ashoka
UNIVERSITY

Ashoka University Economics
Discussion Paper 79

A Deterministic Approximation Approach to the Continuum Logit Dynamic with an Application to Supermodular Games

April 2022

Ratul Lahkar, Ashoka University
Sayan Mukherjee, Indian Statistical Institute, Kolkata
Souvik Roy, Indian Statistical Institute, Kolkata

https://ashoka.edu.in/economics-discussionpapers

ashoka
UNIVERSITY
Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O.Rai, Sonepat, Haryana 131029

A Deterministic Approximation Approach to the Continuum Logit Dynamic with an Application to Supermodular Games

Ratul Lahkar*, Sayan Mukherjee† and Souvik Roy‡

April 4, 2022

Abstract

We consider the logit dynamic in a large population game with a continuum of strategies. The deterministic approximation approach requires us to derive this dynamic as the finite horizon limit of a stochastic process in a game with a finite but large number of strategies and players. We first establish the closeness of this dynamic with a step–wise approximation. We then show that the logit stochastic process is close to the step–wise logit dynamic in a discrete approximation of the original game. Combining the two results, we obtain our deterministic approximation result. We apply the result to large population supermodular games with a continuum of strategies. Over finite but sufficiently long time horizons, the logit stochastic process converges to logit equilibria in a discrete approximation of the supermodular game. By the deterministic approximation approach, so does the logit dynamic in the continuum supermodular game.

Keywords: Logit Dynamic; Deterministic Approximation; Supermodular Games.

JEL classification: C72; C73.

*(Corresponding Author) Department of Economics, Ashoka University, Rajiv Gandhi Education City, Sonipat, Haryana, 131029, India. e-mail: ratul.lahkar@ashoka.edu.in

†Economic Research Unit, Indian Statistical Institute, Kolkata. e-mail: sayanmukherjee012@gmail.com

‡Applied Statistics Unit, Indian Statistical Institute, Kolkata. e-mail: souvik.2004@gmail.com

1. Introduction

The deterministic approach to evolutionary game theory seeks to provide dynamical foundations to Nash equilibria through the use of evolutionary dynamics, which usually take the form of ordinary differential equations (Sandholm (2010)).[1] A large population of agents play a game and the distribution of strategies in that agent constitute the population state. Evolutionary dynamics are defined over such population states. The key focus of most deterministic evolutionary game theoretic models is whether the population state converges to a Nash equilibrium under different evolutionary dynamics.

Two important questions need to be addressed before admitting evolutionary dynamics as valid descriptions of the behavior of a large population. The first is about the behavioral foundations of evolutionary dynamics. This question is resolved by appealing to revision protocols which describe how agents revise strategies when they receive a chance to do so. For example, imitative revision protocols generate the replicator dynamic and the logit best response, which is a particular perturbation of the best response, generates the logit dynamic.[2] The second key question is about the mathematical foundations of such dynamics. Technically, these dynamics are defined for a large population or a continuum of agents, with each agent being of measure zero. But in reality, there is no population or society where there are a continuum of agents or where agents are of measure zero. Even if a population is large, the number of individuals will always be finite. How then to reconcile this aspect of reality with the mathematical abstraction of a continuum of agents?

This paper considers the second question in the context of the logit dynamic with a continuum of strategies (Perkins and Leslie (2014), Lahkar and Riedel (2015)). It does so by taking a *deterministic approximation* approach which we explain more fully below. This approach has been applied earlier for evolutionary dynamics defined on large population games with a finite strategy set (for example, Benaïm and Weibull (2003)). But to the best of our knowledge, this is the first paper that adopts such an approach to an evolutionary dynamic defined for a continuum of strategies. We then apply this approach to an important class of large population games; namely, supermodular games. Hofbauer and Sandholm (2007) analyze such games with finite strategy sets. We extend that notion to the context of continuous strategy sets and establish convergence of the continuum logit dynamic to logit equilibria, which are a perturbation of Nash equilibria, in such games. Again, as far as we know, this is the first paper to consider large population supermodular games with a continuum of strategies.

It is not necessary to have a continuum of measure zero agents to construct evolutionary game theoretic models. One can do so under the more realistic assumption of a finite number of agents by, for example, adopting a stochastic instead of a deterministic approach. In such an approach, we would define a stochastic process using a revision protocol to describe how the social state changes.[3] The focus of attention is then on the stationary distribution of the process and the weight that distribution places on the set of Nash equilibria. A problem with this approach is that it may require astronomically long periods of time to obtain convergence results under this approach; a time period so long that it is no longer meaningful

[1]Well known evolutionary dynamics include the replicator dynamic, the logit dynamic, the Brown–von Neumann–Nash (BNN) dynamic and the pairwise comparison dynamic. Sandholm (2010) provides an extensive exposition of such dynamics for finite strategy games. Such dynamics have also been extended to games with continuous strategy sets. See Cheung and Lahkar (2018) for a review. Not all evolutionary dynamics need to be differential equations. For example, the best response dynamic is a differential inclusion.

[2]See, for example, Lahkar and Sandholm (2008) for some such revision protocols that generate the different important evolutionary dynamics. Also see Cheung (2014) for an extension of the notion of revision protocols to large population games with a continuum of strategies.

[3]See Chapter 10 of Sandholm (2010) for further details of this methodology.

as a prediction for social or economic processes (Ellison (1993)). This is probably the reason why the deterministic approach to evolutionary game theory has been more widely used in economics.

Nevertheless, the stochastic approach can still be extremely useful in resolving the question we wish to address; that of providing the mathematical foundations to deterministic evolutionary dynamics. It is known, at least for large population games with a finite strategy set, that the behavior of a stochastic process can be approximated by a deterministic dynamic over finite time intervals provided the number of agents, while being finite, is large enough. This is the *deterministic approximation* approach to evolutionary dynamics. As a purely mathematical result, deterministic approximation was introduced by Kurtz (1970). Since then, various authors have extended such results to evolutionary game theory in the context of finite strategy games. They include Boylan (1995), Binmore et al. (1995), Börgers and Sarin (1997), Schlag (1998), and Sandholm (2003). The strongest and most general such results can be found in Benaïm and Weibull (2003). Thus, under this approach, a deterministic evolutionary dynamic may be interpreted as the approximation of the way a stochastic evolutionary process would behave over finite time periods.

The deterministic approximation approach, therefore, provides rigorous mathematical foundations to evolutionary dynamics. From a practical point of view, this approach is useful because very often, it is simpler to analyze an evolutionary dynamic instead of the underlying stochastic process. There is, however, a significant gap in the literature on deterministic approach. It is that, as far as we know and as we have already noted, all such deterministic approximation results are for large population games with a finite number of strategies. But by now, there is a significant literature on evolutionary dynamics with a continuum of strategies. These include the replicator dynamic and the more general class of imitative dynamics (Oechssler and Riedel (2001),Oechssler and Riedel (2002), Cheung (2016)), the BNN dynamic (Hofbauer et al. (2009)), the pairwise comparison dynamic (Cheung (2014)), the logit dynamic (Perkins and Leslie (2014), Lahkar and Riedel (2015)) and the more general class of perturbed best response dynamics to which the logit dynamic belongs (Lahkar et al. (2022)). Such dynamics have also been applied to a variety of classical models in economics like public goods, common resource problems and Cournot competition (Lahkar and Mukherjee (2019), Lahkar and Mukherjee (2021), Lahkar and Ramani (2021)). Hence, dynamics with a continuum of strategies are now a well–established part of evolutionary game theory, both from the theoretical perspective as well as the applied. Investigating their mathematical foundations through the deterministic approximation approach would, therefore, be a worthwhile exercise.

