Economics Bulletin

Volume 38, Issue 3

On the purification of mixed strategies

Chaowen Yu
Department of Economics, Keio University

Yuhki Hosoya
Department of Economics, Kanto-Gakuin University

Toru Maruyama
Department of Economics, Keio University

Abstract

We try to examine examples in Aumann et al. (1983) which illuminate the impossibility of purification in a more readable setup. We also provide a rigorous proof of the possibility of approximate purifications that is suggested in Aumann et al. (1983). Further, the relationship between the concept of "conditionally atomless" and that of "weakly conditionaly atomless" is clarified.

We are grateful for Nobusumi Sagara and Takuya Masuzawa for their helpful comments and suggestions. We are also grateful for one anonymous referee for his/her precise advices.
Citation: Chaowen Yu and Yuhki Hosoya and Toru Maruyama, (2018) "On the purification of mixed strategies", *Economics Bulletin*, Volume 38, Issue 3, pages 1655-1675
Contact: Chaowen Yu - yu-chaowen@keio.jp, Yuhki Hosoya - hosoya@kanto-gakuin.ac.jp, Toru Maruyama - maruyama@econ.keio.ac.jp.
Submitted: December 20, 2017. **Published:** September 07, 2018.

1 Introduction

This paper presents a rigorous proof of results on purification of mixed strategies of games. Although we mainly concern impossibility results of purification (Examples 1, 2), we also discuss several possibility results and compare various requirements for purification (Propositions 1-3).

In game theory, there is a reasonable motivation to treat mixed strategies rather than pure strategies. For instance, the existence proof of solutions (e.g., Nash equilibrium, α-core, ...) can be made easier to some extent by admitting mixed strategies. However, in many situations, it appears unnatural for players to choose mixed strategies, because players are required to have a certain device generating random choices. Therefore, we want to obtain an "equivalent" pure strategy to the given mixed strategy, which is called the **purification** of this mixed strategy.

Two kinds of purification are treated in this paper. One is **exact purification**, and the other is **approximate purification**. When the signal space is of Radner-Rosenthal type,[1] there are many strong results on exact purification.[2] On the other hand, Aumann et al. (1983) showed several results on approximate purification under general setup of signals.[3]

The contents of this paper is closely related to results in Aumann et al. (1983). Hence, we start with summarizing their results briefly. First, they showed that if the prior of the signal space is **conditionally atomless** (the definition of this term is explained later), then every mixed strategy can be approximately purified (Theorem, Corollary A). Second, they showed that the "conditional atomless" requirement can be replaced by the "weakly conditionally atomless" requirement if the purified strategy profile is a Nash equilibrium (Corollary B).

Aumann et al. (1983) also mentioned the difficulty of guaranteeing the existence of both exact and approximate purifications under the general setup of signal spaces. For example, in section 3, they argued an example in which the exact purification is impossible. Moreover, in section 7, they provided an example in which the approximate purification is also impossible even if the prior is weakly conditionally atomless. However, their explanations are somewhat vague, which interferes with readers' easy understanding of

[1] See Radner and Rosenthal (1982).

[2] Khan, Rath, and Sun (2006) and Khan and Rath (2009) treated this problem by using the strengthened result of Dvoretsky, Wald, and Wolfowitz (1951). Note that their definition of purification is stronger than the "exact purification" in this paper.

[3] In their setup, the signal space is assumed to be homeomorphic to $[0,1]$. By Kuratowski's theorem, every continuum Polish space with its Borel structure satisfies this requirement. See subsection 2.1.

their counterexamples. One of main purposes of this paper is to clarify these ambiguous arguments by providing rigorous explanations of these examples (Examples 1, 2).

Meanwhile, in sections 3 and 6 of Aumann et al. (1983), they suggested that under some atomless requirement or absolute continuity of the prior, the approximately purifying strategies can be constructed concretely. In this paper, we show these results rigorously (Propositions 1, 2). Further, we clarify the relationship between "conditionally atomless" and "weakly conditionally atomless" criterions. In section 7 of Aumann et al. (1983), they constructed an example of prior that is not conditionally atomless but only weakly conditionally atomless. However, they did not show that conditionally atomless criterion implies weakly conditionally atomless criterion. Hence, the term 'weakly' should be justified by proving that the former implies the latter. We show this result (Proposition 3).

In subsection 2.1, we interpret the setup of the game. In subsections 2.2 and 2.3, we explain the difficulty of exact and approximate purifications, respectively. Section 3 is devoted to the conclusion. All proofs are in the appendix.

2 Results

2.1 Basic Notations

We treat only games with a finite number of players and actions. Let $N = \{1, ..., n\}$ be the player set, $K_i = \{k_i^1, ..., k_i^{m_i}\}$ be the action set for player i, and let $u_i : \prod_{i=1}^n K_i \to \mathbb{R}$ be the payoff function. The probability space $(\Omega, \mathscr{S}, \mu)$ is interpreted as the space of **outside signals**. We assume that $\omega \in \Omega$ does not affect the level of payoff: that is, u_i is not a function on $\Omega \times \prod_{i=1}^n K_i$.[4] Let $x_i : \Omega \to X_i$ be the **observation function**, and assume that player i can observe x_i. Under this structure, we can identify the set Ω as the product set $\prod_{i=1}^n X_i$, and thus we assume that $\Omega = \prod_{i=1}^n X_i$. Throughout this paper, we assume that X_i is a Polish space with continuum cardinality endowed with the Borel structure. By Kuratowski's theorem,[5] X_i with the Borel structure is isomorphic to $[0, 1]$ as a measure space. We call μ **conditionally atomless for player** i if the conditional probability $\mu(\cdot|x_{-i})$ is non-atomic almost surely (a.s.) with respect to (w.r.t.) the marginal probability $\mu_{X_{-i}}$, where $x_{-i} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$ and $X_{-i} = \prod_{j \in N, j \neq i} X_j$ as usual.

[4]See also the notes at the end of this subsection.

[5]See Theorem 3.3.13 of Srivastava (1998).

