Purification of Bayesian-Nash Equilibria in Large Games with Compact Type and Action Spaces*

Guilherme Carmona†
Universidade Nova de Lisboa

August 8, 2008

Abstract

We present a purification result for incomplete information games with a large but finite number of players that allows compact metric spaces for both actions and types. We then compare our framework and findings to the early purification theorems of Rashid (1983), Cartwright and Wooders (2002), Kalai (2004) and Wooders, Cartwright, and Selten (2006). Our proofs are elementary and rely on the Shapley-Folkman theorem.

Journal of Economic Literature Classification Numbers: C72.

Keywords: Bayesian games, Bayesian Nash equilibrium, pure strategies, purification.

*I wish to thank Erik Balder, Mehmet Barlo, Iliyan Georgiev, Konrad Podczeck, Yeneng Sun, Myrna Wooders, Nicholas Yannelis, an anonymous referee, an associate editor, Nizar Allouch, and participants at the XVI European Workshop on General Equilibrium for very helpful comments. I thank also John Huffstot for editorial assistance. Any remaining error is, of course, my own.

†Address: Universidade Nova de Lisboa, Faculdade de Economia, Campus de Campolide, 1099-032 Lisboa, Portugal; Phone: (351) 21 380 1672; Fax: (351) 21 387 0933; email: gcarmona@fe.unl.pt.
1 Introduction

Mixed strategies are usually regarded as unappealing because they are not only hard to interpret, but also, considered as too complex for real players to use. Motivated by this view, game-theorists have provided several purification theorems that describe when mixed strategies can be replaced by (approximately) equivalent pure strategies.

Several purification results have been obtained for games with a large number of players. Such results include those of Rashid (1983) and Wooders, Cartwright, and Selten (2006) for complete information games with finite action spaces and Cartwright and Wooders (2002) for incomplete information games with finite type and countable action spaces.\textsuperscript{1} All these results show that if the number of players in a game is sufficiently large, we can associate each Bayesian Nash equilibrium with a pure strategy approximate equilibrium such that the two strategies yield approximately the same payoff to each type of each player.

However, there are many economic examples that do not satisfy the above assumptions. An example is provided by the following Cournot oligopoly game. In such a game, \( n \) firms produce a nonnegative quantity of a certain good, which due to capacity constraints, cannot exceed a given level \( m \). If firms produce quantities \( x_1, \ldots, x_n \), then the market price of the good is \( P(\sum_{i=1}^{n} x_i/n) \), implying that the revenue of firm \( i \) is \( P(\sum_{i=1}^{n} x_i/n)x_i \). The cost for each firm of producing a quantity \( x \) is \( C(t, x) \), where \( t \) is the firm’s type and can be understood as its productivity level. The payoff (i.e., profit) of firm \( i \) when its type is \( t \) and its quantity is \( x_i \) is then \( P(\sum_{i=1}^{n} x_i/n)x_i - C(t, x_i) \). Since the action space of each firm is the interval \([0, m]\), and so uncountable, none of the above purification results can be applied to this game. Our main result shows that if both functions \( P \) and \( C \) are continuous and each firm’s type is commonly known (i.e., the game is of complete information), then all Nash equilibria of this game can be approximately purified when the number of

\textsuperscript{1}In Wooders, Cartwright, and Selten (2006), each player’s payoff depends on the (empirical) distribution over types and actions, where players’ types, although being deterministic, can belong to a compact type space. In contrast, in Rashid (1983), each player’s payoff depends, besides on his type and action, only on the (average) distribution over actions.
firms is sufficiently large. Furthermore, if each firm’s type is its own private information (i.e., the game is of incomplete information), then the same conclusion holds whenever type spaces are compact and firms’ types are determined independently.

More generally, we consider incomplete information games with finite number of players in which each player’s payoff depends only on his type-action character and either on the average or the empirical distribution of type-action characters of the others. Our purification result shows that if players’ types are independent and their payoff functions are selected from an equicontinuous family, then all Bayesian Nash equilibria of sufficiently large games can be approximately purified.

Our contribution is to provide a purification result that allows for incomplete information games with general compact type and action spaces, a case that has not been dealt with in the literature. Furthermore, we show how our result generalizes those of Rashid (1983) and Wooders, Cartwright, and Selten (2006) both by allowing for incomplete information (with independent types drawn from a compact type space) and a compact action space. However, our result is neither implied nor does it imply the purification theorem of Cartwright and Wooders (2002) as explained in Subsection 5.4 below. Furthermore, we do not consider the concept of self-purification as defined in Kalai (2004). Since self-purification is a stronger requirement than the purification notion we consider (as shown by Cartwright and Wooders (2004)), it follows that our result, although it applies to the setting of Kalai (2004), does not imply his self-purification result. However, since Kalai’s result assumes that both the type and action spaces are finite, our result is also not implied by his.

We wish to emphasize that, although our result differs from the self-purification theorem of Kalai (2004), our framework relates to his. Indeed, in both frameworks, players’ types are drawn independently and each player’s payoff function is selected from an equicontinuous family and depends on his type, action and on the empirical distribution over types and actions. The difference is that, while Kalai (2004) assumes that both the type and action spaces are finite, we allow them to be an arbitrarily compact metric space. A similar comparison can be made regarding the framework of Rashid (1983). Indeed, our framework reduces to that of Rashid (1983) when
we specialize the type space to be a singleton, the action space to be finite and by considering payoff functions that depend on the average distribution over types and actions.

The relationship between our framework and that of Wooders, Cartwright, and Selten (2006) is less direct and will be discussed in detail in Subsection 5.3. In a sense to be made precise, their framework correspond to our framework when players’ types are deterministic, the action space is finite and all players have the same payoff function, which depends on the empirical distribution over types and actions. Finally, as we noted above, there is no formal relationship between our framework and that of Cartwright and Wooders (2002). In fact, we allow for uncountable type and action spaces whereas they allow for noncompact action spaces (a more detailed discussion is provided in Subsection 5.4).

Our approach requires only elementary arguments. We start by establishing a purification result for finitely supported approximate equilibria in games with finitely supported distributions over types. The proof of this result is based on the Shapley-Folkman theorem (see Starr (1969)) in a way similar to that given by Rashid (1983). This basic result is then extended to the general case of games with compact metric spaces of types and actions.

The paper is organized as follows: In Section 2, we present our notation and definitions. Our purification result is presented in Section 3 and its proof in Subsection 3.1. In Section 4 we show that our result applies to the Cournot oligopoly examples described above. Finally, on Section 5, we relate our framework and result to those of Rashid (1983), Cartwright and Wooders (2002), Kalai (2004) and Wooders, Cartwright, and Selten (2006).

2 Notation and Definitions

We start by presenting two classes of Bayesian games. In both of them, each player’s payoff function depends on his type, his action and on a distribution over types and actions induced by the choice of the other players. They differ because we consider
the average distribution in one class and the empirical distribution in the other.

We start by describing the elements that are common to the two classes. In both classes, all players have a common type space $T$ and a pure strategy space $X$.

**Assumption 1** Both $T$ and $X$ are compact metric spaces.

We let $d_T$ and $d_X$ denote the metric on $T$ and $X$, respectively. Since the focus is on a property that depends on the number of players, we will index any game by the number of its players. Thus, $G_n$ is a Bayesian game in which the set of players is $I_n = \{1, \ldots, n\}$, and each has $T$ as his type space and $X$ as his choice set.

Players are allowed to choose randomized strategies. In this context, a randomized strategy for a player is a function from types into a Borel probability measure on the set of his pure strategies. Let $\mathcal{M}(X)$ be the set of Borel probability measures on $X$ endowed with the weak convergence topology. This topology is induced by the Prohorov metric on $\mathcal{M}(X)$. Thus, a randomized strategy for player $i \in I_n$ is a Borel measurable function $\sigma_i : T \rightarrow \mathcal{M}(X)$. A strategy is pure if its values are degenerate probability measures on $X$. Thus, if for all $x \in X$, $1_x$ denotes the probability measure on $X$ that attributes probability 1 to $x$, then a strategy $f_i : T \rightarrow \mathcal{M}(X)$ is pure if for all $t \in T$, there exists $x \in X$ such that $f_i(t) = 1_x$. Let $\Sigma$ denote the set of all (randomized) strategies.