Hence, this paper extends the deterministic approximation approach to evolutionary dynamics with a continuum of strategies. In particular, we focus on the logit dynamic with a continuous strategy set. This dynamic is one of the canonical dynamics of learning and evolutionary game theory. It was introduced by Fudenberg and Levine (1998) for finite strategy games in the context of stochastic fictitious play and has since then been widely used to model evolution in large population games.[4] It is the prototype of the wider class of perturbed best response dynamics (Hofbauer and Sandholm (2002), Hofbauer and Sandholm (2007)) and is generated by perturbing payoffs with the Shannon entropy. Thus, it is an approximation of the best response dynamic but is always uniquely defined and, hence, is a differential equation. Therefore, it is more tractable than the best response dynamic, particulary in large population games where the best response may not even exist.

We focus on the logit dynamic because of its importance in evolutionary game theory. A larger goal, of course, would be to have a general deterministic approximation result for all continuous strategy evolutionary dynamics. Right now, however, that is beyond our reach. In fact, even deriving such a

[4]See, for example, Hofbauer and Hopkins (2005) and Hofbauer and Sandholm (2007). As noted earlier, this dynamic has been extended to continuous strategy sets by Perkins and Leslie (2014) and Lahkar and Riedel (2015).

result for the broader class of continuous strategy perturbed best response dynamics may not be feasible at present. Even though Lahkar et al. (2022) present a general class of continuous strategy perturbed best response dynamics using perturbations other than the Shannon entropy, closed form expressions of such dynamics are not easily derivable. Our analysis in the present paper, on the other hand, relies significantly on the closed form expression of the logit best response, upon which the logit dynamic is based. Nevertheless, the paper does illustrate the possibility of obtaining deterministic approximation results for continuous strategy evolutionary dynamics, even though the technical challenges of extending it more generally to other dynamics may be significant.

We consider a linear game with a continuum of strategies and played by a continuum of agents. Our objective is to provide rigorous foundations to the logit dynamic in this game. We first define a step–wise approximation of the original game and argue that solution trajectories of the original logit dynamic is close to the logit dynamic in the step–wise approximation. We then consider a discrete approximation of the original game with a finite number of strategies and a finite number of players and introduce the logit stochastic process in this game. Using the approach of Benaïm and Weibull (2003), we argue that as the number of players and strategies become large in this discrete game, the behavior of the stochastic process approximates the step–wise logit dynamic which, in turn, approximates the original continuum logit dynamic. This gives us our desired deterministic approximation result. As per this result, the continuum logit dynamic is an approximation of the logit stochastic process in a finite game as the number of strategies and players in that game go to infinity. We should note that even though we apply the finite strategy deterministic approximation framework of Benaïm and Weibull (2003), its extension to a continuum of strategies is not trivial. For example, even determining the notion of closeness in the continuous strategy case raises significant topological considerations like whether to choose the strong or weak topology. All our results are based on the strong topology or the topology induced by the total variation norm. As is typically the case, our deterministic approximation result is a finite horizon result. It holds for finite time intervals but may break down if the time span is of infinite length.

We then apply our deterministic approximation approach to the question of convergence of the continuum logit dynamic in supermodular games. Supermodular games are characterized by strategic complementarities and form an important class of games in economics with applications like search, coordination and modeling of positive externalities (Topkis (1998), Vives (1990), Milgrom and Roberts (1990)). Hofbauer and Sandholm (2007) introduce large population finite strategy supermodular games and establish that a stochastic process generated by a perturbed best response would converge to a perturbed best response equilibrium in a finite but sufficiently long period of time (medium run convergence).[5] We consider a continuous strategy supermodular game and apply the approach of Hofbauer and Sandholm (2007) to argue that the logit stochastic process will converge in the medium run to a logit equilibrium of the discrete approximation of the continuous strategy game. Our deterministic approximation result then allows us to conclude that the logit dynamic converges to logit equilibria in the continuous strategy supermodular game.

As noted earlier, this is the first paper that defines a continuous strategy supermodular games and, therefore, also the first paper that establishes any result on convergence of an evolutionary dynamic in such games. Hofbauer and Sandholm (2007) establish their medium run convergence result for the larger class of perturbed best response dynamics. This is not feasible for us because their result relies on deriving perturbed best response dynamics through stochastic perturbations. But a similar methodology is not yet

[5] A perturbed best response equilibrium is an approximation of a Nash equilibrium. It is a rest point of a perturbed best response dynamic. For the logit dynamic in particular, we refer to its rest points as logit equilibria.

known for games with a continuum of strategies.[6] Hence, we need to rely on the specific form of the logit best response for our analysis which, once again, illustrates the importance of focusing on the logit dynamic in this paper.

The rest of the paper is as follows. Section 2 introduces the model and establishes the step–wise approximation of the continuum logit dynamic. In Section 3, we derive our main deterministic approximation result. Section 4 defines large population supermodular games with a continuum of strategies and establishes the medium run convergence of the logit dynamic in such games. Section 5 concludes.

2. Preliminaries

We consider a population of unit mass with each agent in the population being of measure zero. Let $\mathscr{S} := [0,1] \subset \mathbb{R}$ be the strategy space of the players in the population.[7] We endow $\mathscr{S}$ with the usual Borel sigma-algebra $\mathscr{B}(\mathscr{S})$. Let $\mathscr{P}(\mathscr{S})$ denote the space of probability measures on $\mathscr{S}$. A probability measure $\mu \in \mathscr{P}(\mathscr{S})$ denotes the state of the population which we interpret as the distribution of strategies in the population. Hence, $\mu(A)$ is the proportion of players in the population using strategies in $A \subseteq \mathscr{S}$. Let $\mathscr{L}^\infty(\mathscr{S})$ be the collection of bounded measurable functions on $\mathscr{S}$. A *population game* is a weakly continuous mapping $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$ such that $\varphi_x(\mu)$ denotes the payoff to a player playing strategy x at population state μ. A population state μ° is said to be a Nash equilibrium of the underlying population game φ if for all $x \in \mathrm{Supp}(\mu^\circ)$ and all $y \in \mathscr{S}$, we have $\varphi_x(\mu^\circ) \geq \varphi_y(\mu^\circ)$.[8]

In this paper, we focus on games with a linear payoff structure in a pairwise random matching setting. We define the linear structure using a function $f : \mathscr{S} \times \mathscr{S} \to \mathbb{R}$ such that $f(x,y)$ is the payoff of an agent who plays x when randomly matched with another agent playing y. We can then write payoffs in the associated population game φ as the expected value

$$\varphi_x(\mu) = \int_{\mathscr{S}} f(x,y)\mu(dy), \quad \text{for all } x \in \mathscr{S}. \tag{1}$$

An example of such a linear game is as follows.

Example 2.1. Consider strategies $x, y \in \mathscr{S}$ and let $f(x,y) = m(x)y - c(x)$, where m is an increasing function of x and c is an arbitrary function of x. This linear payoff structure generates a population game φ in which the payoff of an agent playing strategy x at a population state μ is

$$\begin{aligned}\varphi_x(\mu) &= \int_{\mathscr{S}} f(x,y)\mu(dy) \\ &= m(x)\int_{\mathscr{S}} y\mu(dy) - c(x)\int_{\mathscr{S}} \mu(dy) \\ &= m(x)a(\mu) - c(x),\end{aligned} \tag{2}$$

where $a(\mu) = \int_{\mathscr{S}} y\mu(dy)$ is the aggregate strategy level in the population. We note that m being an increasing function in (2) is not required for φ to be linear. But we make this assumption in order to establish supermodularity later. ■

We focus on the logit dynamic for a large population game φ with a continuous strategy set. To

[6] Although Lahkar et al. (2022) extend continuous strategy perturbed best response dynamics beyond the logit dynamic, their extension is still based on deterministic perturbations.

[7] The strategy set being $[0,1]$ is a convenient normalization. Our arguments extend to any strategy set that is a compact convex subset of $\mathbb{R}$.