We call a measurable function $k_i : X_i \to K_i$ a **pure strategy** of player i. Despite the name, the actual action of player i who obeys a pure strategy is randomized by the outside signal x_i. Meanwhile, let $\Delta(K_i)$ be the set of all probabilities on K_i. We call a measurable function $k_i : X_i \to \Delta(K_i)$ a **mixed strategy**. If k_i is a mixed strategy, then $k_i(x_i) = (p_i^1(x_i), ..., p_i^{m_i}(x_i))$, where $p_i^j(x_i)$ is the probability at which player i takes action k_i^j under signal x_i. We call a mixed strategy k_i **simple** if k_i is a constant function. The interpretation of the simple mixed strategy is as follows: player i does not use the outside signal, but instead uses a certain randomizing device that is independent of the signals. $k = (k_1, ..., k_n)$ is called a **strategy profile** if each k_i is a (pure, or mixed) strategy of player i. The expected payoff function U_i for player i is defined by

$$U_i(k) = \int \sum_{j_1, j_2, ..., j_n} \prod_{\ell=1}^{n} p_\ell^{j_\ell}(x_\ell) u_i(k_1^{j_1}, ..., k_n^{j_n}) d\mu.$$

Let k_i be a mixed strategy of player i. A strategy k_i' is **exactly equivalent** to k_i if and only if for every strategy profile k_{-i} of other players and every $\ell \in N$,

$$U_\ell(k_i, k_{-i}) = U_\ell(k_i', k_{-i}).$$

Similarly, a strategy k_i' is **ε-equivalent** to k_i if and only if for every strategy profile k_{-i} of other players and every $\ell \in N$,

$$|U_\ell(k_i, k_{-i}) - U_\ell(k_i', k_{-i})| < \varepsilon.$$

A pure strategy k_i' is called an **exact purification** (resp. ε-purification) of k_i if k_i' is exactly (resp. ε-) equivalent to k_i.

Notes on Basic Notations. Many related studies assumed that u_i is affected by signals, because in many situations, the signal includes information about the type of player i. The reason why we assume that ω does not affect u_i is just for simplicity. Because our purpose is to illuminate examples of impossibility, this simplification does not impose any limitations on our research. Rather, this restriction shows that the impossibility arises **even if** signals do not affect payoffs. For the same reason, the use of a simple mixed strategy in our example indicates that the impossibility result holds **even though** we restrict the theory to simple mixed strategies.

In related researches, our definition of equivalence is sometimes called the **payoff equivalence**. There is another notion of equivalence, called the **distributional equivalence**. k_i and k_i' are distributionally equivalent if for every $k_i^j \in K_i$,

$$\int_{X_i} p_i^j(x_i) d\mu = \int_{X_i} p_i^{j\prime}(x_i) d\mu.$$

However, we can easily deduce the distributional equivalence result from the payoff equivalence result: if we add an artificial player $n+1$, whose strategy space is $\{\emptyset\}$ and whose payoff function is

$$\begin{cases} 1 & \text{if } k_i = k_i^j, \\ 0 & \text{otherwise,} \end{cases}$$

then the payoff equivalence leads to the distributional equivalence.

2.2 First Impossibility Result: on Exact Purification

We simply assume that $N = \{1, 2\}$, $K_1 = K_2 = \{0, 1\}$ and $X_1 = X_2 = [0, 1]$. Define the payoff function as

$$u_1(k_1, k_2) = \begin{cases} 1 & \text{if } k_1 = k_2 = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$
$$u_2(k_1, k_2) \equiv 0.$$

Example 1 (pp.329-330 of Aumann et al. (1983)). Let μ have the following density:

$$h(x_1, x_2) = \begin{cases} 2 & \text{if } x_1 \le x_2, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Then, there exists no exact purification of the simple $(1/2, 1/2)$-mixed strategy for player 1.[6] The proof is in the appendix.

Therefore, in many cases there is no exact purification.

This problem can be avoided by using the ε-purification. Fix any $p \in]0, 1[$, and let k_1 be the simple $(p, 1-p)$-mixed strategy.[7] Consider the following sequence:

$$S_m = \left[0, \frac{p}{m}\right] \cup \left[\frac{1}{m}, \frac{1+p}{m}\right] \cup ... \cup \left[\frac{m-1}{m}, \frac{m-1+p}{m}\right].$$

Then, the following proposition holds:

[6] If the signal structure satisfies the assumptions of Radner and Rosenthal (1982), then for every mixed strategy, we can obtain an exact purification. However, in this example, the signal structure is not the Radner-Rosenthal type.

[7] That is, p is the probability with which k_1 chooses 0 regardless of the observed signal.

Proposition 1. Suppose that $\mu(\cdot|x_2)$ has a density function $h(\cdot|x_2)$ for almost all x_2 w.r.t. μ_{X_2}. Then, the pure strategy $k_1^m = 1_{S_m}$ satisfies

$$\sup_{k_2} |U_1(k_1^m, k_2) - U_1(k_1, k_2)| \to 0 \text{ as } m \to \infty.$$

This proposition can be extended by using the Fourier analysis techniques.

Proposition 2. Suppose that μ is conditionally atomless for player 1. Define the pure strategies $k_1^m = 1_{S_m}$. Then, there exists a subsequence $(k_1^{m(n)})$ such that

$$\sup_{k_2} |U_1(k_1^{m(n)}, k_2) - U_1(k_1, k_2)| \to 0 \text{ as } n \to \infty.$$

Note that in both propositions, u is not assumed to be defined by (1). This result is very general. Therefore, we can treat the ε-purification under conditionally atomless condition of the prior μ in the signal space.[8]

Clearly, without the conditionally atomless requirement, ε-purification may be impossible. For example, consider $\mu(A) = \lambda(\{x \in [0,1] | (1,x) \in A\})$, where λ is the Lebesgue measure. Then, every pure strategy of player 1 is actually constant, and thus the simple $(1/2, 1/2)$-mixed strategy clearly cannot be ε-purified for sufficiently small ε.

2.3 Second Impossibility Result: under Only Weakly Conditionally Atomless Requirement

Let $\mu_{ij} = \mu_{X_i \times X_j}$ be the marginal probability of μ. We call μ **weakly conditionally atomless** for player i if for every $j \in N$ with $i \neq j$, μ_{ij} is conditionally atomless for player i. The next proposition justifies this name.

Proposition 3. If μ is conditionally atomless for every player, then μ is weakly conditionally atomless.

If the number of players is two, then clearly μ is conditionally atomless if and only if μ is weakly conditionally atomless. Moreover, according to Aumann et al. (1983), the weakly conditionally atomless requirement of the prior assures the possibility of the approximate purification of the strategy profile provided that it is a Nash equilibrium, even if the number of players is greater than two (Corollary B). Therefore, one might think that for

[8] Actually, both propositions hold even when u_i depends on signals: construct another prior ν as in the proof of Theorem of Aumann et al. (1983), and apply Lemma 1 or 2. Note that the requirement $K_1 = K_2 = \{0, 1\}$ can be weakened to the general case by changing the definition of S_m.

approximate purification, only weakly conditionally atomless requirement is needed, and conditionally atomless requirement is not needed.

However, the next example shows that this conjecture is not true; that is, for ε-purification, the weakly conditionally atomless requirement is not sufficient. Note that, this example also shows the existence of the prior that is not conditionally atomless but only weakly conditionally atomless.