For all $i \in I_n$, player $i$’s type is decided by nature according to $\tau_i \in \mathcal{M}(T)$. This probability measure, together with a strategy $\sigma_i : T \rightarrow \mathcal{M}(X)$, defines a measure $\tau_i \otimes \sigma_i$ on $T \times X$ as the unique measure satisfying
\[
\tau_i \otimes \sigma_i(A \times B) = \int_A \sigma_i(B|t)d\tau_i(t) \tag{1}
\]
for all Borel measurable subsets $A$ of $T$ and $B$ of $X$ (see Ash (1972, Theorem 2.6.2, p. 97)). We let $\rho$ denote the Prohorov metric in $\mathcal{M}(T \times X)$.

In both classes, players’ payoff functions are defined using a function on types, actions and distributions. To each player $i$, we associate a function $V^n_i : T \times X \times \mathcal{M}(T \times X) \rightarrow \mathbb{R}$ with the following interpretation: $V^n_i(t, x, \mu)$ is player $i$’s payoff when he is of type $t$, plays action $x$ and faces the distribution $\mu$.5
Assumption 2 For all $i \in I_n$, $V_i^n$ is continuous.

In both classes of games that we consider, a game is defined using the same elements: $I_n$, $V_n$, $\tau_n$, $T$ and $X$. Thus, we represent a game by $G_n = (I_n, V_n, \tau_n, T, X)$.

We let $\mathcal{U}$ denote the space of all continuous real-valued functions on $T \times X \times \mathcal{M}(T \times X)$ endowed with the sup norm. For each game $G_n$, let $V_n(I_n) = \{V_i^n : i \in I_n\} \subseteq \mathcal{U}$. A subset $K$ of $\mathcal{U}$ is equicontinuous if for all $\eta > 0$ there exists a $\delta > 0$ such that $|v(t, x, \mu) - v(w, y, \nu)| < \eta$ whenever $\max\{d_T(t, w), d_X(x, y), \rho(\mu, \tau)\} < \delta$, $t, w \in T$, $x, y \in X$, $\mu, \tau \in \mathcal{M}(T \times X)$ and $v \in K$. A subset $K$ of $\mathcal{U}$ is bounded if there exists $\xi > 0$ such that $|v(t, x, \mu)| \leq \xi$ for all $t \in T$, $x \in X$, $\mu \in \mathcal{M}(T \times X)$ and $v \in K$.

The two frameworks we consider differ on how players’ payoff functions are defined from $V_i^n$, as will be described in the following subsections.

2.1 Average Distribution

In this case, each player’s payoff depends on his type-action character and on the average distribution of type-action characters of the others. Given a strategy $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma^n$ in a game $G_n$ with $n$ players, the average distribution is:

$$\bar{\sigma} = \sum_{i \in I_n} \frac{\tau_i \otimes \sigma_i}{n}. \quad (2)$$

Similarly, for all $i \in I_n$, the average distribution of the other players is

$$\bar{\sigma}_{-i} = \sum_{l \neq i} \frac{\tau_l \otimes \sigma_l}{n - 1}. \quad (3)$$

Note that both $\bar{\sigma}$ and $\bar{\sigma}_{-i}$ are elements of $\mathcal{M}(T \times X)$. Then, for all strategies $\sigma$, player $i$’s payoff function is

$$U_i^n(\sigma) = \int_T \int_X V_i^n(t, x, \bar{\sigma}_{-i})d\sigma_i(x|t)d\tau_i(t). \quad (4)$$

We denote this class of games by $\mathcal{G}_R$, which is formally defined in the following definition.

Definition 1 A Bayesian game $G_n = (I_n, V_n, \tau_n, T, X)$ belongs to the class $\mathcal{G}_R$ if $U_i^n$ is defined by (4) for all $i \in I_n$ and both Assumptions 1 and 2 are satisfied.
2.2 Empirical Distribution

In this case, players’ payoff functions are defined (from \( V^n_i \)) in the following way. Let \( C = T \times X, i \in I_n \text{ and } c_{-i} \in C_{-i} = C^{n-1} \). Define a measure \( \text{emp}_{c_{-i}} \in \mathcal{M}(T \times X) \) by

\[
\text{emp}_{c_{-i}}(B) = \frac{|\{j \neq i : c_j \in B\}|}{n-1} \tag{5}
\]

for all Borel measurable subsets \( B \) of \( T \times X \). The measure \( \text{emp}_{c_{-i}} \) is the empirical distribution and the quantity \( \text{emp}_{c_{-i}}(B) \) is simply the fraction of players other than player \( i \) with type-action characters in the set \( B \).

The following lemma shows that if two vectors \( c_{-i} \) and \( \hat{c}_{-i} \) are close, then so will be \( \text{emp}_{c_{-i}} \) and \( \text{emp}_{\hat{c}_{-i}} \).

**Lemma 1** Let \( i \in I_n \) and \( \eta > 0 \). If \( \max_{j \neq i} d(c_j, \hat{c}_j) \leq \eta \), then \( \rho(\text{emp}_{c_{-i}}, \text{emp}_{\hat{c}_{-i}}) \leq \eta \).

**Proof.** Let \( D \) be a Borel measurable subset of \( T \times X \). Since \( c_j \in D \) implies that \( \hat{c}_j \in \overline{B}_\eta(D) \), then

\[
\text{emp}_{c_{-i}}(D) = \frac{|\{j \neq i : c_j \in D\}|}{n-1} \leq \frac{|\{j \neq i : \hat{c}_j \in \overline{B}_\eta(D)\}|}{n-1} = \text{emp}_{\hat{c}_{-i}}(\overline{B}_\eta(D)). \tag{6}
\]

Thus, \( \rho(\text{emp}_{c_{-i}}, \text{emp}_{\hat{c}_{-i}}) \leq \eta \). \( \blacksquare \)

In particular, Lemma 1 implies that the function \( \text{emp} : c_{-i} \mapsto \text{emp}_{c_{-i}} \) is continuous, which, in turn, implies that the function \( c \mapsto V^n_i(c_i, \text{emp}_{c_{-i}}) \) is continuous.

For all strategies \( \sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma^n \) and \( i \in I_n \), let \( \tau \otimes \sigma = \prod_{l \in I_n} \tau_l \otimes \sigma_l \) and \( [\tau \otimes \sigma]_{-i} = \prod_{l \neq i} \tau_l \otimes \sigma_l \), and define

\[
U^n_i(\sigma) = \int_{\mathcal{C}^n} V^n_i(c_i, \text{emp}_{c_{-i}})d\tau \otimes \sigma(c) = \int_{\mathcal{C}} \int_{\mathcal{C}^{n-1}} V^n_i(c_i, \text{emp}_{c_{-i}})d[\tau \otimes \sigma]_{-i}(c_{-i})d\tau_i \otimes \sigma_i(c_i). \tag{7}
\]

We denote this class of games by \( \mathcal{G}_K \).

**Definition 2** A Bayesian game \( G_n = (I_n, V_n, \tau_n, T, X) \) belongs to the class \( \mathcal{G}_K \) if \( U^n_i \) is defined by (7) for all \( i \in I_n \) and Assumptions 1 and 2 are satisfied.
2.3 Equilibria and Purification

Let $G = G_R \cup G_K$. The following notions and notation apply to all $G_n$ in $G$.

For convenience, let, for all $i \in I_n$, $t \in T$, $x \in X$ and $\sigma \in \Sigma^n$, $U^n_i(x, \sigma_{-i}|t) = \int_{C_{-i}} V^n_i(t, x, \text{emp}_{c_{-i}}) d[\tau \otimes \sigma]_{-i}(c_{-i})$ if $G_n \in G_K$ and $U^n_i(x, \sigma_{-i}|t) = V^n_i(t, x, \bar{\sigma}_{-i})$ if $G_n \in G_R$ be player $i$’s payoff when his type is $t$, he plays $x$ and the other players play according to $\sigma_{-i}$. Furthermore, let $U^n_i(\sigma|t) = \int_X U^n_i(x, \sigma_{-i}|t) d\sigma_i(x|t)$ for all $i \in I_n$, $t \in T$ and $\sigma \in \Sigma^n$ be player $i$’s payoff when his type is $t$ and $\sigma$ is played. Using this notation, we can write $U^n_i(\sigma) = \int_T U^n_i(\sigma|t) d\tau_i(t)$.