[8] We denote the support of a probability measure μ by $\mathrm{Supp}(\mu)$.

define the dynamic, fix a parameter $\eta > 0$. Then the *continuum strategy logit best response* is a mapping $\mathfrak{L}_\eta : \mathscr{P}(\mathscr{S}) \to \mathscr{P}(\mathscr{S})$ such that for all $\mu \in \mathscr{P}(\mathscr{S})$,

$$\mathfrak{L}_\eta[\mu](A) := \int_A \frac{\exp(\eta^{-1}\varphi_x(\mu))dx}{\int_{\mathscr{S}} \exp(\eta^{-1}\varphi_y(\mu))dy}, \quad \text{for all } A \in \mathscr{B}(\mathscr{S}). \tag{3}$$

Intuitively, the logit best response is an approximation of the best response, particularly for η small. In that case, the probability measure $\mathfrak{L}_\eta[\mu]$ puts nearly all its mass on the set of best responses to μ. Unlike best responses, though, the logit best response is uniquely defined and varies smoothly with μ. From (3), we then obtain the *continuum strategy logit dynamic* (Perkins and Leslie (2014); Lahkar and Riedel (2015))

$$\dot{\mu} = \mathfrak{L}_\eta(\mu) - \mu \tag{4}$$

induced by φ. Like any other evolutionary dynamic, the logit dynamic also measures the rate of change of the population state. Under the logit dynamic, a population state moves in the direction of the logit best response.

The definition of the logit dynamic does not depend upon φ being linear. But our subsequent results will depend greatly on the linear structure. A probability measure μ° is said to be a *logit equilibrium* if $\mathfrak{L}_\eta(\mu^\circ) = \mu^\circ$. Theorem 3.1 in Lahkar and Riedel (2015) establishes the existence of such equilibria. As $\eta \to 0$, any such logit equilibrium converges weakly to a Nash equilibrium of φ (Theorem 3.2, Lahkar and Riedel (2015)). We use the notation $\mathrm{LE}_\eta(\varphi)$ to denote the collection of logit best response equilibrium.

2.1 Step-wise Approximation of the Logit Dynamic

We now introduce a step–wise approximation of the logit dynamic that will be useful for our analysis that follows. Our objective is to show that this step–wise approximation is close to the original logit dynamic (4).The methodology we apply has certain similarities with the one employed by Oechssler and Riedel (2002) to establish the closeness of the continuous strategy replicator dynamic with its step–wise approximation.[9]

For this purpose, consider $n \geq 1$, $j \in \{0, 1, \ldots, 2^n - 1\}$ and denote the dyadic rational number $\frac{j}{2^n}$ by $\alpha_{n,j}$. Further, let S_n be the collection of left-closed right-open intervals with dyadic rational endpoints in $\mathscr{S}$ defined in the following way:

$$\mathscr{S}_n := \left\{[\alpha_{n,j}, \alpha_{n,j+1}) : j = 0, \ldots, 2^n - 2\right\} \cup [\alpha_{n,2^n-1}, 1]. \tag{5}$$

Thus, $\mathscr{S}_n$ is a collection of 2^n disjoint intervals which covers the original strategy space $\mathscr{S}$. Recall the payoff $f(x,y)$ in (1). Our next step is to construct a step function approximation of f. For $n \geq 1$, let $f_n : \mathscr{S} \times \mathscr{S} \to \mathbb{R}$ is such that for all $(x,y) \in \mathscr{S} \times \mathscr{S}$,

$$f_n(x,y) = \begin{cases} f(\alpha_{n,j}, \alpha_{n,k}) & \text{if } (x,y) \in [\alpha_{n,j}, \alpha_{n,j+1}) \times [\alpha_{n,k}, \alpha_{n,k+1}) \text{ and } j,k = 0,1,\ldots,2^n-2 \\ f(\alpha_{n,2^n-1}, \alpha_{n,2^n-1}) & \text{if } (x,y) \in [\alpha_{n,2^n-1}, 1] \times [\alpha_{n,2^n-1}, 1]. \end{cases} \tag{6}$$

Hence, by definition, $(f_n)_{n\geq 1}$ is a sequence of step functions such that $f_n \to f$ uniformly as $n \to \infty$.

Now, similar to the definition of the linear population game φ in (1), let $\varphi^n : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$ be

[9] See Section 7 of Oechssler and Riedel (2002).

such that for all $\mu \in \mathscr{P}(\mathscr{S})$,

$$\varphi_x^n(\mu) := \int_{\mathscr{S}} f_n(x,y)\mu(dy), \quad \text{for all } x \in \mathscr{S}. \tag{7}$$

Note from (6) that we can write (7) as

$$\varphi_x^n(\mu) = \begin{cases} \sum_{k=0}^{2^n-1} f(\alpha_{n,j},\alpha_{n,k}) \int_{[\alpha_{n,j},\alpha_{n,j+1})} \mu(dy) \text{ if } x \in [\alpha_{n,j},\alpha_{n,j+1}) \text{ and } j = 0,1,\ldots,2^n-2 \\ \sum_{k=0}^{2^n-1} f(\alpha_{n,2^n-1},\alpha_{n,k}) \int_{[\alpha_{n,2^n-1},1]} \mu(dy) \text{ if } x \in [\alpha_{n,2^n-1},1]. \end{cases} \tag{8}$$

We note that for every $\mu \in \mathscr{P}(\mathscr{S})$, φ^n is also a step function approximation of φ as defined in (1). This follows from the definition of f_n in (6). For a given $n \geq 1$ and parameter $\eta > 0$, we can then define the *logit best response induced by* φ_n as a mapping $\mathfrak{L}_\eta^n : \mathscr{P}(\mathscr{S}) \to \mathscr{P}(\mathscr{S})$ such that for all $\mu \in \mathscr{P}(\mathscr{S})$,

$$\mathfrak{L}_\eta^n[\mu](A) := \int_A \frac{\exp(\eta^{-1}\varphi_x^n(\mu))dx}{\int_{\mathscr{S}} \exp(\eta^{-1}\varphi_y^n(\mu))dy}, \quad \text{for all } A \in \mathscr{B}(\mathscr{S}). \tag{9}$$

Comparing (3) and (9), it is evident that the probability density function of the logit best response induced by φ_n is also a step–wise approximation of the probability density function of the logit best response induced by φ. Finally for $n \geq 1$ and $\eta > 0$, we define the logit dynamic induced by φ_n as

$$\dot{\mu} = \mathfrak{L}_\eta^n(\mu) - \mu. \tag{10}$$

Letting $t \geq 0$ be time, we denote as $\mu(t) \in \mathscr{P}(\mathscr{S})$ and $\mu^n(t) \in \mathscr{P}(\mathscr{S})$, $t \geq 0$, the solution trajectories of the two logit dynamics (4) and (10) respectively from the initial state $\mu(0)$. Theorem 3.4 in Lahkar and Riedel (2015) establishes the existence of such unique solution trajectories of the logit dynamic from every initial state in $\mathscr{P}(\mathscr{S})$ provided the underlying population game is bounded and Lipschitz continuous with respect to the variational norm.[10] Such conditions are automatically satisfied if φ is linear as defined in (1) since the underlying two player game f, and hence $(f_n)_{n\geq 1}$ are bounded. Intuitively, it is also reasonable to expect that these two solution trajectories would be close to each other since the logit best response (9) is itself a step function approximation of (3). We formalize this intuition in the following theorem.