Example 2 (pp.338-340 of Aumann et al. (1983)). Consider $N = \{1, 2, 3\}$, $K_i = \{0, 1\}$, $X_1 =]0, 1]^2$, $X_2 = X_3 =]0, 1]$ and

$$u_1(k_1, k_2, k_3) = \begin{cases} 1 & \text{if } k_1 = k_2 = k_3 = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

$$u_2 \equiv 0, \ u_3 \equiv 0.$$

Let $\mathscr{Q}_m$ be an algebra generated by $\left]\frac{j}{4^m}, \frac{j+1}{4^m}\right]$ for $j = 0, ..., 4^m - 1$, and $\mathscr{Q}_m^2 = \mathscr{Q}_m \otimes \mathscr{Q}_m$. Let $\mathscr{Q} = \cup_m \mathscr{Q}_m^2$. Because $\mathscr{Q}_m^2$ is finite, $\mathscr{Q}$ is countable infinite, and thus can be numbered; that is, $\mathscr{Q} = \{A_0, A_1, ...\}$. Without loss of generality, we can assume that $A_m \in \mathscr{Q}_m^2$. Define

$$C_{ij}^m = \left]\frac{i}{4^m}, \frac{i+1}{4^m}\right] \times \left]\frac{j}{4^m}, \frac{j+1}{4^m}\right].$$

Now, let θ_0 be the usual Lebesgue measure on $]0, 1]^2$, and inductively define

$$f^{m+1}(x) = \begin{cases} 2^{m+1}, & \text{if } x \in C_{i'j'}^{m+1} \subset C_{ij}^m, C_{ij}^m \subset A_m, \\ & \theta_m(C_{ij}^m) > 0, i' + j' \text{ is even,} \\ & \text{or } x \in C_{i'j'}^{m+1} \subset C_{ij}^m, C_{ij}^m \not\subset A_m, \\ & \theta_{m-1}(C_{ij}^m) > 0, i' + j' \text{ is odd,} \\ 0, & \text{otherwise,} \end{cases}$$

and θ_m as the probability measure whose density function is f^m. Then, θ_m converges to some θ^* w.r.t. the Prohorov metric.[9] For a Borel set $A \subset X_1 \times X_2 \times X_3$, let

$$\theta(A) = \theta^*(\{(x, y) | (x, y, x, y) \in A\}).$$

Then, θ is weakly conditionally atomless. However, the simple $(1/2, 1/2)$-mixed strategy of player 1 cannot be ε-purified for $\varepsilon = 0.05$.[10]

[9] The definition of the Prohorov metric is in the appendix.

[10] It is clear that $\theta(\cdot | x_3, x_4) = \delta_{(x_3, x_4)}$, and thus, it is not conditionally atomless for player 1.

3 Conclusion

We have presented two examples. One indicates the impossibility of exact purification under general setups, though approximate purification is possible. The other indicates that the conditionally atomless requirement is crucial for approximate purification. Along with these examples, we have stated three propositions. The first says that under general setups and the absolute continuity requirement, we can construct concretely the approximate purification strategy. The second says that the absolute continuity requirement can be weakened to the conditionally atomless requirement. The third says that the weakly conditionally atomless requirement is weaker than the conditionally atomless requirement.

We have not treated the Nash equilibria. To restrict our attention to the Nash equilibrium strategy profiles, it is known that there exists an approximate purification under the weakly conditionally atomless requirement. See Aumann et al. (1983).

It is known that the countability of K_i is crucial for exact purification. If K_i is uncountable, then the signal space must be **saturated** even when the Radner-Rosenthal requirement holds. For purification results in saturated spaces, see, for example, Khan and Zhang (2014).

Meanwhile, we guess that for approximate purification, the saturated signal space is not needed even when K_i is uncountable. We think that the Polish requirement of the signal space is somewhat natural, and any Polish space is not saturated. Therefore, we want to reconsider the approximate purification. However, this is still a future task.

A Proofs

A.1 Proof of Claim in Example 1

We can easily verify that μ_{X_2} has the following density:

$$h_{X_2}(x_2) = 2x_2,$$

and if $x_2 > 0$, then $\mu(\cdot|x_2)$ has the following density

$$h(x_1|x_2) = \begin{cases} \frac{1}{x_2} & \text{if } x_1 \le x_2, \\ 0 & \text{otherwise.} \end{cases}$$

Now, let $k_1(x_1) \equiv (1/2, 1/2)$ and $k_1' = 1_S$ be an exact purification of k_1, where $S \subset [0, 1]$ is some Borel set. Let $k_2 = 1_T$ be a pure strategy of player

2, where $T \subset [0,1]$ is some Borel set. Then, by a simple calculation,

$$U_1(k_1, k_2) = \int \frac{1}{2} 1_T(x_2) d\mu = \int_T \frac{1}{2} d\mu_{X_2},$$
$$U_1(k_1', k_2) = \int 1_S(x_1) 1_T(x_2) d\mu = \int_T \left(\int 1_S(x_1) \mu(dx_1|x_2) \right) d\mu_{X_2},$$

and thus, by the uniqueness of the Radon-Nikodym derivative, we have

$$\frac{1}{2} = \int_{X_1} 1_S(x_1) \mu(dx_1|x_2)$$
$$= \frac{\int_0^{x_2} 1_S(x_1) dx_1}{x_2}$$

for almost all $x_2 \in [0,1]$ w.r.t. μ_{X_2}, and thus,

$$\int_0^y \frac{1}{2} dx = \int_0^y 1_S(x) dx$$

for all $y \in [0,1]$.[11] By the usual arguments (which includes the use of monotone class lemma and monotone convergence theorem), we have

$$\int_A \frac{1}{2} dx = \int_A 1_S(x) dx$$

for every Borel subset $A \subset [0,1]$, and thus, again by the uniqueness of the Radon-Nikodym derivative, we have

$$\frac{1}{2} = 1_S(x) \text{ for almost all } x \text{ w.r.t. the Lebesgue measure,}$$

which is absurd.

A.2 Proof of Proposition 1

First, we will show the following lemma:

Lemma 1. If μ is a probability measure defined on $[0,1]^2$ and the conditional probability $\mu(\cdot|x_2)$ has a density function $h(\cdot|x_2)$ for almost all x_2 w.r.t. μ_{X_2}, then

$$\mu(S_m|x_2) \to p \text{ for almost all } x_2 \text{ w.r.t. } \mu_{X_2}. \tag{4}$$

[11] Note that the Lebesgue measure is absolutely continuous w.r.t. μ_{X_2}, and the both side of above equation are continuous in y.