For all $\varepsilon \geq 0$, we say that a strategy $\sigma$ is a strong $\varepsilon$ – equilibrium of a game $G_n$ if $U^n_i(x, \sigma_{-i}|t) \geq U^n_i(\hat{x}, \sigma_{-i}|t) - \varepsilon$ for all $i \in I_n$, $t \in T$, $x \in \text{supp}(\sigma_i(t))$, and $\hat{x} \in X$. Thus, in a strong $\varepsilon$ – equilibrium all players assign a strictly positive probability only to the actions that are within $\varepsilon$ of their best response. A strategy $\sigma$ is a Bayesian Nash equilibrium of $G_n$ if $\sigma$ is a strong $\varepsilon$ – equilibrium of $G_n$ for $\varepsilon = 0$.

We remark that this definition is equivalent to the more common one, according to which a strategy $\sigma$ is a Bayesian Nash equilibrium if for all $i \in I_n$ and $t \in T$, $\sigma_i(t)$ maximizes player $i$’s payoff given that the others are using $\sigma_{-i}$.

Let $\sigma$ be a strong $\zeta$ – equilibrium of a game $G_n$. Then, we say that $g$ is an $\epsilon$ – purification of $\sigma$ if $g$ is a pure strong $\zeta + \epsilon$ – equilibrium and

$$|U^n_i(g|t) - U^n_i(\sigma|t)| < \zeta + \epsilon,$$

for all $i \in I_n$ and $t \in T$. Thus, under $g$, all players are playing actions that are $\zeta + \epsilon$ – best responses to $g_{-i}$. Furthermore, their payoff under $g$ is close to their payoff under $\sigma$ for all possible types. In particular, when $\sigma$ is a Bayesian-Nash equilibrium of $G_n$ and $g$ is an $\epsilon$ – purification of $\sigma$, then $g$ is a strong $\epsilon$ – equilibrium and each player’s payoff under $g$ is, for each type, within $\epsilon$ of that under $\sigma$.

3 Purification of Equilibria

Our main result says that in all sufficiently large games, all Bayesian Nash equilibria can be $\epsilon$ – purified, provided that players’ payoff functions are selected from a bounded

equicontinuous family (and that we can dispense with the boundedness of the family of payoff functions for games in the class $G_R$).

**Theorem 1** Let $K$ be an equicontinuous subset of $U$. Then, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ with the following property:

If $n \geq N$ and $G_n = (I_n, V_n, \tau_n, T, X) \in G_R$ is such that $V_n(I_n)$ is a subset of $K$, then all Bayesian Nash equilibria of $G_n$ can be $\varepsilon$–purified.

If, in addition, $K$ is bounded, then the same conclusion holds for the class $G_K$.

### 3.1 Proof of Theorem 1

The proof of our purification result follows two steps in both of the frameworks. First, we establish a purification result for finitely supported approximate equilibria in games with finitely supported distributions over types. Second, we show that such games and equilibria can be used to approximate general equilibria in general games.

**Lemma 2** Let $K$ be an equicontinuous subset of $U$ and $m, p \in \mathbb{N}$. Then, for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that:

If $n \geq N$ and $G_n = (I_n, V_n, \tau_n, T, X) \in G_R$ is such that $V_n(I_n) \subseteq K$ and there exists a finite subset $\hat{T}$ of $T$ with $|\hat{T}| = p$ and $\text{supp}(\tau_n^i) \subseteq \hat{T}$ for all $i \in I_n$, then all strong $\zeta$–equilibrium $\nu$ of $G_n$, $\zeta \geq 0$, can be $\varepsilon$–purified provided that there exists a finite subset $\hat{X}$ of $X$ with $|\hat{X}| = m$ and $\text{supp}(\nu_i(t)) \subseteq \hat{X}$ for all $i \in I_n$ and $t \in T$.

Furthermore, we can choose the $\varepsilon$–purification $g$ so that $g_i(t) = g_i(t')$ for all $i \in I_n$, $t \in T \setminus \hat{T}$ and $t' \in \hat{T}$, such that $\nu_i(t) = \nu_i(t')$.

If, in addition, $K$ is bounded, then the same conclusion holds for the class $G_K$.

Although Lemma 2 implies the purification theorem of Rashid (1983), which corresponds to the case of $\zeta = 0$, $X$ finite and $T$ singleton, its proof follows closely Rashid’s arguments.\(^2\) As there, the critical result for the construction of the purification strategies is the Shapley-Folkman Theorem.

---

\(^2\)Although the equicontinuity assumption was not made explicit in Rashid (1983), it is needed as we have shown in Carmona (2004).
Proof. Let \( \varepsilon > 0 \). Let \( 0 < \eta < \varepsilon/2B \) with \( B = 1 \) if \( G_n \in G_R \) and \( B = 4\xi + 3 \) if \( G_n \in G_K \), where \( \xi > 0 \) is such that \( |v(t, x, \mu)| \leq \xi \) for all \( t \in T, x \in X \) and \( \mu \in M(T \times X) \) (such \( \xi \) exists whenever \( K \) is bounded).

Since \( K \) is equicontinuous, let \( \delta > 0 \) be such that \( |v(t, x, \tau) - v(w, y, \mu)| < \eta \) whenever \( \max \{d_T(t, w), d_X(x, y), \rho(\tau, \mu)\} < \delta \), \( x, y \in X \), \( \tau, \mu \in M(T \times X) \) and \( v \in K \).

Finally, let \( N \in \mathbb{N} \) be such that \( n > m, 4m^2p/(n-1) < \delta \) and \( 2mpc^{-2n(\delta/2mp)} < \eta \) whenever \( n \geq N \). In particular, \( m^2p/n < \delta/2 \) and \( 2/(n-1) < \delta/2 \) for all \( n \geq N \).

Let \( n \geq N \) and \( G_n \) be a game in \( G \) and \( \nu \) be a strong \( \zeta \) - equilibrium satisfying the assumptions in the statement of the lemma.

Let \( t \in \hat{T} \) and define for all \( i \in I_n \)

\[
S_{i,t} = \left\{ \frac{\tau_i(t)}{n} e_j : \nu_i(x_j|t) > 0 \right\},
\]

(9)

where \( E = \{e_1, \ldots, e_m\} \) is the standard basis of \( \mathbb{R}^m \). In particular, note that \( S_{i,t} = \{0\} \) if \( \tau_i(t) = 0 \). However, if \( \tau_i(t) > 0 \), then \( (n/\tau_i(t))S_{i,t} \subseteq E \).

Note also that \( (\tau_i(t)\nu_i(x_1|t)/n, \ldots, \tau_i(t)\nu_i(x_m|t)/n) \in \text{co}(S_{i,t}) \) for all \( i \in I_n \) and \( t \in \hat{T} \). This implies that

\[
(\tilde{\nu}(t, x_1), \ldots, \tilde{\nu}(t, x_m)) = \sum_{i=1}^{n} \frac{\tau_i(t)}{n} (\nu_i(x_1|t), \ldots, \nu_i(x_m|t))
\]

is an element of \( \text{co} \left( \sum_{i=1}^{n} S_{i,t} \right) = \sum_{i=1}^{n} \text{co}(S_{i,t}) \). By the Shapley-Folkman Theorem (see Rashid (1983)), it follows that there are \( n \) points \( \alpha_i(t) = (\alpha_i(x_1|t), \ldots, \alpha_i(x_m|t)) \in \text{co}(S_{i,t}), \) with \( i \in I_n \), such that \( |\{i \in I_n : \alpha_i(t) \notin S_{i,t}\}| \leq m \) and

\[
(\tilde{\nu}(t, x_1), \ldots, \tilde{\nu}(t, x_m)) = \sum_{i=1}^{n} (\alpha_i(x_1|t), \ldots, \alpha_i(x_m|t)).
\]

Since \( S_{i,t} = \text{co}(S_{i,t}) = \{0\} \) if \( \tau_i(t) = 0 \), then \( \{i \in I_n : \alpha_i(t) \notin S_{i,t}\} \subseteq \{i \in I_n : \tau_i(t) > 0\} \). Hence, \( |\{i \in I_n : \tau_i(t) > 0 \text{ and } n\alpha_i(t)/\tau_i(t) \notin E\}| \leq m \).