Proposition 2.1. *Consider the linear supermodular game φ as defined in (1) and its step function approximation φ^n as defined in (7). Denote by $(\mu(t))_{t\geq 0}$ the solution to the logit dynamic (4) induced by φ and by $(\mu^n(t))_{t\geq 0}$ the solution to the logit dynamic (10) induced by φ_n for $n \geq 1$. Suppose that $\mu(0) = \mu^n(0)$ for every $n \geq 1$. Then $(\mu^n(t))_{t\geq 0}$ converges uniformly on compact sets to $(\mu(t))_{t\geq 0}$ in the variational norm, that is, for every finite time horizon $T > 0$,*

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \|\mu(t) - \mu^n(t)\|_{TV} = 0.$$

Proof. See Appendix 6.1. ■

Proposition 2.1 is our desired result on the step–wise approximation of the logit dynamic. It shows that the solution trajectories of the two dynamics are close to each other the strong topology, or the topology

[10] The total variation distance is a mapping $\|\cdot\|_{\mathbf{TV}} : \mathscr{P}(\mathscr{S}) \times \mathscr{P}(\mathscr{S}) \to \mathbb{R}_+$ defined as

$$\|\mu - \nu\|_{\mathbf{TV}} := 2 \sup_{A\in\mathscr{B}(\mathscr{S})} |\mu(A) - \nu(A)|, \quad \text{for all } \mu, \nu \in \mathscr{P}(\mathscr{S}).$$

induced by the total variation norm. Of course, the result requires that the initial states be identical. Also, n needs to be sufficiently large so that the step function f_n defined in (6) is a reasonably close approximation of the original linear payoff f in (1). Intuitively, the probability mass on each interval of the finite set S_n defined in (5) under the two trajectories $\mu(t)$ and $\mu^n(t)$ become increasingly identical as $n \to \infty$.

3. Deterministic Approximation

The deterministic approximation approach to the continuum logit dynamic would require us to approximate it as the behavior of a stochastic process in a discrete game where the number of strategies and the number of players are finite but sufficiently large. With this objective, we now define the discrete analogue of our population game φ. Consider the finite set

$$S_n = \{\alpha_{n,0}, \alpha_{n,1}, \alpha_{n,2}, \ldots, \alpha_{n,2^n-1}\}, \tag{11}$$

where, as defined previously, $\alpha_{n,j} = \frac{j}{2^n}$. Thus, S_n is a finite approximation of $\mathscr{S} = [0,1]$. Like $\mathscr{S}_n$ in (5), it contains 2^n elements with the difference being it consists of points instead of intervals. Each element of S_n is the leftmost point of the 2^n intervals in $\mathscr{S}_n$. Now, let

$$\Delta(S_n) = \left\{ \mathbf{p}^n \in \mathbb{R}_+^{2^n} : \sum_{0 \le k \le 2^n-1} \mathbf{p}_i^n = 1 \right\} \tag{12}$$

be the set of probability distributions on S_n. Hence, $\mathbf{p}_j^n$ is the probability on $\alpha_{n,j}$ for $j \in \{0, 1, \ldots, 2^n - 1\}$. Recalling the step function f_n from (6), we then define payoffs in the discrete version of φ as follows.

Definition 3.1. For $n \geq 1$, the **level-n discretization of a population game** φ is the mapping $\mathscr{D}_\varphi^n : \Delta(S_n) \to \mathbb{R}^{2^n}$ such that for all $\mathbf{p}^n \in \Delta(S_n)$,

$$\mathscr{D}_\varphi^n[\mathbf{p}^n](\alpha_{n,j}) := \langle f_n(\alpha_{n,j}, \cdot), \mathbf{p}^n \rangle = \sum_{0 \le k \le 2^n-1} f(\alpha_{n,j}, \alpha_{n,k}) \mathbf{p}_k^n, \quad \text{for all } j = 0, 1, \ldots, 2^n - 1. \quad \blacksquare \tag{13}$$

We can, therefore, envisage $\mathscr{D}_\varphi^n$ as the population game induced by the matrix game whose entries are given by $(f_n(\alpha_{n,j}, \alpha_{n,k}))_{0 \le j,k \le 2^n-1}$. Equivalently, $\mathscr{D}_\varphi^n$ is a finite strategy population game in which players' strategy set is S_n and the proportion of players playing $\alpha_{n,j}$ is $\mathbf{p}_j^n$. Notice from (8) and (13) that if we have $\mu \in \mathscr{P}(\mathscr{S})$ and $\mathbf{p}^n \in \Delta(S_n)$ such that $\mathbf{p}_j^n = \int_{[\alpha_{n,j}, \alpha_{n,j+1})} \mu(dy)$ for $j = 0, 1, \ldots, 2^n - 2$, then the payoffs in φ^n and $\mathscr{D}_\varphi^n$ are identical.

We focus on the discrete game $\mathscr{D}_\varphi^n$ introduced in Definition 3.1. Further, we now assume that the game is being played by a finite number of players, with $N > 1$ being the number of players . Recall our interpretation of $\mathbf{p}_j^n$ as being the mass of agents playing $\alpha_{n,j} \in S_n$ in $\mathscr{D}_\varphi^n$. For $n \geq 1$, and $\mathbf{p}^n \in \Delta(S_n)$, we also define the atomic probability measure on $\mathscr{S}$ as

$$\mu_{\mathbf{p}^n}^N := \sum_{0 \le j \le 2^n-1} \mathbf{p}_j^n \delta_{\alpha_{n,j}}, \tag{14}$$

where δ_x denotes the Dirac measure at $x \in \mathscr{S}$ and the superscript N indicates the dependence on the number of players N.[11] Thus, the measure $\mu_{\mathbf{p}^n}$ on $\mathscr{S}$ puts probability $\mathbf{p}_j^n$ on $\alpha_{n,j} = \frac{j}{2^n}$, $j \in \{0, 1, \cdots, 2^n - 1\}$ and 0

[11] This dependence will arise from the fact that with N being the number of players, $\mathbf{p}_j^n$ can take values only in $\{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\}$.

on all other points. This atomic measure now allows us to treat this discretized game as being played on $\mathscr{S}$ with the proportion of players using strategy $\alpha_{n,j}$ being $\mathbf{p}^n_j$.

We now introduce the logit best response on the discrete game $\mathscr{D}^n_\varphi$. Given the parameter $\eta > 0$ and, this logit best response $L^n_\eta : \Delta(S_n) \to \Delta(S_n)$ at the state $\mathbf{p}^n$ takes the standard finite dimensional form (Fudenberg and Levine (1998))

$$\begin{aligned} L^n_{\eta,i}[\mathbf{p}^n] &= \frac{\exp(\eta^{-1}\mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j}))}{\sum_{0\le k\le 2^n-1}\exp(\eta^{-1}\mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,k}))} \\ &= \frac{\exp(\eta^{-1}\langle f_n(\alpha_{n,i},\cdot),\mathbf{p}^n\rangle)}{\sum_{0\le k\le 2^n-1}\exp(\eta^{-1}\langle f_n(\alpha_{n,k},\cdot),\mathbf{p}^n\rangle)}. \end{aligned} \tag{15}$$

Thus, given η, (15) is the probability with which strategy $\alpha_{n,i} \in \mathscr{S}$ would get chosen under the logit best response in the population game $\mathscr{D}^n_\varphi$ at the population state $\mu^N_{\mathbf{p}^n}$.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be an arbitrary probability space. We now consider the following discrete time stochastic process $(\mathbb{X}^N_n(k))_{k\ge 1}$, where $\mathbb{X}^N_n(k) : \Omega \to \mathscr{P}(\mathscr{S})$ describes the state of the population at time k for every $k \ge 1$. We use the discrete logit best response (15) to define the transition probability function as

$$\mathbb{P}\left(\mathbb{X}^N_n\left(k+\frac{1}{N}\right) = \mu^N_{\mathbf{q}^n} | \mathbb{X}^N_n(k) = \mu^N_{\mathbf{p}^n}\right) = \begin{cases} \mathbf{p}^n_i L^n_{\eta,j}[\mathbf{p}^n], \text{ if } \mathbf{q}^n = \mathbf{p}^n + \dfrac{1}{N}(\mathbf{e}_j - \mathbf{e}_i) \text{ and } j \ne i \\ \sum_{i=0}^{2^n-1} \mathbf{p}^n_i L^n_{\eta,i}[\mathbf{p}^n] \text{ if } \mathbf{q}^n = \mathbf{p}^n \\ 0, \text{ otherwise.} \end{cases} \tag{16}$$