Proof of Lemma 1. Fix any $\varepsilon > 0$. Let $h_{x_2}(x_1) = h(x_1|x_2)$ if it can be defined. Because h_{x_2} is integrable, there exists a continuous function $\varphi : [0,1] \to \mathbb{R}$ such that[12]

$$\int_0^1 |\varphi(x) - h_{x_2}(x)|dx < \varepsilon.$$

Because $h_{x_2}(x)$ is a density function of a probability measure, we have

$$\left|\int_0^1 \varphi(x)dx - 1\right| = \left|\int_0^1 (\varphi(x) - h_{x_2}(x))dx\right| \le \int_0^1 |\varphi(x) - h_{x_2}(x)|dx < \varepsilon. \quad (5)$$

On the other hand,

$$\begin{aligned}\left|\int_{S_m} \varphi(x)dx - \mu(S_m|x_2)\right| &= \left|\int_{S_m} (\varphi(x) - h_{x_2}(x))dx\right| \\ &\le \int_0^1 |\varphi(x) - h_{x_2}(x)|dx < \varepsilon. \quad (6)\end{aligned}$$

Because ϕ is continuous, it is Riemann integrable on $[0,1]$. For $k = 1, ..., 2m$, define

$$\Delta_k^m = \left[\frac{k-1}{2m}, \frac{k-1+2p}{2m}\right],\ \xi_k^m = \frac{k-1+2p}{2m}$$

if k is odd, and

$$\Delta_k^m = \left[\frac{k-2+2p}{2m}, \frac{k}{2m}\right],\ \xi_k^m = \frac{k-2+2p}{2m}$$

if k is even. Then, the Riemann sum is

$$s(\phi, \Delta^m, \xi^m) = \sum_{k=1}^m \varphi(\xi_{2k-1}^m)\frac{p}{m} + \sum_{k=1}^m \varphi(\xi_{2k}^m)\frac{1-p}{m} = \frac{1}{m}\sum_{k=1}^m \varphi(\xi_{2k-1}^m),$$

and thus, if we define

$$\varphi_m = \sum_{k=1}^m \varphi(\xi_{2k-1}^m) 1_{\Delta_{2k-1}^m},$$

then

$$\int_0^1 \varphi_m(x)dx = ps(\phi, \Delta^m, \xi^m) \to p\int_0^1 \varphi(x)dx,$$

[12] See Theorem 3.14 of Rudin (1987).

as $m \to \infty$. Therefore, for sufficiently large m, we have

$$\left| \int_0^1 (\varphi_m(x) - p\varphi(x))dx \right| < \varepsilon, \tag{7}$$

and thus, by (5) and (7),

$$\left| \int_0^1 \varphi_m(x)dx - p \right| \leq \left| \int_0^1 (\varphi_m(x) - p\varphi(x))dx \right| + \left| \int_0^1 p[\varphi(x) - 1]dx \right| < 2\varepsilon. \tag{8}$$

Because φ is continuous on the compact set $[0, 1]$, it is also uniformly continuous, and thus

$$\sup_{x \in [0,1]} |\varphi(x)1_{S_m}(x) - \varphi_m(x)| < \varepsilon$$

for sufficiently large m. Therefore,

$$\begin{aligned} \left| \int_{S_m} \varphi(x)dx - \int_0^1 \varphi_m(x)dx \right| &= \left| \int_0^1 (\varphi(x)1_{S_m}(x) - \varphi_m(x))dx \right| \\ &\leq \int_0^1 |\varphi(x)1_{S_m}(x) - \varphi_m(x)|dx \leq \varepsilon. \end{aligned} \tag{9}$$

By (6), (8), and (9), we have

$$\begin{aligned} &|\mu(S_m|x_2) - p| \\ \leq &\left| \mu(S_m|x_2) - \int_{S_m} \varphi(x)dx \right| + \left| \int_{S_m} \varphi(x)dx - \int_0^1 \varphi_m(x)dx \right| + \left| \int_0^1 \varphi_m(x)dx - p \right| \\ < &4\varepsilon, \end{aligned}$$

as desired. This completes the proof of Lemma 1. ■

Let

$$a_i^1 = u_i(0, 0),\ a_i^2 = u_i(0, 1),\ a_i^3 = u_i(1, 0),\ a_i^4 = u_i(1, 1).$$

Then, for every strategy k_2 of player 2,

$$\begin{aligned} |U_i(k_1^m, k_2) - U_i(k_1, k_2)| \leq &\left| \int_{X_2} (\mu(S_m|x_2) - p)(p_2^0(x_2)a_i^3 + p_2^1(x_2)a_i^4)d\mu_{X_2} \right| \\ &+ \left| \int_{X_2} (p - \mu(S_m|x_2))(p_2^0(x_2)a_i^1 + p_2^1(x_2)a_i^2)d\mu_{X_2} \right| \\ \leq &\int_{X_2} |\mu(S_m|x_2) - p| \times \max_j |a_i^j| d\mu_{X_2} \to 0 \text{ as } m \to \infty \end{aligned}$$

by dominated convergence theorem, as desired. This completes the proof. ■

A.3 Proof of Proposition 2

As in the proof of Proposition 1, it suffices to show the following lemma.

Lemma 2. Suppose that μ is a probability measure on $[0,1]^2$ and the conditional probability $\mu(\cdot|x_2)$ is atomless for almost all x_2 w.r.t. μ_{X_2}. Then, there exists a subsequence $(S_{k(n)})$ of (S_m) such that

$$\mu(S_{k(n)}|x_2) - p \to 0 \text{ as } m \to \infty \text{ for almost all } x_2 \text{ w.r.t. } \mu_{X_2}. \tag{10}$$

Proof of Lemma 2. Define

$$f^m(x) = 1_{S_m}(x) - p$$

on $[0,1[$. If $f : \mathbb{R} \to \mathbb{R}$ is a periodic function with period 1 and $f(x) = f^1(x)$ for $x \in [0,1[$, then

$$f^m(x) = f(mx).$$

Define

$$c_m(x_2) = \int_0^1 f(mx)\mu_{x_2}(x),$$

where $\mu_{x_2}(\cdot) = \mu(\cdot|x_2)$. It suffices to show that there exists an increasing sequence $k : \mathbb{N} \to \mathbb{N}$ such that the function $c_{k(n)} \to 0$ a.s. w.r.t. μ_{X_2}. As is well known, any sequence of functions that converges to zero in L^1 has a subsequence that converges to zero a.s.. Hence, it suffices to show that there exists an increasing sequence $m : \mathbb{N} \to \mathbb{N}$ such that

$$\int |c_{m(n)}(x_2)| d\mu_{X_2}(x_2) \to 0. \tag{11}$$

Choose any small $\varepsilon > 0$, and define $T = \cup_{i \in \mathbb{Z}}[i, i+p]$ and

$$f^{\varepsilon}(x) = \max\{0, 1 - \varepsilon^{-1} \inf\{|y - x| | y \in T\}\} - p,$$

$$f^{-\varepsilon}(x) = \max\{0, 1 - \varepsilon^{-1} \inf\{|y - x| | \inf\{|z - y| | z \in T^c\} \geq \varepsilon\}\} - p.$$

Then, both f^{ε} and $f^{-\varepsilon}$ are periodic with period 1. Figures 1-2 depicts both functions $f^{\pm\varepsilon}$, where $p = 0.4$ and $\varepsilon = 0.1$.