For all \( t \in \hat{T} \), let \( P_t = \{i \in I_n : \tau_i(t) > 0 \text{ and } n\alpha_i(t)/\tau_i(t) \notin E\} \). Define a pure strategy \( g \) as follows: if \( i \in P_t \), let \( e_j \) be such that \( n\alpha_i(t)/\tau_i(t) = e_j \) and define \( g_i(t) = 1_{x_j} \); if \( i \in P_t^c \), choose \( 1 \leq j \leq m \) such that \( \nu_i(x_j|t) > 0 \) and define \( g_i(t) = 1_{x_j} \). Finally, for all \( t \in T \setminus \hat{T} \), let \( x \in \text{supp}(\nu_i(t)) \) and set \( g_i(t) = 1_x \). Since the only requirement
that we impose on \( g_i(t) \) when \( t \in T \setminus \hat{T} \) is that \( g_i(t) \in \text{supp}(\nu_i(t)) \), it is clear that we can let \( g_i(t) = g_i(t') \) whenever \( t' \in \hat{T} \) and \( \nu_i(t) = \nu_i(t') \). It follows from the definition of \( g_i \) that \( g_i(t) \in \text{supp}(\nu_i(t)) \) for all \( t \in T \) and so \( U^n_i(g_i(t), \nu_{-i}|t) \geq U^n_i(x, \nu_{-i}|t) - \zeta \) for all \( i \in I_n, t \in T \) and \( x \in X \).

Note that, for all \( j \in \{1, \ldots, m\} \) and \( t \in \hat{T} \), \( \bar{g}(t, x_j) = \sum_{i \in P_t} \alpha_i(x_j|t) + \frac{1}{n} \sum_{i \in P_t^i} \tau_i(t)g_i(x_j|t) \).

Hence,

\[
|\bar{\bar{\nu}}(t, x_j) - \bar{\bar{\nu}}(t, x_j)| = \frac{1}{n} \left| \sum_{i \in P_t^i | \tau_i(t) > 0} \tau_i(t) \left( \frac{n\alpha_i(x_j|t)}{\tau_i(t)} - g_i(x_j|t) \right) \right| \leq \frac{m}{n}. \tag{10}
\]

The following claim states that conditional payoffs under \( \nu \) are close to those under \( g \).

**Claim 1** For all \( i \in I_n, x \in X \) and \( t \in T \), \( |U^n_i(x, \nu_{-i}|t) - U^n_i(x, g_{-i}|t)| < \varepsilon/2 \).

**Proof of Claim 1.** We consider two cases: when \( G_n \in G_R \) and when \( G_n \in G_K \). In both cases, the following result is used.

**Claim 2** If \( \Psi \) is a finite subset of a metric space \( Y \) and both \( \mu \) and \( \pi \) are probability measures supported in \( \Psi \) such that \( |\pi(\{l\}) - \mu(\{l\})| \leq q \) for all \( l \in \Psi \) and some \( q > 0 \), then \( \rho(\pi, \mu) \leq \|\Psi\|q \).

**Proof.** Let \( q > 0 \) and \( B \subseteq Y \) be Borel measurable. Then,

\[
\mu(B) = \sum_{l \in \Psi \cap B} \mu(\{l\}) \leq \sum_{l \in \Psi \cap B} (\pi(\{l\}) + q) \leq \sum_{l \in \Psi \cap B} \pi(\{l\}) + |\Psi|q \leq \pi(B|\Psi|q) + |\Psi|q.
\]

Symmetrically, we have that \( \pi(B) \leq \mu(B|\Psi|q) + |\Psi|q \). This implies that \( \rho(\mu, \pi) \leq |\Psi|q \). ■

**Case 1** \( G_n \in G_R \).

It follows from Claim 2 and (10) that \( \rho(\bar{\bar{\nu}}, \bar{\bar{g}}) \leq m^2p/n < \delta/2 \). As can be easily verified, we also have that \( \rho(\bar{\bar{\sigma}}, \bar{\bar{\sigma}}_{-i}) \leq \frac{1}{n-1} \) for all \( i \in I_n \) and strategies \( \sigma \). Thus, it follows that \( \rho(\bar{\bar{\nu}}_{-i}, \bar{\bar{g}}_{-i}) \leq \rho(\bar{\bar{\nu}}_{-i}, \bar{\bar{v}}) + \rho(\bar{\bar{v}}, \bar{\bar{g}}) + \rho(\bar{\bar{g}}, \bar{\bar{g}}_{-i}) \leq 1/(n-1) + \delta/2 + 1/(n-1) < \delta \) for all \( i \in I_n \). Therefore, \( |U^n_i(x, \nu|t) - U^n_i(x, g|t)| = |V^n_i(t, x, \nu_{-i}) - V^n_i(t, x, g_{-i})| < \eta < \varepsilon/2 \).
Case 2 $G_n \in \mathcal{G}_K$.

Note that, for all strategies $\sigma$, $i \in I_n$ and $t \in T$,

$$U^n_i(x, \sigma_{-i}|t) = \int_{T_i \times X_i} V^n_i(t, x, \text{emp}_{(t-i, x-i)})d[\tau \otimes \nu]_{-i}(t-i, x-i)$$

$$= \int_{T_i \times X_i} V^n_i(t, x, \text{emp}_{(t-i, x-i)})d[\tau \otimes \nu]_{-i}(t-i, x-i).$$

(11)

Thus, it is enough to consider $(t_{-i}, x_{-i}) \in \hat{T}_{-i} \times \hat{X}_{-i}$.

Let $i \in I_n$ and define

$$\text{nbd}(\bar{\nu}_{-i}, \delta') = \{(t_{-i}, x_{-i}) \in \hat{T}_{-i} \times \hat{X}_{-i} : \max_{(t,x) \in T \times X} |\text{emp}_{(t-i, x-i)}(t, x) - \bar{\nu}_{-i}(t, x)| \leq \delta'\}. \quad (12)$$

Note that if $(t_{-i}, x_{-i})$ is such that $\text{emp}_{(t_{-i}, x_{-i})}$ belongs to $\text{nbd}(\bar{\nu}_{-i}, \delta/2mp)$, then $\rho(\text{emp}_{(t_{-i}, x_{-i})}, \bar{\nu}_{-i}) \leq |\hat{T} \times \hat{X}|\delta/2mp = \delta/2 < \delta$ (by Claim 2). Thus, $|V^n_i(t, x, \text{emp}_{(t_{-i}, x_{-i})}) - V^n_i(t, x, \bar{\nu}_{-i})|$ is smaller than $\eta$ if $\text{emp}_{(t_{-i}, x_{-i})} \in \text{nbd}(\bar{\nu}_{-i}, \delta/2mp)$ and smaller than $2\xi$ if $\text{emp}_{(t_{-i}, x_{-i})} \notin \text{nbd}(\bar{\nu}_{-i}, \delta/2mp)$.