To understand these transition probabilities, we suppose that at discrete time intervals of length $\frac{1}{N}$, a single agent gets the opportunity to change strategy.[12] If that agent is currently playing strategy $\alpha_{n,i}$ and chooses to shift to $\alpha_{n,j}$, then the social state would change from $\mu_{\mathbf{p}^n}$ to $\mu_{\mathbf{q}^n}$, with $\mathbf{q}^n = \mathbf{p}^n + \frac{1}{N}(\mathbf{e}_j - \mathbf{e}_i)$. The probability of this transition is the mass of agents playing $\alpha_{n,i}$, which is $\mathbf{p}^n_i$, multiplied by the logit probability of playing $\alpha_{n,j}$, which is $L^n_{\eta,j}[\mathbf{p}^n]$. Alternatively, the agent can continue playing the same strategy, the probability of which is given by the second line of (16).

In the rest of this section, we show that the behavior of the logit stochastic process in $\mathscr{D}^n_\varphi$ approximates the behavior of the continuum logit dynamic (4) in φ over finite time spans as the number of players, N, and the number of strategies, 2^n, becomes large. That would be the interpretation of the continuum logit dynamic as a deterministic approximation of the logit stochastic process in a finite strategy and finite player game. For that, we now extend the process (16) defined over discrete time to one over continuous time. Thus, let $(\widehat{\mathbb{X}}^N_n(t))_{t\ge 0}$ denote the continuous time Markov process obtained by affine interpolation of the process $(\mathbb{X}^N_n(k))_{k\ge 1}$ as defined in (16). We call this process the **interpolated logit stochastic process**. We establish our deterministic approximation result with respect to this interpolated process. Note that this process is defined on the discrete game $\mathscr{D}^n_\varphi$. Establishing that result requires us to define the notion of smoothing of a probability measure.

Definition 3.2. For $n \ge 1$, the **smoothing of a probability measure** is a mapping $\mathfrak{s}_n : \mathscr{P}(S_n) \to \mathscr{P}(\mathscr{S})$ such that the following two conditions are satisfied:[13]

[12] Here, $\mathbf{e}_j$, $j \in \{0,1,\dots,2^n-1\}$, is the standard basis vector on $\mathbb{R}^{2^n}$. Under our atomic measures, the strategies that receive positive probability are $\alpha_{n,j} = \frac{j}{2^n}$, $j \in \{0,1,\dots,2^n-1\}$. Hence, $\mathbf{e}_j$ assigns value 1 to strategy $\frac{j}{2^n}$ and 0 to all other strategies.

[13] Here, $\mathscr{P}(S_n)$ denotes the collection of atomic probability measures like (14) with support S_n.

- $\mathfrak{s}_n(\mu^N_{\mathbf{p}^n})$ is absolutely continuous with respect to the Lebesgue measure λ, for all $\mathbf{p}^n \in \Delta(S_n)$.[14]
- The Radon-Nikodym derivative of the probability measure $\mathfrak{s}_n(\mu^N_{\mathbf{p}^n})$ is expressed as

$$\frac{\mathfrak{s}_n(\mu^N_{\mathbf{p}^n})}{d\lambda}(x) := \sum_{0\leq k\leq 2^n-2} 2^n \mu^N_{\mathbf{p}^n,k} \mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x), \quad \text{for all } x \in \mathscr{S}. \quad \blacksquare \tag{17}$$

The smoothing of a probability measure transforms atomic measures on $\mathscr{S}$ into probability measures admitting a probability density function. Suppose that we have an atomic measure $\mu_{\mathbf{p}^n}$ on $\mathscr{S}$, where the probabilities are concentrated on the points $\alpha_{n,j}$. Then $\mathfrak{s}_n(\mu_{\mathbf{p}^n})$ denotes the probability measure obtained by distributing the probability mass at $\alpha_{n,j}$ uniformly on the interval $[\alpha_{n,j}, \alpha_{n,j+1})$. We, thereby, construct a probability measure with a probability density function that is a step function. This density function is given by (17). Applying Definition 3.2 to $(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$, we can derive the smoothed interpolated logit stochastic process. As an intermediate step towards our desired result, we first establish a deterministic approximation result with respect to the step–wise logit dynamic (10).

Proposition 3.1. *Fix $n, N \geq 1$. Suppose that the interpolated logit stochastic process $(\widehat{\mathbb{X}}^N_n(k))_{k\geq 1}$ has as initial condition the atomic measure $\mu^N_{\mathbf{p}^n} \in \mathscr{P}(\mathscr{S})$, for some $\mathbf{p}^n \in \Delta(S_n)$. Let $(\mu^n(t))_{t\geq 0}$ denote the solution to the level-n step logit dynamic (10) induced by the population game φ_n with initial condition $\mu^n(0) = \mu^N_{\mathbf{p}^n}$. Then for every $T > 0$, there exists $\bar{C} > 0$ (independent of $\mu^n(0)$) such that for all $\varepsilon > 0$, we have that*

$$\mathbb{P}_{\mu^N_{\mathbf{p}^n}}\left(\sup_{0\leq t\leq T} \left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)) - \mu^n(t)\right\|_{TV} \geq \varepsilon\right) \leq 2^{n+1} e^{-N\bar{C}\varepsilon^2}$$

for sufficiently large N.

Proof. See Appendix 6.2. $\blacksquare$

Proposition 3.1, therefore, approximates the interpolated logit stochastic process in $\mathscr{D}^n_\varphi$ with the step–wise deterministic logit dynamic (10) over finite time horizons as the number of strategies and players in $\mathscr{D}^n_\varphi$ become large. Its proof can broadly be divided into two steps. First, we show that for every $\eta > 0$, the level-n step logit choice $\mathfrak{L}^n_\eta$ is "approximately Lipschitz continuous" for sufficiently large n in a sense made precise in Appendix 6.2. Then, along the lines of Benaïm and Weibull (2003), we bound the total variation norm between the smoothed stochastic evolutionary process and the solution to the level-n step logit dynamic uniformly over finite time horizons using Grönwall's inequality and some other relevant estimates to complete the proof.

Thus, Proposition 3.1 is a deterministic approximation result but with respect to the step–wise logit dynamic and not the original continuum logit dynamic. As noted in the Introduction, such results are well known in finite strategy games where a stochastic evolutionary process is well approximated by a deterministic evolutionary dynamic over finite time horizons. But as far as we know, a deterministic approximation result hasn't been established for any of the canonical continuous strategy evolutionary dynamics. We are, however, now in a position to use Proposition 3.1 along with Proposition 2.1 to state such a theorem for the actual continuum logit dynamic and not just its step–wise approximation.