Let g^{ε} (resp. $g^{-\varepsilon}$) be a trigonometric polynomial that uniformly approximates f^{ε} (resp. $f^{-\varepsilon}$), that is,[13]

$$\|f^{\pm\varepsilon} - g^{\pm\varepsilon}\|_{\infty} \equiv \sup_{x \in \mathbb{R}} |f^{\pm\varepsilon}(x) - g^{\pm\varepsilon}(x)| < \varepsilon.$$

[13] For the existence of such trigonometric polynomials, use Stone-Weierstrass's theorem. See Theorem 2.4.11 of Dudley (2002).

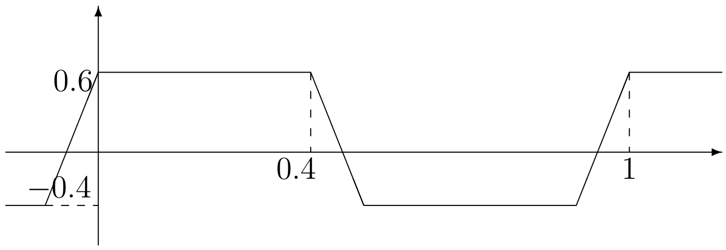

Figure 1: Graph of f^{ε}

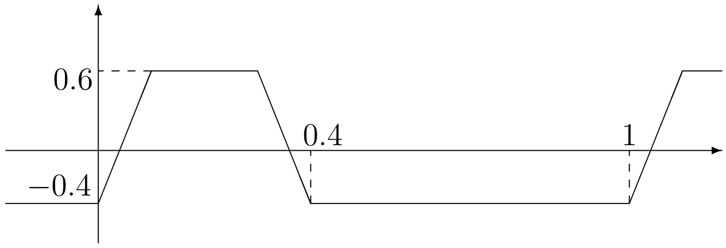

Figure 2: Graph of $f^{-\varepsilon}$

The following equation expresses $g^{\pm\varepsilon}$ concretely:

$$g^{\pm\varepsilon}(x) = \sum_{m\in\mathbb{Z}} a_m^{\pm\varepsilon} e^{2\pi mix},$$

where $a_m^{\pm\varepsilon} = 0$ for all m with sufficiently large $|m|$. We can obtain

$$\left|\int_0^1 f^{\pm\varepsilon}(x)dx - \int_0^1 g^{\pm\varepsilon}(x)dx\right| \leq \int_0^1 |f^{\pm\varepsilon}(x) - g^{\pm\varepsilon}(x)|dx \leq \varepsilon.$$

Because

$$\left|\int_0^1 f^{\pm\varepsilon}(x)dx\right| = \varepsilon,\ \int_0^1 g^{\pm\varepsilon}(x)dx = a_0^{\pm\varepsilon},$$

we have $|a_0^{\pm\varepsilon}| \leq 2\varepsilon$. Define

$$c_m^{\pm\varepsilon}(x_2) = \int_0^1 f^{\pm\varepsilon}(mx)d\mu_{x_2}(x),$$

$$d_m^{\pm\varepsilon}(x_2) = \int_0^1 g^{\pm\varepsilon}(mx)d\mu_{x_2}(x).$$

By the Cauchy-Schwarz inequality,

$$\frac{1}{2M+1}\sum_{m=-M}^{M}|\hat{\mu}_{x_2}(m)| \leq \sqrt{\frac{1}{2M+1}\sum_{m=-M}^{M}|\hat{\mu}_{x_2}(m)|^2}, \tag{12}$$

where $\hat{\mu}_{x_2}$ is the Fourier transform of μ_{x_2}, i.e.,

$$\hat{\mu}_{x_2}(m) = \int_0^1 e^{-2\pi imx} d\mu_{x_2}(x).$$

By Wiener's theorem,[14] for every atomless Radon measure ν on $]0,1]$,

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{j=-N}^{N}|\hat{\nu}(j)|^2 = 0,$$

where $\hat{\nu}(j) = \int_0^1 e^{-2\pi ijx}d\nu(x)$. Applying this theorem, we have that the right-hand side of (12) converges to zero as $M \to \infty$ whenever μ_{x_2} is atomless. Thus,

$$\frac{1}{2M+1}\sum_{m=-M}^{M}|d_m^{\pm\varepsilon}(x_2) - a_0^{\pm\varepsilon}| \to 0 \text{ as } M \to \infty,$$

for every x_2 such that μ_{x_2} is atomless. Because

$$\sum|c_m^{\pm\varepsilon}(x_2)| \leq \sum|c_m^{\pm\varepsilon}(x_2) - d_m^{\pm\varepsilon}(x_2)| + |d_m^{\pm\varepsilon}(x_2)|,$$

we have

$$\limsup_{M\to\infty}\frac{1}{2M+1}\sum_{m=-M}^{M}|c_m^{\pm\varepsilon}(x_2)| \leq 3\varepsilon,$$

for every x_2 such that μ_{x_2} is atomless. Since $f^\varepsilon \geq f \geq f^{-\varepsilon}$, we have $c_m^\varepsilon(x_2) \geq c_m(x_2) \geq c_m^{-\varepsilon}(x_2)$. Hence, $|c_m(x_2)| \leq |c_m^\varepsilon(x_2)| + |c_m^{-\varepsilon}(x_2)|$, and thus, for every such x_2, we have

$$\limsup_{M\to\infty}\frac{1}{2M+1}\sum_{m=-M}^{M}|c_m(x_2)| \leq 6\varepsilon,$$

which implies that,

$$\lim_{M\to\infty}\frac{1}{2M+1}\sum_{m=-M}^{M}|c_m(x_2)| = 0.$$

[14] See section 1.7 or 6.2 of Katznelson (2004).