Let $\gamma = [\tau \otimes \nu]_{-i}(\{(t_{-i}, x_{-i}) \in \hat{T}_{-i} \times \hat{X}_{-i} : \text{emp}_{(t_{-i}, x_{-i})} \notin \text{nbd}(\bar{\nu}_{-i}, \delta/2mp)\})$. By Kalai (2004, Lemma 4), then $\gamma < 2mpe^{-2n(\delta/2mp)^2}$ and so

$$\left|\int_{\hat{T}_{-i} \times \hat{X}_{-i}} V^n_i(t, x, \text{emp}_{(t_{-i}, x_{-i})})d[\tau \otimes \nu]_{-i}(t_{-i}, x_{-i}) - V^n_i(t, x, \bar{\nu}_{-i})\right| < \gamma 2\xi + (1-\gamma)\eta < 4mp\xi e^{-2n(\delta/2mp)^2} + \eta < (2\xi + 1)\eta. \quad (13)$$

Similarly, we have that

$$\left|\int_{\hat{T}_{-i} \times \hat{X}_{-i}} V^n_i(t, x, \text{emp}_{(t_{-i}, x_{-i})})d[\tau \otimes g]_{-i}(t_{-i}, x_{-i}) - V^n_i(t, x, \bar{g}_{-i})\right| < (2\xi + 1)\eta. \quad (14)$$

As shown in Case 1 above, we have that $\rho(\bar{\nu}_{-i}, \bar{g}_{-i}) < \delta$ and $|V^n_i(t, x, \bar{\nu}_{-i}) - V^n_i(t, x, \bar{g}_{-i})| < \eta$. Hence, $|U^n_i(x, \nu_{-i}|t) - U^n_i(x, g_{-i}|t)| < (4\xi + 3)\eta < \varepsilon/2$. ■

It follows from Claim 1 that, for all $i \in I_n$, $t \in T$ and $x \in X$,

$$U^n_i(g_i(t), g_{-i}|t) > U^n_i(g_i(t), \nu_{-i}|t) - \frac{\varepsilon}{2} \geq U^n_i(t, \nu_{-i}|t) - \frac{\varepsilon}{2} - \zeta > U^n_i(t, g_{-i}|t) - \varepsilon - \zeta. \quad (15)$$

Therefore, $g$ is a pure $\zeta + \varepsilon$-equilibrium of $G_n$. 

12
Since $\nu$ is a strong $\zeta$-equilibrium, it follows that $|U^n_i(x, \nu_{-i}|t) - U^n_i(\hat{x}, \nu_{-i}|t)| \leq \zeta$ for all $i \in I_n$, $t \in T$ and $x, \hat{x} \in \text{supp}(\nu_i(t))$. Hence,

\[
\left| \sum_{j=1}^{m} \nu_i(x_j|t)U^n_i(x_j, \nu_{-i}|t) - U^n_i(g_i(t), g_{-i}|t) \right| \leq \varepsilon.
\]

(16)

Thus, $|U^n_i(\nu|t) - U^n_i(g|t)| < \zeta + \varepsilon$ for all $i \in I_n$ and $t \in T$ and so $g$ is an $\varepsilon$-purification of $\nu$. 

We turn to the proof of Theorem 1. It consists of approximating any Bayesian Nash equilibrium in any game with a compact type and action space with a finitely supported approximate equilibrium in a game with finitely supported distributions over types, and then applying Lemma 2 to the latter.

**Proof of Theorem 1.** Let $\varepsilon > 0$ and let $0 < \gamma < \varepsilon/20$. Since $K$ is equicontinuous, there exists $\delta > 0$ such that $\max\{d_T(t, w), d_X(x, y), \rho(\mu, \nu)\} < \delta$ implies that $|v(t, x, \mu) - v(w, y, \nu)| < \gamma$ for all $t, w \in T$, $x, y \in X$, $\mu, \nu \in \mathcal{M}(T \times X)$ and $v \in K$. Since $X$ is a compact metric space, let $\hat{X} = \{x_1, \ldots, x_m\}$ be a finite subset of $X$ and $\{X_j\}_{j=1}^{m}$ be a finite measurable partition of $X$ such that $x_j \in X_j \subseteq B_{\delta/4}(x_j)$ for all $j = 1, \ldots, m$. Similarly, let $\hat{T} = \{t_1, \ldots, t_p\}$ be a finite subset of $T$ and $\{T_l\}_{l=1}^{p}$ be a finite measurable partition of $T$ such that $t_l \in T_l \subseteq B_{\delta/4}(t_l)$ for all $l = 1, \ldots, p$. Let $N \in \mathbb{N}$ be given by Lemma 2 and corresponding to $K, \gamma, m$ and $p$.

Let $G_n = (I_n, V_n, \tau_n, T, X)$ be such that $V_n(I_n)$ is a subset of $K$. Let $\sigma$ be a Bayesian Nash equilibrium of $G_n$. Define a strategy $\varphi$ of $G_n$ by

\[
\varphi_i(x_j|t_l) = \begin{cases} \int_{T_l} \sigma_i(X_j|t)d\tau_i(t)/\tau_i(T_l) & \text{if } \tau_i(T_l) > 0, \\
\sigma_i(X_j|t_l) & \text{if } \tau_i(T_l) = 0 \end{cases}
\]

for all $i \in I_n$, $1 \leq l \leq p$ and $1 \leq j \leq m$ and $\varphi_i(t) = \varphi_i(t_l)$ for all $t \in T_l$.

We next claim that the payoff under $\sigma$ is close to that under $\varphi$ for all players and types.

**Claim 3** For all $i \in I_n$, $t, t' \in T$ and $x, x' \in X$ such that $\max\{d_T(t, t'), d_X(x, x')\} < \delta$, we have that $|U^n_i(x', \sigma_{-i}|t') - U^n_i(x, \varphi_{-i}|t)| < 2\gamma$. 

13
Proof of Claim 3.

We consider two cases: when \( G_n \in \mathcal{G}_R \) and when \( G_n \in \mathcal{G}_K \).

Case 1 \( G_n \in \mathcal{G}_R \).

We start this case with the following claim.

Claim 4 For all \( i \in I_n \), \( \rho(\tau_i \otimes \sigma_i, \tau_i \otimes \varphi_i) \leq \delta/2 \).

Proof of Claim 4. Note that \( \tau_i \otimes \sigma_i(T_l \times X_j) = \tau_i \otimes \varphi_i(T_l \times X_j) \) for all \( l \) and \( j \). For convenience, let \( \mu = \tau_i \otimes \sigma_i \), \( \nu = \tau_i \otimes \varphi_i \) and \( \beta = \delta/4 \). Let \( D \subseteq T \times X \) be Borel measurable. Let \( A = \{(l, j) : \mu((T_l \times X_j) \cap D) > 0\} \) and \( B = \{(l, j) : \nu((T_l \times X_j) \cap \overline{B}_{2\beta}(D)) > 0\} \). Note that if \( (l, j) \in A \), then \((T_l \times X_j) \cap D \neq \emptyset\), which implies that \((t_l, x_j) \in \overline{B}_{2\beta}(D)\). Hence, \((t, x) \in \overline{B}_{2\beta}(D)\) for all \( t \in T_l \) and \( x \in X_j \), i.e., \( T_l \times X_j \subseteq \overline{B}_{2\beta}(D) \). It follows that \( \nu((T_l \times X_j) \cap \overline{B}_{2\beta}(D)) = \nu(T_l \times X_j) = \mu(T_l \times X_j) > 0 \).

Thus, \( A \subseteq B \).

Therefore,

\[
\mu(D) = \sum_{(l,j) \in A} \mu((T_l \times X_j) \cap D) \leq \sum_{(l,j) \in A} \mu(T_l \times X_j) = \sum_{(l,j) \in A} \nu(T_l \times X_j) = \sum_{(l,j) \in A} \nu((T_l \times X_j) \cap \overline{B}_{2\beta}(D)) \leq \sum_{(l,j) \in B} \nu((T_l \times X_j) \cap \overline{B}_{2\beta}(D)) = \nu(\overline{B}_{2\beta}(D)) < \nu(\overline{B}_{2\beta}(D)) + 2\beta.
\]

By Billingsley (1999, p. 72), we also have that \( \nu(D) \leq \mu(\overline{B}_{2\beta}(D)) + 2\beta \) and so \( \rho(\mu, \nu) \leq 2\beta = \delta/2 \). \( \blacksquare \)

Since \( \rho(\tau_i \otimes \sigma_i, \tau_i \otimes \varphi_i) \leq \delta/2 \) for all \( i \in I_n \), it follows that \( \rho(\tilde{\tau} \otimes \tilde{\sigma}, \tilde{\tau} \otimes \tilde{\varphi}) \leq \delta/2 \).