Theorem 3.1. *Fix $n, N \geq 1$. Suppose that the interpolated logit stochastic process $(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$ has initial condition $\mu^N_{\mathbf{p}^n}$, for some $\mathbf{p}^n \in \Delta(S_n)$. Let $(\mu(t))_{t\geq 0}$ denote the solution to the continuum strategy logit*

[14] A probability measure μ is said to be absolutely continuous with respect to the Lebesgue measure λ if for all $A \in \mathscr{B}(\mathscr{S})$, we have $\lambda(A) = 0 \implies \mu(A) = 0$.

dynamic induced by the population game φ with initial condition $\mu(0) = \mu^N_{p^n}$. Then for every $T > 0$, and every $\varepsilon > 0$, we have that

$$\lim_{n\to\infty}\lim_{N\to\infty}\mathbb{P}_{\mu^N_{p^n}}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)) - \mu(t)\right\|_{TV} \ge \varepsilon\right) = 0.$$

Proof. See Appendix 6.3. ■

Theorem 3.1 is our desired deterministic approximation result and is a consequence of Propositions 2.1 and 3.1. By Proposition 3.1, the step–wise logit dynamic provides a deterministic approximation of the logit stochastic process. But Proposition 2.1 implies that the step–wise logit dynamic itself is an arbitrarily close approximation of the original continuum logit dynamic. Hence, intuitively, the original logit dynamic should also be a deterministic approximation of the logit stochastic process. Notice that in Proposition 3.1, we are also able to establish the rate of convergence of the step–wise dynamic to the stochastic process. This is due to the simpler structure of the step–wise dynamic. In comparison, Theorem 3.1 is a weaker result as it only establishes convergence of the original logit dynamic and not the rate of convergence. Nevertheless, it suffices in establishing the conclusion we desire.

This theorem provides microfoundations to the continuum logit dynamic (4). It implies that over finite time horizons, the behavior of the logit stochastic process (16) or, equivalently, its smoothed version dynamic $\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))$ in $\mathscr{D}^n_\varphi$, is arbitrarily well approximated by the solution trajectory from the same initial state of the logit dynamic (4) in the continuum game φ provided the number of strategies and the number of players in $\mathscr{D}^n_\varphi$ are sufficiently high. Thus, the continuum logit dynamic may be interpreted as approximating evolution in situations where the underlying population game have a finite but large number of strategies and players and where the strategic behavior of players is governed by the logit stochastic process. This is important because the direct analysis of the stochastic process itself may be difficult. In such situations, we can make the problem significantly more tractable by focusing, instead, on the behavior of the continuum logit dynamic.

4. Medium Run Convergence in Supermodular Games

As an application of our deterministic approximation approach, we now consider supermodular games. Our objective is to show that over sufficiently large but finite time horizon, the continuum logit dynamic converges to the set of logit equilibria in supermodular games. For finite strategy supermodular games, Hofbauer and Sandholm (2007) establish results on convergence in such games for the class of stochastically perturbed best response dynamics.[15] But since we do not have a general theory of stochastic perturbation of best response in continuous strategy games, we confine ourselves to the logit dynamic.

To define supermodular games with a continuous strategy set, we introduce the notation $\bar{F}_\mu : \mathscr{S} \to [0,1]$ to denote the inverse cumulative distribution function of a probability measure $\mu \in \mathscr{P}(\mathscr{S})$. Thus, $\bar{F}_\mu(x) := \mu((x,1])$ for all $x \in \mathscr{S}$. Let $\preceq$ be a partial order on $\mathscr{P}(\mathscr{S})$ defined as follows:

$$[\mu \preceq \nu] \iff [\bar{F}_\mu(x) \le \bar{F}_\nu(x), \quad \text{for all } x \in \mathscr{S}]. \tag{18}$$

Therefore, under this partial order, a population state ν is "ranked higher" than another state μ if the proportion of agents playing higher strategies is greater under ν than μ. We then arrive at the following

[15]For other canonical evolutionary dynamics like the replicator dynamic, the Brown–von Neumann–Nash dynamic, payoff comparison dynamics and the best response dynamic, general results on convergence in supermodular games are not available.

definition of a large population supermodular game, which is an extension of the definition provided by Hofbauer and Sandholm (2007) to our setting of a continuous strategy set. Intuitively, the definition implies that in a supermodular game, higher strategy choices by other agents make a higher strategy more attractive for every agent. Such games are, therefore, characterized by strategic complementarities.

Definition 4.1. A population game $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$ is said to be a **continuum supermodular game on** $\mathscr{S}$ if for all $\mu, \nu \in \mathscr{P}(\mathscr{S})$, the following condition is satisfied:

$$[\mu \preceq \nu] \implies [\varphi_y(\mu) - \varphi_x(\mu) \leq \varphi_y(\nu) - \varphi_x(\nu), \quad \text{for all } 0 \leq x < y \leq 1]. \tag{19}$$

As an illustration, the game in Example 2.1 is a supermodular game provided the function $m(x)$ is strictly increasing. To check that the relevant payoff (2) in this example satisfies supermodularity, notice that $\varphi_y(\mu) - \varphi_x(\mu) = (m(y) - m(x))a(\mu) - (c(y) - c(x))$. Further, if $\mu \preceq \nu$, then $a(\mu) \leq a(\nu)$. Hence, with m increasing, it must then be that $\varphi_y(\mu) - \varphi_x(\mu) < \varphi_y(\nu) - \varphi_x(\nu)$ if $x < y$. Thus, (19) is satisfied.

Our result on convergence of the logit dynamic in supermodular games will depend upon our key deterministic approximation result as presented in Theorem 3.1. Since such convergence is over finite time horizons, we follow Hofbauer and Sandholm (2007) and call it medium run convergence. Hofbauer and Sandholm (2007) also extend such convergence to an infinite time horizon by directly computing the stationary distribution of a stochastic process like (16). We do not pursue this long run analysis here as that is independent of the deterministic approximation approach to the logit dynamic.

We now consider a linear supermodular game φ such as Example 2.1. Recall that $\mathrm{LE}_\eta(\varphi)$ is the set of logit equilibria in a continuum population game φ. For $\varepsilon > 0$, let $\mathscr{O}^\varepsilon(\mathrm{LE}_\eta(\varphi))$ be the open ball of radius ε under the total variation norm around $\mathrm{LE}_\eta(\varphi)$ in a linear supermodular game φ. Applying the methodology used to derive Theorem 3.1, we first discretize our supermodular game as in Definition 4.2. The following theorem establishes the relationship between a continuum supermodular game and its discrete version. As is to be expected, the two notions are equivalent. Before presenting the theorem, we state the definition of a discrete supermodular game.

Definition 4.2. A population game $\mathscr{D}^n_\varphi : \Delta(S_n) \to \mathbb{R}^{2^n}$ as defined in Definition 3.1 is a supermodular game if it exhibits strategic complementarities, that is, for all $\mathbf{p}^n, \mathbf{q}^n \in \Delta(S_n)$, the following condition is satisfied:

$$[\mathbf{p}^n \preceq \mathbf{q}^n] \implies [\mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j+1}) - \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j}) \leq \mathscr{D}^n_\varphi[\mathbf{q}^n](\alpha_{n,j+1}) - \mathscr{D}^n_\varphi[\mathbf{q}^n](\alpha_{n,j})], \tag{20}$$

for all $j = 0, 1, \ldots, 2^n - 2$. ■

In words, if $\mathbf{q}^n$ stochastically dominates $\mathbf{p}^n$, then for any strategy $\alpha_{n,j}$, the payoff advantage of $\alpha_{n,j+1}$ over $\alpha_{n,j}$ is greater at $\mathbf{q}^n$ than at $\mathbf{p}^n$. We can also write the strategic complementarity condition (20) equivalently by using partial derivatives as[16]

$$\frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j+1})}{\partial \mathbf{p}^n_{i+1}} - \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j+1})}{\partial \mathbf{p}^n_i} \geq \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j})}{\partial \mathbf{p}^n_{i+1}} - \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j})}{\partial \mathbf{p}^n_i}, \tag{21}$$

for all $i, j = 0, 1, \ldots, 2^n - 2$. Thus, if a small of mass of players switches from strategy $\alpha_{n,i}$ to $\alpha_{n,i+1}$, then the improvement in the payoff of strategy $\alpha_{n,j+1}$ is greater than that of $\alpha_{n,j}$. The result we now seek is as follows.

[16] See Section 3.3 in Hofbauer and Sandholm (2007) and Theorem 3.4.2 in Sandholm (2010).