Because $|c_m(x_2)| \leq 1$, by dominated convergence theorem, we have

$$\lim_{M\to\infty} \frac{1}{2M+1} \sum_{m=-M}^{M} \int_0^1 |c_m(x_2)| d\mu_{X_2}(x_2) = 0. \tag{13}$$

Now, define

$$a_m = \int_0^1 [|c_m(x_2)| + |c_{-m}(x_2)|] d\mu_{X_2}(x_2).$$

If there is no subsequence $(a_{m(n)})$ that satisfies $\lim_{n\to\infty} a_{m(n)} = 0$, then there exists $\varepsilon > 0$ and N_0 such that for every $m > N_0$, $a_m \geq \varepsilon$. Therefore, for every sufficiently large N,

$$\frac{1}{N} \sum_{m=1}^{N} a_m \geq \frac{\varepsilon}{2},$$

which contradicts (13). Thus, there must exist a subsequence $(a_{m(n)})$ such that $\lim_{n\to\infty} a_{m(n)} = 0$, and consequently,

$$0 \leq \int_0^1 |c_{m(n)}(x_2)| d\mu_{X_2}(x_2) \leq a_{m(n)} \to 0,$$

which implies (11). This completes the proof of Lemma 2 and Proposition 2. ■

A.4 Proof of Proposition 3

Because every X_i is a Polish space with continuum cardinarity, we can assume that $X_i = [0, 1]$ without loss of generality. Suppose that ν is a probability measure on $[0, 1]$. First, if ν is atomless, then $\nu(\{x\}) = 0$ for every $x \in [0, 1]$. Second, suppose that $\nu(\{x\}) = 0$ for every $x \in [0, 1]$, and define $F(x) = \nu([0, x])$. Then, $F(0) = 0$ and F is continuous. Because $[0, 1]$ is compact, $F(0) = 0$ and F is uniformly continuous. Third, let $F(x) = \nu([0, x])$ satisfy that $F(0) = 0$ and be uniformly continuous. Then, for every Borel set $B \subset [0, 1]$,

$$\max_{i=1,\ldots,m} \nu\left(B \cap \left[\frac{i-1}{m}, \frac{i}{m}\right]\right) \leq F\left(\frac{i}{m}\right) - F\left(\frac{i-1}{m}\right) \to 0 \text{ as } m \to \infty,$$

which implies that ν is atomless. To summarize the above arguments, the following three conditions are equivalent.

1. ν is atomless.

2. $\nu(\{x\}) = 0$ for every $x \in [0,1]$.

3. If $F(x) = \nu([0,x])$, then $F(0) = 0$ and F is uniformly continuous.

Now, define $\mu_j = \mu_{X_j}$ and $\mu_{-i} = \mu_{X_{-i}}$. It suffices to show that if $\mu(\cdot|x_{-i})$ is atomless for almost all x_{-i} w.r.t. μ_{-i}, then $\mu_{ij}(\cdot|x_j)$ is atomless for almost all x_j w.r.t. μ_j.

Define

$$\varphi_k(x_{-i}) = \sup_{0<x_i'-x_i<\frac{1}{k}, x_i, x_i' \in [0,1]\cap\mathbb{Q}} \mu([x_i, x_i']|x_{-i}),$$

$$\psi_k(x_j) = \sup_{0<x_i'-x_i<\frac{1}{k}, x_i, x_i' \in [0,1]\cap\mathbb{Q}} \mu_{ij}([x_i, x_i']|x_j),$$

and

$$\varphi = \inf_{k\in\mathbb{N}} \varphi_k, \ \psi = \inf_{k\in\mathbb{N}} \psi_k.$$

Then, all of the above functions are measurable. By the above arguments, we have that $\mu(\cdot|x_{-i})$ is atomless for almost all x_{-i} w.r.t. μ_{-i} if and only if

$$\int_{X_{-i}} \varphi(x_{-i}) d\mu_{-i} = 0,$$

and $\mu_{ij}(\cdot|x_j)$ is atomless for almost all x_j w.r.t. μ_j if and only if

$$\int_{X_j} \psi(x_j) d\mu_j = 0.$$

Therefore, it suffices to show that

$$\int_{X_{-i}} \varphi(x_{-i}) d\mu_{-i} = 0 \Rightarrow \int_{X_j} \psi(x_j) d\mu_j = 0. \tag{14}$$

Choose two Borel subset $B, D \subset [0,1]$. Then,

$$\begin{aligned}\mu\left(B \times D \times \prod_{\ell\neq i,j} X_\ell\right) &= \int_{D\times\prod_{\ell\neq i,j} X_\ell} \mu(B|x_{-i}) d\mu_{-i}(x_{-i}) \\ &= \int_D \mu_{ij}(B|x_j) d\mu_j(x_j) \\ &= \int_{D\times\prod_{\ell\neq i,j} X_\ell} \mu_{ij}(B|\mathrm{proj}_j(x_{-i})) d\mu_{-i}(x_{-i}),\end{aligned}$$

by Theorem 4.1.11 of Dudley (2002), where $\mathrm{proj}_j : X_{-i} \to X_j$ is the projection mapping; that is,

$$\mathrm{proj}_j(x_{-i}) = x_j.$$

Therefore, we have

$$E(\mu(B|\cdot)|\mathrm{proj}_j) = \mu_{ij}(B|\mathrm{proj}_j(\cdot)).$$

By monotone convergence theorem of conditional expectation, we have

$$\begin{aligned}
0 &\le \int_{X_j} \psi(x_j) d\mu_j(x_j) \\
&= \int_{X_{-i}} \psi(\mathrm{proj}_j(x_{-i})) d\mu_{-i}(x_{-i}) \\
&= \int_{X_{-i}} \inf_{k\in\mathbb{N}} \left(\sup_{0<x_i'-x_i<\frac{1}{k}, x_i, x_i' \in [0,1]\cap\mathbb{Q}} \mu_{ij}([x_i, x_i']|\mathrm{proj}_j(x_{-i})) \right) d\mu_{-i}(x_{-i}) \\
&= \int_{X_{-i}} \inf_{k\in\mathbb{N}} \left(\sup_{0<x_i'-x_i<\frac{1}{k}, x_i, x_i' \in [0,1]\cap\mathbb{Q}} E(\mu([x_i, x_i']|\cdot)|\mathrm{proj}_j)(x_{-i}) \right) d\mu_{-i}(x_{-i}) \\
&\le \int_{X_{-i}} \inf_{k\in\mathbb{N}} E(\varphi_k|\mathrm{proj}_j)(x_{-i}) d\mu_{-i}(x_{-i}) \\
&= \int_{X_{-i}} E(\varphi|\mathrm{proj}_j)(x_{-i}) d\mu_{-i}(x_{-i}) \\
&= \int_{X_{-i}} \varphi(x_{-i}) d\mu_{-i}(x_{-i}) = 0,
\end{aligned}$$

which implies (14). This completes the proof. ■

A.5 Proof of Claims in Example 2

First, we confirm the definition of the Prohorov metric. Let (X, ρ_X) be a metric space and P, Q be Borel probability measures defined on X. For every $A \subset X$ and $\varepsilon > 0$, let $A^\varepsilon = \cup_{x\in A}\{y \in X|\rho_X(x, y) \le \varepsilon\}$. Define

$$\rho(P, Q) = \inf\{\varepsilon|P(A) \le Q(A^\varepsilon) + \varepsilon \text{ for all Borel set } A\}.$$

This ρ is known to satisfy all requirements of a metric, and is called the Prohorov metric. If X is a second-countable complete metric space, then ρ is also complete.[15]

In our case, θ_m is a probability measure on $]0, 1]^2$, which is a second-countable complete metric space. Thus, to verify the convergence of (θ_m), it suffices to show that it is a Cauchy sequence.