Thus, \( |U^n_i(x', \sigma_\sim |l') - U^n_i(x, \varphi_\sim |l)| = |V^n_i(t', x', \tilde{\sigma}_\sim) - V^n_i(t, x, \tilde{\varphi}_\sim)| < \gamma \).

Case 2 \( G_n \in \mathcal{G}_K \).

The partitions \( \{X_1, \ldots, X_m\} \) and \( \{T_1, \ldots, T_p\} \) induce a partition \( \{C_{1}, \ldots, C_{mp}\} \) of \( T \times X \). Since \( \tau_i \otimes \varphi_i(T_l \times X_j) = \tau_i \otimes \sigma_i(T_l \times X_j) \) for all \( i \in I_n \), \( l = 1, \ldots, p \) and \( j = 1, \ldots, m \), then, \( \tau \otimes \varphi|_{\tau_i}(\prod_{j \neq i} C_{r_j}) = \tau \otimes \sigma|_{\tau_i}(\prod_{j \neq i} C_{r_j}) \) for all \( i \in I_n \), \( j \neq i \) and \( r_j \in \{1, \ldots, mp\} \). Let \( \Xi_\sim = \{\prod_{j \neq i} C_{r_j} : r_j \in \{1, \ldots, mp\} \text{ for all } j \neq i\} \).
Clearly, it is enough to show that
\[ \left| \int_A V_i^n(t, x, \text{emp}_{c_{-i}})d[\tau \otimes \varphi]_{-i}(c_{-i}) - \int_A V_i^n(t', x', \text{emp}_{c_{-i}})d[\tau \otimes \sigma]_{-i}(c_{-i}) \right| < 2\gamma[\tau \otimes \sigma]_{-i}(A) \] (18)
for all \( A \in \Xi_{-i} \).

Let \( A \in \Xi_{-i} \) and \( \hat{c}_{-i} \in A \). Hence, if \( c_{-i} \in A \), then \( \max\{d_T(t_j, \hat{t}_j), d_X(x_j, \hat{x}_j)\} < \delta/2 \) for all \( j \neq i \) and so \( \rho(\text{emp}_{c_{-i}}, \text{emp}_\hat{c}_{-i}) \leq \delta/2 \) by Lemma 1. Since \( [\tau \otimes \varphi]_{-i}(A) = [\tau \otimes \sigma]_{-i}(A) \), it follows that
\[
\left| \int_A V_i^n(t, x, \text{emp}_{c_{-i}})d[\tau \otimes \varphi]_{-i}(c_{-i}) - \int_A V_i^n(t', x', \text{emp}_{c_{-i}})d[\tau \otimes \sigma]_{-i}(c_{-i}) \right|
\leq \left| \int_A V_i^n(t, x, \text{emp}_{c_{-i}})d[\tau \otimes \varphi]_{-i}(c_{-i}) - [\tau \otimes \varphi]_{-i}(A)V_i^n(t', x', \text{emp}_{c_{-i}}) \right|
+ \left| [\tau \otimes \sigma]_{-i}(A)V_i^n(t', x', \text{emp}_{c_{-i}}) - \int_A V_i^n(t', x', \text{emp}_{c_{-i}})d[\tau \otimes \sigma]_{-i}(c_{-i}) \right|
< 2\gamma[\tau \otimes \sigma]_{-i}(A). \] (19)

Thus, we obtain the desired conclusion. ■

We next show that \( \varphi \) is a strong \( 4\gamma \) - equilibrium of \( G_n \). Fix \( i \in I_n \) and \( t \in T \). Let \( t_i \in \hat{T} \) be such that \( \varphi_i(t) = \varphi_i(t_i) \) and \( x_j \in \hat{X} \) be such that \( x_j \in \text{supp}(\varphi_i(t)) \).
The definition of \( \varphi_i(t) \) implies that there exists \( t' \in T_i \) and \( x' \in X_j \) such that \( x' \in \text{supp}(\sigma_i(t')) \). Since \( \sigma \) is a Bayesian Nash equilibrium, it follows by Claim 3 that, for all \( x \in X \),
\[ U^n_i(x, \varphi_{-i}|t) > U^n_i(x', \sigma_{-i}|t') - 2\gamma > U^n_i(x, \sigma_{-i}|t') - 2\gamma = U^n_i(x, \varphi_{-i}|t) - 4\gamma. \] (20)
Thus, \( \varphi \) is a strong \( 4\gamma \) - equilibrium of \( G_n \).

Furthermore, \( |U^n_i(\varphi|t) - U^n_i(\sigma|t)| < 7\gamma \) for all \( i \in I_n \) and \( t \in T \). This claim can be established as follows. Let \( l \in \{1, \ldots, p\} \) be such that \( t \in T_l \) and \( j \in \{1, \ldots, m\} \) be such that \( x_j \in \text{supp}(\varphi_l(t)) \). Then, \( |U^n_i(\varphi|t) - U^n_i(x_j, \varphi_{-i}|t)| < 4\gamma \). Also, let \( t' \in T_l \) and \( x' \in X_j \) be such that \( x' \in \text{supp}(\sigma_l(t')) \). Since \( \sigma \) is a Bayesian Nash equilibrium, then \( U^n_i(\sigma|t) = \max_x U^n_i(x, \sigma_{-i}|t) \) and \( U^n_i(\sigma|t') = \max_x U^n_i(x, \sigma_{-i}|t') = U^n_i(x', \sigma_{-i}|t') \). Since \( d_T(t, t') < \delta \), then \( |U^n_i(\sigma|t) - U^n_i(x', \sigma_{-i}|t')| < \gamma \).\(^3\) Therefore, it follows from Claim 4 and from \( \max\{d_T(t, t'), d_X(x_j, x')\} < \delta \) that \( |U^n_i(\varphi|t) - U^n_i(\sigma|t)| < 7\gamma. \)

\(^3\)This conclusion follows from a standard “maximum theorem” argument (see Berge (1997, p. 116)). In fact, \( \max_x U^n_i(x, \sigma_{-i}|t') = U^n_i(x', \sigma_{-i}|t') < U^n_i(x_j, \varphi_{-i}|t) + \gamma \leq \max_x U^n_i(x, \sigma_{-i}|t) + \gamma \)
Consider the game \( \hat{G}_n = (I_n, V_n, \hat{\tau}_n, T, X) \) with \( \hat{\tau}_n(t_l) = \tau_n(t_l) \) for all \( l \in \{1, \ldots, p\} \) and \( i \in I_n \). Let \( \hat{U}_i^n \) denote player \( i \)'s payoff function in \( \hat{G}_n \), which is defined as \( U_i^n \) but with \( \hat{\tau}_n^k \) in place of \( \tau_n^k \) for all \( k \in I_n \). Since \( \hat{\tau}_i \otimes \varphi_i(T_i \times X_j) = \hat{\tau}_i \otimes \varphi_i(T_i \times X_j) \) for all \( i \in I_n, l \in \{1, \ldots, p\} \) and \( j \in \{1, \ldots, m\} \), then, using an argument similar to that in Claim 3, it follows that

\[
|\hat{U}_i^n(x, \varphi_{-i}|t) - U_i^n(x', \varphi_{-i}|t')| < 2\gamma
\]  
(21)

for all \( i \in I_n, t, t' \in T \) and \( x, x' \in X \) such that \( \max\{d_T(t, t'), d_X(x, x')\} < \delta \). Thus, \( \varphi \) is a strong \( 8\gamma \) - equilibrium of \( \hat{G}_n \). Furthermore, \( |\hat{U}_i^n(\varphi|t) - U_i^n(\varphi|t)| \leq \int_X |\hat{U}_i^n(x, \varphi_{-i}|t) - U_i^n(x, \varphi_{-i}|t)|d\varphi(x|t) < 2\gamma \) for all \( i \in I_n \) and \( t \in T \).