Proposition 4.1. *Fix a weakly continuous population game $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$. Then φ is a supermodular if and only if for every $n \geq 1$, the level-n discretization $\mathscr{D}^n_\varphi : \Delta(S_n) \to \mathbb{R}^{2^n}$ of the population game φ is a supermodular game.*

Proof. See Appendix 6.4. ■

Recall now the step–wise approximation of φ, φ^n, as defined in (7) or (8). Also recall our comments about the relationship between the payoffs in $\mathscr{D}^n_\varphi$ and φ^n in the paragraph following Definition 3.1. We then obtain the following corollary from Proposition 4.1. The proof is obvious and is left to the reader.

Corollary 4.1. *Fix a population game $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$. Then φ is a supermodular if and only if φ_n is a supermodular game for every $n \geq 1$.*

Next, we introduce the discrete logit process $\mathbb{X}^N_n(k)$ as defined in (16) on the discrete supermodular game $\mathscr{D}^n_\varphi$, construct the interpolated logit stochastic process $\widehat{\mathbb{X}}^N_n(t)$ and smoothen the process as described in Definition 3.2 to obtain $\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))$. In addition, we also need the following definition of an *irreducible discrete supermodular game* from Hofbauer and Sandholm (2007). In stating this definition, we recall the finite strategy set $S_n = \{\alpha_{n,0}, \alpha_{n,1}, \ldots, \alpha_{n,2^n-1}\}$ over which we defined the discrete game $\mathscr{D}^n_\varphi$ in Definition 3.1. We also recall the notation $\alpha_{n,j}$ to denote the strategy $\frac{j}{2^n}$ in (5).

Definition 4.3. Let $\widehat{S}_n = S_n \setminus \alpha_{n,2^n-1} = \{\alpha_{n,0}, \alpha_{n,1}, \ldots, \alpha_{n,2^n-2}\}$. For $n \geq 1$, the level-n discretization of a supermodular game φ is said to be *irreducible* if for all states $\mathbf{p}^n \in \Delta(S_n)$ and all $j \in \widehat{S}_n$, there exists at least one $i \in \widehat{S}_n$ such that (21) holds with strict inequality.

Definition 4.3 is a mild strengthening of the strategic complementarity condition (21) for a discrete supermodular game. It requires that at every state $\mathbf{p}^n$, a movement from strategy j to $j+1$ improves the relative performance of at least one strategy. As an illustration, we show that the discrete version of the game in Example 2.1 does satisfy irreducibility.

Proposition 4.2. *For every $n \geq 1$, the discrete version $\mathscr{D}^n_\varphi$ of the game φ in Example 2.1 is irreducible.*

Proof. See Appendix 6.5. ■

With this irreducibility condition, we can now establish our main result about supermodular games which, following Hofbauer and Sandholm (2007), we call a medium run convergence theorem. Before considering that theorem, we state the following lemma which will be useful to us.

Lemma 4.1. *Let φ be a linear population game as defined in (1) and $\mathscr{D}^n_\varphi$ be its discretization. Suppose $(\boldsymbol{p}^*_n)_{n\geq 1}$ is a sequence of logit equilibria of the discretized games $(\mathscr{D}^n_\varphi)_{n\geq 1}$. Recall from (14) the atomic measure $\mu_{\boldsymbol{p}^*_n}$ derived from $\boldsymbol{p}^*_n$ on $\mathscr{S}$ and the smoothing $\mathfrak{s}_n(\mu_{\boldsymbol{p}^*_n})$ of $\mu_{\boldsymbol{p}^*_n}$ (Definition 3.2). Then the accumulation points of the sequence $(\mathfrak{s}_n(\mu_{\boldsymbol{p}^*_n}))_{n\geq 1}$ under the total variation norm is non-empty. Furthermore, any such accumulation point is a logit equilibrium of the game φ.*

Proof. See Appendix 6.6. ■

Lemma 4.1 is independent of supermodularity. It holds for any linear population game. It is well known that the set of probability measures $\mathscr{P}(\mathscr{S})$ is compact under the weak topology. Hence, the sequence $\mathfrak{s}_n(\mu_{\mathbf{p}^*_n})$ that we obtain from the logit equilibria of the discretized game $\mathscr{D}^n_\varphi$ will have an accumulation point in $\mathscr{P}(\mathscr{S})$. We show that this accumulation point is a logit equilibrium of the original game φ. In addition, we also need to show that in case of such logit equilibria, convergence under the weak topology also implies convergence under the strong topology, or the topology induced by the total variation norm. This follows because as $n \to \infty$, the probability mass put by the logit equilibrium $\mathbf{p}_n^*$ on any strategy in the finite strategy set of $\mathscr{D}_\varphi^n$ goes to zero.

We now establish our medium run convergence theorem. Suppose φ is a continuum supermodular game and that its discretized form $\mathscr{D}_\varphi^n$ satisfies irreducibility. Consider the interpolated logit stochastic process on $\mathscr{D}_\varphi^n$ and let the number of strategies and players in $\mathscr{D}_\varphi^n$ become large. The theorem states that over sufficiently long but finite time horizons (hence the name "medium run"), the continuum logit dynamic spends most of its time near the set of logit equilibria of φ.

Theorem 4.1. *Suppose that the two player game $f : \mathscr{S} \times \mathscr{S} \to \mathbb{R}$ be Lipschitz continuous. Suppose the the population game $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$ is supermodular game such that for every n, the level-n discretization of φ is irreducible. Then for every $\varepsilon > 0$, there exists $T_\varepsilon \geq 0$ such that for all $T \geq T_\varepsilon$, $(\mu(t))_{t\geq 0} \in \mathscr{O}^{3\varepsilon}(LE_\eta(\varphi))$ for all $T_\varepsilon \leq t \leq T$.*

Proof. See Appendix 6.7. ∎

The proof of Theorem 4.1 proceeds in several steps. First, we discretize φ into $\mathscr{D}_\varphi^n$. Since φ is supermodular, Proposition 4.1 implies that $\mathscr{D}_\varphi^n$ is also supermodular. We then consider the interpolated logit stochastic process $\widehat{\mathbb{X}}_n^N(t)$, define it appropriately on $\Delta(S_n)$ and allow it to run in $\mathscr{D}_\varphi^n$. By Theorem 3.1, once we smoothen $\widehat{\mathbb{X}}_n^N(t)$ and $N, n \to \infty$, the behavior of the stochastic process in $\mathscr{D}_\varphi^n$ is well approximated by the continuum logit dynamic (4) on φ. Therefore, it is legitimate to analyze the convergence of the logit dynamic on φ by considering the convergence of the logit stochastic process on $\mathscr{D}_\varphi^n$. We do so by appealing to results in Hofbauer and Sandholm (2007) for finite strategy supermodular games and applying some of our earlier results.

Since the discretized or finite–strategy game $\mathscr{D}_\varphi^n$ is supermodular, Theorem 4.1 (part (iii)) in Hofbauer and Sandholm (2007) implies that as the number of players $N \to \infty$, the logit stochastic process converges to the set of logit equilibria of $\mathscr{D}_\varphi^n$. We then show that once we smoothen a logit equilibrium of $\mathscr{D}_\varphi^n$, we obtain a logit equilibrium of the step–wise approximation of φ, φ^n. Corollary 4.1 then implies that φ^n is also supermodular. Hence, the smoothed interpolated process $\mathfrak{s}_n(\widehat{\mathbb{X}}_n(t))$ lies near the set of logit equilibria of φ^n over sufficiently long but finite time horizons as $N \to \infty$. By the deterministic approximation result for the step–wise logit dynamic (Proposition 3.1), $\mathfrak{s}_n(\widehat{\mathbb{X}}_n(t))$ is itself close to the trajectory of the step–wise logit dynamic, $\mu^n(t)$, over such finite time horizons. Hence, over a finite time horizon, $\mu^n(t)$ converges to the set of logit equilibria of φ^n. Now, as $n \to \infty$, a logit equilibrium of φ^n approaches a logit equilibrium of φ under the strong topology. This follows from Lemma 4.1. Moreover, as $n \to \infty$, $\mu^n(t)$ itself is close to $\mu(t)$ by Proposition 2.1. Combining all these arguments, we conclude that as both $N, n \to \infty$, a trajectory $\mu(t)$ of the continuum logit dynamic converges to a logit equilibrium of the supermodular game φ over finite but sufficiently long time periods.