[15] See Corollary 11.5.5 of Dudley (2002).

Second, we note that by definition of θ_m, we have that if $C^{m+1}_{i'j'} \subset C^m_{ij}$,

$$\theta_m(C^{m+1}_{i'j'}) = \begin{cases} \frac{\theta_m(C^m_{ij})}{8} & \text{if } C^m_{ij} \subset A_m, i'+j' \text{ is even,} \\ & \text{or } C^m_{ij} \not\subset A_m, i'+j' \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

We will show that $\rho(\theta_m, \theta_{m+1}) \leq \frac{1}{4^{m+1}}$, which implies that (θ_m) is actually a Cauchy sequence. Choose any Borel set $A \subset]0,1]^2$ and let $A^m_{ij} = A \cap C^m_{ij}$ and $B^m_{ij} = (A^m_{ij})^{\frac{1}{4^{m+1}}} \cap C^m_{ij}$. It suffices to show that $\theta_m(A^m_{ij}) \leq \theta_{m+1}(B^m_{ij})$. If $\theta_m(C^m_{ij}) = 0$, then it is clear. Therefore, we assume that $\theta_m(C^m_{ij}) = \frac{1}{8^m}$. Let $C^{m+1}_{i'j'} \subset C^m_{ij}$ and $\theta_{m+1}(C^{m+1}_{i'j'}) > 0$. If i' is odd, then $k' = i' - 1$ is even and $C^{m+1}_{k'j'} \subset C^m_{ij}$. In this case, B^m_{ij} includes $D^{m+1}_{k'j'} = A^{m+1}_{k'j'} + \left(\frac{1}{4^{m+1}}, 0\right)$. If i' is even, then $k' = i' + 1$ is odd and $C^{m+1}_{k'j'} \subset C^m_{ij}$. In this case, B^m_{ij} includes $D^{m+1}_{k'j'} = A^{m+1}_{k'j'} - \left(\frac{1}{4^{m+1}}, 0\right)$. In both cases,

$$\begin{aligned} \theta_m(A^{m+1}_{i'j'}) + \theta_m(A^{m+1}_{k'j'}) &= \int_{A^{m+1}_{i'j'}} 2^m dx + \int_{A^{m+1}_{k'j'}} 2^m dx \\ &= \int_{A^{m+1}_{i'j'}} 2^m dx + \int_{D^{m+1}_{k'j'}} 2^m dx \\ &\leq \int_{B^m_{ij} \cap C^{m+1}_{i'j'}} 2^{m+1} dx = \theta_{m+1}(B^m_{ij} \cap C^{m+1}_{i'j'}). \end{aligned}$$

Thus, we have

$$\theta_m(A^m_{ij}) \leq \theta_{m+1}(B^m_{ij}),$$

as desired.

Therefore, we have $\theta_m \to \theta^*$ as $m \to \infty$. Next, we will show that

$$\theta^*(C^m_{ij}) = \theta_m(C^m_{ij}). \quad (16)$$

Note that, by definition of C^m_{ij}, we have

$$\theta_{m+k}(C^m_{ij}) = \theta_m(C^m_{ij})$$

for every $k \geq 0$.

First, suppose that $\theta_m(C^m_{ij}) = \frac{1}{8^m}$. Define

$$g^k(x) = \max\{0, 1 - 4^{m+k} \inf\{\|y - x\| | y \in C^m_{ij}\}\},$$

$$h^k(x) = \max\{0, 1 - 4^{m+k} \inf\{\|y - x\| | \inf\{\|z - y\| | z \notin C^m_{ij}\} \geq 4^{-m-k}\}\}.$$

Then, clearly g^k, h^k are bounded continuous functions on $]0,1]^2$, and

$$\begin{aligned}\theta_m(C_{ij}^m) - \frac{4 \times 4^k}{8^{m+k}} &\leq \int h^k(x) d\theta_{m+\ell} \\ &\leq \theta_m(C_{ij}^m) \\ &\leq \int g^k(x) d\theta_{m+\ell} \\ &\leq \theta_m(C_{ij}^m) + \frac{4 \times (4^k + 1)}{8^{m+k}},\end{aligned}$$

for every k, ℓ with $\ell \geq k$. Therefore,

$$\begin{aligned}\theta_m(C_{ij}^m) - \frac{4 \times 4^k}{8^{m+k}} &\leq \int h^k(x) d\theta^* \\ &\leq \int g^k(x) d\theta^* \\ &\leq \theta_m(C_{ij}^m) + \frac{4 \times (4^k + 1)}{8^{m+k}},\end{aligned}$$

and thus,

$$\theta^*(\text{int. } C_{ij}^m) = \theta^*(\text{cl. } C_{ij}^m) = \theta_m(C_{ij}^m),$$

which implies (16) in this case. Because the number of such (i,j) is exactly 8^m, we have if $\theta_m(C_{ij}^m) = 0$, then $\theta^*(C_{ij}^m) = 0$, and (16) holds.

