon) := \{\gamma \in \mathcal{R} \mid \frac{1}{n^{\beta\gamma}_T} \leq \epsilon\}$. Then, Assumption (A2) is equivalent to the following condition: $(*)$ for all $h_\infty \in H^\infty$, if $n^{\beta}_T \to \infty$, then for all ϵ, $\eta > 0$, there exists $T_{\epsilon\eta}$ such that $\sum_{\gamma\in\mathcal{R}^{\beta}_T(\epsilon)} \frac{n^{\beta\gamma}_T}{n^{\beta}_T} \geq 1 - \eta$ for all $T \geq T_{\epsilon\eta}$. Thus, for all $h_\infty \in H^\infty$, if $n^{\beta}_T \to \infty$, then

$$\lim_{T\to\infty} \frac{1}{n^{\beta}_T} \sum_{\gamma\in\mathcal{R}^{\beta}_T} \sum_{m=1}^{n^{\beta\gamma}} o^{\beta\gamma}_m = \lim_{T\to\infty} \sum_{\gamma\in\mathcal{R}^{\beta}_T} \frac{n^{\beta\gamma}_T}{n^{\beta}_T}[\frac{1}{n^{\beta\gamma}_T} \sum_{m=1}^{n^{\beta\gamma}} o^{\beta\gamma}_m] = 0$$

where $\mathcal{R}^{\beta}_T$ is the set of all categories that have been effective in $\beta-$active periods up to time T.

Step 4: The range of regret vectors is bounded. It, together with Assumption (A1), implies that there exists a constant $C_1 > 0$ such that for any effective category γ in $\beta-$active periods from time T_0 on, $\|\bar{g}^{\beta\gamma}_T - \bar{g}^{\gamma}_T\| \leq \frac{T_0}{n^{\gamma}_T} \cdot C_1$ for all $T \geq T_0$, where T_0 is a calendar time in Assumption (A1). Further, there exists $C_2 > 0$ such that for all γ and all T, $| c^{\beta\gamma}_{n^{\beta\gamma}_T} | \leq C_2 \cdot \|\partial P(\bar{g}^{\beta\gamma}_{T-1}) - \partial P(\bar{g}^{\gamma}_{T-1})\|$. From these and Condition $(*)$ it follows that for all $h_\infty \in H^\infty$, if $n^{\beta}_T \to \infty$, $\lim_{T\to\infty} \frac{1}{n^{\beta}_T} \sum_{\gamma\in\mathcal{R}^{\beta}_T} \sum_{m=1}^{n^{\beta\gamma}_T} c^{\beta\gamma}_m = 0$.

Step 5: Finally, from Steps 2, 3 and 4 it follows that for all $\beta \in \Omega$, if $n_T^\beta \to \infty$,

$$
\begin{aligned}
& \limsup_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^\beta} P(\bar{g}_T^{\beta\gamma}) \\
= & \limsup_{T\to\infty} [\frac{1}{n_T^\beta} \sum_{t=1}^{T} (n_t^\beta S_t^\beta - n_{t-1}^\beta S_{t-1}^\beta) - \frac{1}{n_T^\beta} \sum_{\gamma\in\mathcal{R}_T^\beta} \sum_{m=1}^{n_T^{\beta\gamma}} c_m^{\beta\gamma} - \frac{1}{n_T^\beta} \sum_{\gamma\in\mathcal{R}_T^\beta} \sum_{m=1}^{n_T^{\beta\gamma}} o_m^{\beta\gamma}] \\
= & \lim_{T\to\infty} \bar{X}_T[\beta] \\
\leq & \lim_{T\to\infty} [\bar{X}_T]_+[\beta] \\
= & 0,\ \mu_{(\sigma,\rho)} - a.s.
\end{aligned}
$$

Further, by (P2), $\liminf_{T\to\infty} \sum_{\gamma\in\mathcal{R}} \frac{n_T^{\beta\gamma}}{n_T^\beta} P(\bar{g}_T^{\beta\gamma}) \geq 0$ for all $h_\infty \in H^\infty$. Thus the desired result obtains. ¥

References

Fudenberg, D., and D. K. Levine (1995) "Consistency and Cautious Fictitious Play," J. Econ. Dynam. Control **19**, 1065-1089.

Fudenberg, D., and D. K. Levine (1999) "Conditional Universal Consistency," Games and Econ. Behav. **29**, 104-130.

Hart, S., and A. Mas-Colell (2000) "A Simple Adaptive Procedure Leading to Correlated Equilibrium," Econometrica **68**, 1127-1150.

Hart, S., and A. Mas-Colell (2001) "A General Class of Adaptive Strategies," J. Econ. Theory **98**, 26-54.

Lehrer, E. (2002) "Approachability in Infinite Dimensional Spaces," Int. J. Game Theory **31**, 253-268.

Lehrer, E. (2003) "A Wide Range No-Regret Theorem," Games and Econ. Behav. **42**, 101-115.

Noguchi, Y. (2000) "Optimal Properties of Conditional Fictitious Play," in Ph.D. Thesis at Harvard University.

Noguchi, Y. (2002) "Note on Wide Range No-Regret," Mimeo., Hitotsubashi University.

Noguchi, Y. (2003) "Optimal Properties of Conditional Adaptive Strategies," Mimeo., Hitotsubashi University.