Since \( |\hat{T}| = p \) and \( |\hat{X}| = m \), it follows by Lemma 2 that there is a \( \gamma \) - purification \( g \) of \( \varphi \) satisfying \( g_i(t) = g_i(t_l) \) for all \( t \in \hat{T}_l \) and all \( l \in \{1, \ldots, p\} \). Thus, \( g \) is a strong \( 9\gamma \) - equilibrium of \( \hat{G}_n \) and \( |\hat{U}_i^n(g|t) - \hat{U}_i^n(\varphi|t)| < 9\gamma \). Furthermore, as above, we have that

\[
|\hat{U}_i^n(x, g_{-i}|t) - U_i^n(x', g_{-i}|t')| < 2\gamma
\]  
(22)

for all \( i \in I_n, t, t' \in T \) and \( x, x' \in X \) such that \( \max\{d_T(t, t'), d_X(x, x')\} < \delta \).

We first claim that \( g \) is an \( \epsilon \) – equilibrium of \( G_n \). Let \( i \in I_n, t \in T \) and \( x \in X \). Let \( l \in \{1, \ldots, p\} \) be such that \( t \in \hat{T}_l \) and \( i \in \{1, \ldots, m\} \) be such that \( x \in X_j \). Then, \( g_i(t_l) = g_i(t_l) \) and \( U_i^n(g_i(t_l), g_{-i}|t_l) > \hat{U}_i^n(g_i(t_l), g_{-i}|t_l) - 2\gamma \geq \hat{U}_i^n(x_j, g_{-i}|t_l) - 11\gamma > U_i^n(x, g_{-i}|t_l) - 13\gamma \) and \( 13\gamma < \epsilon \).

Finally, \( |U_i^n(g|t) - U_i^n(\varphi|t)| \leq |U_i^n(g|t) - \hat{U}_i^n(\varphi|t)| + |\hat{U}_i^n(\varphi|t) - \hat{U}_i^n(\varphi|t)| + |\hat{U}_i^n(\varphi|t) - U_i^n(\varphi|t)| < (2 + 9 + 2)\gamma = 13\gamma \), for all \( i \in I_n \) and \( t \in T \). Thus, \( |U_i^n(g|t) - U_i^n(\sigma|t)| < 20\gamma < \epsilon \) for all \( i \in I_n \) and \( t \in T \) and so \( g \) is an \( \epsilon \) – purification of \( \sigma \). ■

### 4 An example: Cournot Oligopoly

In this section we show that Theorem 1 can be applied to the Cournot oligopoly game described in the Introduction.
First, note that it belongs to the class \( \mathcal{G}_K \). Indeed, letting \( \text{id} \) denote the identity function in \( \mathbb{R} \) and \( \mu_X \) denote the marginal of \( \mu \) on \( X \) for all \( \mu \in \mathcal{M}(T \times X) \), we simply have to define \( V_i^n(t, x, \mu) = P\left(\frac{n-1}{n} \int \text{id} d\mu_X + \frac{x}{n}\right)x - C(t, x) \). Letting, for all \( c_{-i} = (t_{-i}, x_{-i}) \in C_{-i} \), \( \mu_{c_{-i}} \) denote the marginal emp\(_{c_{-i}}\),\( X \) of emp\(_{c_{-i}}\), on \( X \) for all \( \mu \in \mathcal{M}(T \times X) \), we easily obtain that \( \mu_{c_{-i}}(A) = \frac{|\{ j \neq i : x_j \in A\}|}{(n-1)} \) for all Borel measurable subsets \( A \) of \( X \) and that \( \int \text{id} d\mu_{c_{-i}} = \sum_{j \neq i} x_j / (n-1) \). Thus, for all \( i \in I_n \), \( t \in T \), \( x \in X \) and \( \sigma \in \Sigma \),

\[
U_i(x, \sigma_{-i}|t) = \int_{C_{-i}} V_i^n(t, x, \text{emp}_{c_{-i}}) d[\tau \otimes \sigma]_{-i} = \int_{C_{-i}} P\left(\frac{x + \sum_{j \neq i} x_j}{n}\right) d[\tau \otimes \sigma]_{-i}(t_{-i}, x_{-i}) - C(t, x) = \int_{T_{-i}} \left[ \int_{X_{-i}} P\left(\frac{x + \sum_{j \neq i} x_j}{n}\right) d\sigma_{-i}(x_{-i}|t_{-i}) \right] d\tau_{-i}(t_{-i}) - C(t, x),
\]

which, when \( \sigma \) is a pure strategy, corresponds to the expression given in the Introduction. Second, note that \( K = \{(t, x, \mu) \mapsto P\left(\frac{n-1}{n} \int \text{id} d\mu_X + \frac{x}{n}\right)x - C(t, x) : n \in \mathbb{N}\} \) is equicontinuous due to the continuity of the functions \( P \), \( C \) and \( \mu \mapsto \int \text{id} d\mu_X \) and the compactness of \( T \) and \( X \). Thus, it follows from Theorem 1 that all Bayesian Nash equilibria of the Cournot competition game can be approximately purified whenever the number of firms is sufficiently large.

5 Related Literature

In this section we discuss the relationship between our framework and result and those of Rashid (1983), Cartwright and Wooders (2002), Kalai (2004) and Wooders, Cartwright, and Selten (2006). The main findings are the following: (1) Our result allows us to generalize those of Rashid (1983) and Wooders, Cartwright, and Selten (2006). (2) Our framework is more general than that of Kalai (2004), but our result is formally unrelated to his. (3) Our framework and result is formally unrelated with that of Cartwright and Wooders (2002).
5.1 Relationship with Rashid (1983)

Rashid (1983) considers the framework described in Subsection 2.1 with the following assumptions: $T$ is a singleton and $X$ is finite. Since we allow both $T$ and $X$ to be an arbitrarily compact metric space, it follows that our framework is more general than his, thus, our Theorem 1 extends his purification result.

5.2 Relationship with Kalai (2004)

Kalai (2004) considers the framework described in Section 2.2 with the following assumptions: both $T$ and $X$ are finite. Thus, our framework is more general than his. However, our result is formally unrelated to his: Even though our result requires weaker assumptions, it reaches a weaker conclusion. In fact, our purification notion is weaker than the notion of self-purification used in Kalai (2004) (see Cartwright and Wooders (2004)).

5.3 Relationship with Wooders, Cartwright and Selten (2006)

The relationship between our framework with that of Wooders, Cartwright, and Selten (2006) is less direct when compared to those discusses above. Nevertheless, in this subsection we show how to obtain the purification result of Wooders, Cartwright, and Selten (2006). Their framework is defined by a compact metric type space $\Omega$ (which is referred to as the attribute space), a finite action space $S$ and a function $h$. Using our notation, let $T = \Omega$ and $X = S$. Players’ types are defined by an attribute function $\alpha : I_n \to \Omega$, with the interpretation that $\alpha_i$ is player $i$’s type. Thus, we can define $\tau_i = 1_{\alpha_i}$. Finally, for all pure strategies $f = (f_1, \ldots, f_n)$, player $i$’s payoff is $h(\alpha_i, f_i, w_{\alpha, f})$ where $w_{\alpha, f}(t, x) = |\{i \in I_n : \alpha_i = t \text{ and } f(i) = x\}|$ for all $(t, x) \in T \times X$.

We claim that, when the large game property of Wooders, Cartwright, and Selten (2006) is satisfied, $h$ can be used to define a continuous real-valued function $v$ on $T \times X \times M(T \times X)$. This claim can be established as follows: for all finite subsets $\hat{T}$ of $T$, $n \in \mathbb{N}$ and functions $z : \hat{T} \to \{0, \ldots, n\}$ such that $\sum_{t \in \hat{T}} z(t) = n$, let $M(\hat{T}, n, z)$ be the set of those $\nu \in M(T \times X)$ satisfying $\text{supp}(\nu) \subseteq \hat{T} \times X$, $n \nu(t, x) \in \{0, \ldots, n\}$
for all \((t, x) \in \hat{T} \times X\) and \(\sum_{x \in X} n \nu(t, x) = z(t)\). Then, we obtain the following three conclusions. First, note that \(v(t, x, \mu)\) is easily defined for all \(t \in T, x \in X\) and \(\mu \in M(\hat{T}, n, z)\). In fact, for each \(\mu \in M(\hat{T}, n, z)\), we can find an attribute function \(\alpha : I_n \to \hat{T}\) and a pure strategy \(f\) such that \(\mu = w_{\alpha,f}/n\).\(^4\) Thus, set \(v(t, x, \mu) = h(t, x, n \mu)\).