We shall show that θ is weakly conditionally atomless for player 1. We write $]0,1]^2 = Y_1 \times Y_2$, where $Y_i =]0,1]$. It suffices to show that $\theta^*(\cdot|y_2)$ is atomless for almost all y_2 w.r.t. $\theta^*_{Y_2}$ and $\theta^*(\cdot|y_1)$ is atomless for almost all y_1 w.r.t. $\theta^*_{Y_1}$. We treat only the former, because the proof of the latter is symmetrical. First, we call $Q \in \mathscr{Q}_m$ an atom of $\mathscr{Q}_m$ if $Q =]\frac{i}{4^m}, \frac{i+1}{4^m}]$ for some i. We will show that if Q is an atom of $\mathscr{Q}_m$ and $B \in \mathscr{Q}_\ell$, then

$$\theta^*(B \times Q) \leq 2^{\min\{\ell,m\}} \lambda(B)\lambda(Q) \leq \lambda(B)\sqrt{\lambda(Q)}, \tag{17}$$

where λ is the Lebesgue measure. By additivity of both sides in B, we can assume that B is also an atom of $\mathscr{Q}_\ell$. If $\ell = m$, then $B \times Q = C_{ij}^m$ for some i, j, and thus

$$\theta^*(B \times Q) \leq \frac{1}{8^m} = 2^m \times 4^{-m} \times 4^{-m} = 2^m \lambda(B)\lambda(Q),$$

as desired. If $\ell < m$, then $B \times Q$ is a finitely many union of Cartesian products of atoms in $\mathscr{Q}_m$ and $2^{m-\ell}$ of these have positive measure w.r.t. θ^*, and thus,

$$\begin{aligned}\theta^*(B \times Q) &= 2^{m-\ell} \times 8^{-m} \\ &= 2^{m-\ell} \times 2^m \times 4^{\ell-m} \times 4^{-\ell} \times 4^{-m} \\ &= 2^\ell \lambda(B)\lambda(Q),\end{aligned}$$

as desired. The proof of the case $\ell > m$ is symmetrical.

Particularly,

$$\theta^*_{Y_2}(B) = \lambda(B)$$

if $B \in \mathscr{Q}_\ell$ for some ℓ. Because the algebra $\cup_\ell \mathscr{Q}_\ell$ is a base of the Borel σ-algebra, By monotone class lemma, we have

$$\theta^*_{Y_2} = \lambda.$$

Define

$$A = \left\{ y_2 \in]0,1] \,\middle|\, \lim_{m\to\infty} \max_{i\in\{0,\dots,4^m-1\}} \theta^* \left(\left] \frac{i}{4^m}, \frac{i+1}{4^m} \right] \,\middle|\, y_2 \right) > 0 \right\}.$$

If $\lambda(A) > 0$, then there exists $\delta > 0$ such that $\lambda(A_\delta) > 0$, where

$$A_\delta = \left\{ y_2 \in]0,1] \,\middle|\, \lim_{m\to\infty} \max_{i\in\{0,\dots,4^m-1\}} \theta^* \left(\left] \frac{i}{4^m}, \frac{i+1}{4^m} \right] \,\middle|\, y_2 \right) > \delta \right\}.$$

Therefore, for every m, there exists i_m such that $\lambda(A^m_\delta) > 0$, where

$$A^m_\delta = \left\{ y_2 \in]0,1] \,\middle|\, \theta^* \left(\left] \frac{i_m}{4^m}, \frac{i_m+1}{4^m} \right] \,\middle|\, y_2 \right) > \delta \right\}.$$

If $Q_m = \left] \frac{i_m}{4^m}, \frac{i_m+1}{4^m} \right]$, then by (17) and monotone class lemma,

$$\theta^*(A^m_\delta \times Q_m) \le \lambda(A^m_\delta)\sqrt{\lambda(Q_m)}.$$

On the other hand,

$$\theta^*(A^m_\delta \times Q_m) \ge \int_{A_m} \delta d\theta_{Y_2} = \delta\lambda(A^m_\delta),$$

which is absurd for sufficiently large m. Therefore, $\lambda(A) = 0$. By the same argument as in the first paragraph of the proof of Proposition 3, we have $\theta^*(\cdot|y_2)$ is atomless for almost all y_2 w.r.t. $\theta^*_{Y_2}$.

Next, let k_1 be the simple $(1/2, 1/2)$-mixed strategy, and $k_1' = 1_A$, where $A = A_m$ for some m. Suppose that k_1' is 0.1-purification of k_1. Choosing $k_2 \equiv 1$ and $k_3 \equiv 1$, we have

$$|U_1(k_1, k_2, k_3) - U_1(k_1', k_2, k_3)| = \left|\theta^*(A_m) - \frac{1}{2}\right| < 0.1,$$

and thus, $\theta^*(A_m) > 0.4$. Next, let B, C be the union of $\left]\frac{i}{4^{m+1}}, \frac{i+1}{4^{m+1}}\right]$ for odd i, and let $k_2 = 1_B$ and $k_3 = 1_C$. If D is a product of atoms of $\mathscr{Q}_m$, then

$$\theta(D \times B \times C) = \begin{cases} \frac{\theta^*(D)}{2} & \text{if } D \subset A_m, \\ 0 & \text{otherwise,} \end{cases}$$

and thus,

$$\theta(A_m \times B \times C) = \frac{1}{2}\theta^*(A_m),$$
$$\theta(A_m^c \times B \times C) = 0.$$

Therefore,

$$|U_1(k_1, k_2, k_3) - U_1(k_1', k_2, k_3)| = \frac{1}{4}\theta^*(A_m) > 0.1,$$

a contradiction.

Finally, suppose that $k_1' = 1_A$ for some Borel set A. Then, there exists a set A_m such that $\theta^*(A \Delta A_m) < 0.05$. In this case, we can easily check that k_1' is not 0.05-purification of k_1. This completes the proof. ■

References

[1] Aumann, R. J., Katznelson, Y., Radner, R., Rosenthal, R. W., and Weiss, B. (1983) "Approximate Purification of Mixed Strategies". *Mathematics of Operations Research* 8, 327-341.

[2] Dudley, R. M. (2002) *Real Analysis and Probability*, 2nd ed. Cambridge University Press, New York.

[3] Katznelson, Y. (2004) *An Introduction to Harmonic Analysis*, 3rd ed. Cambridge University Press, New York.

[4] Khan, M. A. and Rath, K. P. (2009) "On Games with Incomplete Information and the Dvoretsky-Wald-Wolfowitz Theorem with Countable Partitions". *Journal of Mathematical Economics* 45, 830-837.

[5] Khan, M. A., Rath, K. P., and Sun, Y. (2006) "The Dvoretzky-Wald-Wolfowitz Theorem and Purification in Atomless Finite-Action Games". *International Journal of Game Theory* 34, 91-104.

[6] Khan, M. A. and Zhang, Y. (2014) "On the Existence of Pure-Strategy Equilibria in Games with Private Information: A Complete Characterization". *Journal of Mathematical Economics* 50, 197-202.

[7] Radner, R. and Rosenthal, R. W. (1982) "Private Information and Pure-Strategy Equilibria". *Mathematics and Operations Research* 7, 401-409.

[8] Rudin, W. (1987) *Real and Complex Analysis*, 3rd ed. McGraw-Hill, Singapole.

[9] Srivastava, S. M. (1998) *A Course on Borel Sets*. Springer-Verlag, New York.