Second, due to the large game property, \(v\) is uniformly continuous in \(T \times X \times M(\hat{T}, n, z)\). This can be established as follows. Note that the expression

\[
\frac{1}{n} \sum_{x \in X} \sum_{\nu \in \alpha(N)} |w_{\alpha,f}(t, x) - w_{\alpha,g}(t, x)|
\]

(defined by \(\rho_{\alpha,f}(g, \nu)\) in Wooders, Cartwright, and Selten (2006)) is equal to \(\|\text{emp}_{\alpha,f} \cdot \text{emp}_{\alpha,g}\|_1 = \sum_{(t,x) \in T \times X} |\text{emp}_{\alpha,f}\{(t, x)\} - \text{emp}_{\alpha,g}\{(t, x)\}|\). Moreover, note that if \(\mu, \nu \in M(\hat{T}, n, z)\) and \(\rho(\mu, \nu) < \gamma\), then \(\|\mu - \nu\|_1 < |\hat{T}| |X| \gamma\) for all \(\gamma > 0\) such that \(\overline{B}_\gamma\{(t, x)\} \cap (\hat{T} \times X) = \{(t, x)\}\) for all \((t, x) \in \hat{T} \times X\). Therefore, for all \(\varepsilon > 0, t, t' \in T, x \in X\) and \(\mu, \nu \in M(\hat{T}, n, z)\) satisfying \(d_T(t, t') < \delta/|\hat{T}| |X|\) and \(\rho(\mu, \nu) < \delta/|\hat{T}| |X|\), where \(\delta > 0\) is given by the large game property and corresponding to \(\varepsilon/2\) (and is also chosen to satisfy \(\overline{B}_{\delta/|\hat{T}| |X|}\{(t, x)\} \cap (\hat{T} \times X) = \{(t, x)\}\) for all \((t, x) \in \hat{T} \times X\), we obtain that \(|v(t, x, \mu) - v(t', x, \nu)| \leq |v(t, x, \mu) - v(t, x, \nu)| + |v(t, x, \nu) - v(t', x, \nu)| < \varepsilon\).

Hence, \(v\) is indeed uniformly continuous in \(T \times X \times M(\hat{T}, n, z)\).

Third, the collection \(\{M(\hat{T}, n, z)\}\), with \(\hat{T}\) being a finite subset of \(T\), \(n \in \mathbb{N}\) and \(z : \hat{T} \to \{0, \ldots, n\}\) satisfying \(\sum_{t \in \hat{T}} z(t) = n\), is dense in \(\mathcal{M}(T \times X)\). In fact, the set of finitely supported probability measures on \(T \times X\) with rational values is dense in \(\mathcal{M}(T \times X)\) and any such probability measure belongs to some set \(M(\hat{T}, n, z)\).

Therefore, it follows from the above facts that \(v\) can indeed be extended to a continuous real-valued function on \(T \times X \times \mathcal{M}(T \times X)\) (see DePree and Swartz (1988, Lemma 14, p. 279))). Letting \(v\) also denote such extension, we clearly obtain that \(K = \{v\}\) is equicontinuous.

Finally, we claim that players’ payoffs are defined by (7). Note that if \(\text{emp}_{\alpha_i,f_i} = \text{emp}_{\alpha_i,g_i}\) and \(f_i = g_i\), then \(w_{\alpha,f} = n\text{emp}_{\alpha,f} = n\text{emp}_{\alpha,g} = w_{\alpha,g}\) and so \(h(t, x, w_{\alpha,f}) = \)

\(^4\)The set of players is \(I_n\), the attribute function is chosen to satisfy \(|\{i \in I_n : \alpha_i = t\}| = z(t)\) and \(f\) is chosen to satisfy \(|\{i \in I_n : \alpha_i = t \text{ and } f_i = x\}| = n\mu(t, x)\).
h(t, x, w_{α,g}) by the global interaction property. This result implies that we can actually write \( h(t, x, w_{α,f}) = v(t, x, \text{emp}_{(α,f)}) = v(t, x, \text{emp}_{(α_i,f_i)}) \). Thus, for every pure strategy \( f \) and player \( i \in I_n \) in an \( n \)-player game, her payoff is

\[
h(α_i, f_i, w_{α,f}) = v(α_i, f_i, \text{emp}_{α_i,f_i}) = \int_{T_{-i} \times X_{-i}} v_1(α_i, f_i, \text{emp}_{t_{-i},x_{-i}}) d[τ ⊗ f]_{-i}(t_{-i}, x_{-i}),
\]

which is precisely (7) specialized to the case of a pure strategy.

Thus, it follows from Theorem 1 that for all \( ε > 0 \), there exists \( N \in \mathbb{N} \) such that all (Bayesian) Nash equilibria of any game with \( n \geq N \) players induced by \((Ω, S, h)\) can be \( ε \)-purified.

### 5.4 Relationship with Cartwright and Wooders (2002)

We finally comment on the purification result by Cartwright and Wooders (2002). Although their result neither is implied by nor implies our main result, the following comparison can be made. In fact, their result is weaker to the extent that their type space is finite and their action space is countable, whereas they can be countable or uncountable in ours. However, their result is stronger in the following aspects: their action space is not required to be compact and players’ types are not required to be independent. Furthermore, some comparison can be made regarding the continuity assumptions used in the two papers. For all \( μ, ν \in \mathcal{M}(T \times X) \), let

\[
s(μ, ν) = \sup_{(t,x) ∈ T \times X} |μ(\{(t, x)\}) − ν(\{(t, x)\})|.
\]

When \( T \times X \) is finite, endowing it with the metric \( s \) or \( ρ \) yields the same notion of continuity. Furthermore, when \( T \times X \) is countable, \( s \) is also a metric and the topology on \( \mathcal{M}(T \times X) \) it induces is stronger than that induced by \( ρ \).\(^5\) Hence, in the case where \( T \times X \) is countable, the continuity assumption used by Cartwright and Wooders (2002) is weaker than ours. Finally, when \( T \times X \) is not countable, \( s \) is a pseudo-metric but the topology it induces is neither weaker nor stronger than that induced by \( ρ \).

\(^5\)It can be easily shown that if \( s(μ_k, μ) \to 0 \), then \( \liminf_k μ_k(A) ≥ μ(A) \) for all Borel measurable subsets \( A \) of \( T \times X \). In particular, the inequality holds when \( A \) is open, and this implies that \( ρ(μ_k, μ) \to 0 \).
Our approach is, essentially, topological: once we strengthen Rashid’s purification result for the model with a finite number of actions, we use compactness and continuity to extend it. In contrast, Cartwright and Wooders (2002) address the purification problem directly by approximating mixed strategies with pure strategies. Their approach is important since it provides tools to directly address the countable case, and yields a purification result which is quite strong along many important dimensions. However, it seems that this strength needs to be compensated somehow, by placing assumptions on the cardinality of $T$ and $X$.

Finally, we note that our approach can be combined with theirs to obtain a purification result for games in the class $G_R$ with $\sigma$-compact actions spaces, i.e., the action space is the union of compact sets (an analogous remark has made before in Cartwright and Wooders (2002)). The reason is as follows: First, we can obtain a Bayesian game in the framework of Cartwright and Wooders (2002) starting with a game in $G_R$. The argument is simpler than the one in Subsection 5.3 since we do not need to extend players’ payoff functions to a larger domain, but rather, have to restrict them to a smaller one. Second, we can use their purification theorem to obtain a purification result for countably supported strategies (obtaining a result analogous to Lemma 2 in this fashion). Third, and finally, we can use the arguments used in the proof of Theorem 1 to extend the previous result to general (not necessarily countably supported) strategies on a $\sigma$-compact action space. Indeed, any strategy can be approximated with a strategy with a finite support within each compact set belonging to the countable collection whose union is the action space.

References