Corollary 4.2. *Suppose $\varphi : \mathscr{P}(\mathscr{S}) \to \mathscr{L}^\infty(\mathscr{S})$ be a supermodular game such that for every n, the level-n discretization of φ is irreducible. Then for every $\varepsilon > 0$, there exists $T_\varepsilon \geq 0$ such that for all $T \geq T_\varepsilon$, we have*

$$\lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}_{\mu^n}\left(\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) \in \mathscr{O}^\varepsilon(LE_\eta(\varphi)) \text{ for all } T_\varepsilon \leq t \leq T\right) = 1.$$

Proof. The proof follows from Theorem 4.1. ∎

5. CONCLUSION

The deterministic approach to evolutionary game theoretic dynamics seeks to provide mathematical foundations to such dynamics as the limiting outcome of a stochastic process as the number of players become large. In this paper, we have adopted such a deterministic approach to the logit dynamic in a large population game with a continuous strategy set. We first construct a step–wise approximation of such a game and show that the logit dynamic in that approximate game is close the dynamic in the original game. We then introduce a discrete approximation of the original game with a finite number of strategies and players. Over finite time horizons, the logit stochastic process in this discrete game approximates the behavior of the step–wise logit dynamic, which itself approximates the original continuum logit dynamic. This gives us our deterministic approximation result. The continuum logit dynamic is approximation of the logit stochastic process of a game in which both the number of players and strategies are finite but large.

We then apply our approach to large population supermodular games with a continuous strategy set. If the discrete version of the supermodular game satisfies a condition called irreducibility, then over time horizons that are finite but sufficiently long, the continuum logit dynamic converges to the set of logit equilibria of the supermodular game. To establish this result, we consider the logit stochastic process in the discrete version of the game. Results from finite strategy supermodular games imply that this process converges. Our deterministic approximation result then show that the logit dynamic must also converge.

This paper is limited to the logit dynamic. An obvious research question for the future is to extend the deterministic approach to other well–known continuous strategy evolutionary dynamics. For example, should it prove feasible to establish such a result for the larger class on continuous strategy perturbed best response dynamics, then it may be possible to extend our result on convergence of the logit dynamic in supermodular games to such dynamics as well.

6. APPENDIX

6.1 PROOF OF PROPOSITION 2.1

Fix $A \in \mathscr{B}(\mathscr{S})$. We then observe that,

$$
\begin{aligned}
|\mu[t](A) - \mu^n[t](A)| &\leq \left| \int_0^t \dot{\mu}[s](A)\mathrm{d}s - \int_0^t \dot{\mu}^n[s](A)ds \right| \\
&= \left| \int_0^t \mathfrak{L}[\mu(s)](A)\mathrm{d}s - \int_0^t \mathfrak{L}^n[\mu^n(s)](A)ds \right| + \left| \int_0^t \mu[s](A) - \mu^n[s](A) \right| ds \\
&\leq \left| \int_0^t \mathfrak{L}[\mu(s)](A)ds - \int_0^t \mathfrak{L}[\mu^n(s)](A)ds \right| + \left| \int_0^t \mu[s](A) - \mu^n[s](A) \right| ds \\
&\qquad + \left| \int_0^t \mathfrak{L}(\mu^n(s))(A)\mathrm{d}s - \int_0^t \mathfrak{L}^n(\mu^n(s))(A)ds \right| \\
&\leq \int_0^t \underbrace{(\|\mathfrak{L}(\mu(s)) - \mathfrak{L}(\mu^n(s))\|_{\mathbf{TV}} + \|\mu(s) - \mu^n(s)\|_{\mathbf{TV}})}_{=:\mathscr{T}_1^n(s)} ds
\end{aligned}
$$

$$+\int_0^t \int_A \underbrace{\left| \frac{\exp(\varphi_x(\mu^n(s)))}{\int_{\mathscr{S}} \exp(\varphi_y(\mu^n(s)))dy} - \frac{\exp(\varphi_x^n(\mu^n(s)))}{\int_{\mathscr{S}} \exp(\varphi_y^n(\mu^n(s)))dy} \right|}_{=:\mathscr{T}_2^n(s,x)} dxds$$

It follows from Lahkar and Riedel (2015) that there exists $\kappa > 0$ such that

$$\mathscr{T}_1^n(s) := \|\mathfrak{L}(\mu(s)) - \mathfrak{L}(\mu^n(s))\|_{\mathbf{TV}} \leq (1+\kappa)\|\mu(s) - \mu^n(s)\|_{\mathbf{TV}}, \quad \text{for all } s \geq 0. \tag{22}$$

We now state and prove a lemma which provides an appropriate upper bound on the second term $\mathscr{T}_2(s,x)$ in order to apply Grönwall's lemma.

Lemma 6.1. *Fix $n \geq 1$. Then for every $x \in \mathscr{S}$ and every $s \geq 0$, it holds that:*

$$\mathscr{T}_2^n(s,x) \leq 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty). \tag{23}$$

Proof. Fix $x \in \mathscr{S}$. It then follows from the definition of $\mathscr{T}_2^n(s,x)$ that,

$$\begin{aligned}
\mathscr{T}_2^n(s,x) &= \left| \frac{\int_S \exp(\varphi_x(\mu^n(s)))\exp(\varphi_y^n(\mu^n(s)))dy - \int_S \exp(\varphi_x^n(\mu^n(s)))\exp(\varphi_y(\mu^n(s)))dy}{\left(\int_S \exp(\varphi_y(\mu^n(s)))dy\right)\left(\int_S \exp(\varphi_y^n(\mu^n(s)))dy\right)} \right| \\
&\leq \left| \int_S \exp(\varphi_x(\mu^n(s)))\exp(\varphi_y^n(\mu^n(s)))dy - \int_S \exp(\varphi_x^n(\mu^n(s)))\exp(\varphi_y(\mu^n(s)))dy \right| \\
&= \left| \int_S \exp[\varphi_x(\mu^n(s)) + \varphi_y^n(\mu^n(s))]dy - \int_S \exp[\varphi_x^n(\mu^n(s)) + \varphi_y(\mu^n(s))]dy \right| \\
&\leq \int_S \underbrace{|\exp[\varphi_x(\mu^n(s)) + \varphi_y^n(\mu^n(s))] - \exp[\varphi_x^n(\mu^n(s)) + \varphi_y^n(\mu^n(s))]|}_{=:\mathscr{T}_{21}^n(s,x)} dy \\
&\qquad + \int_S \underbrace{|\exp[\varphi_x^n(\mu^n(s)) + \varphi_y^n(\mu^n(s))] - \exp[\varphi_x^n(\mu^n(s)) + \varphi_y(\mu^n(s))]|}_{=:\mathscr{T}_{22}^n(s,x)} dy.
\end{aligned}$$

The first inequality in the above string of inequalities occurs since the payoff functions $(f_n)_{n\geq 1}$ and f are bounded below by 0, in which case the integrals $\int_S \exp(\varphi_y(\mu^n(s)))\mathrm{d}y$ and $\int_S \exp(\varphi_y^n(\mu^n(s)))\mathrm{d}y$ are bounded below by 1. Now we bound the terms $\mathscr{T}_{21}^n(s,x)$ and $\mathscr{T}_{22}^n(s,x)$ separately. For any bounded function $g : \mathscr{S} \to \mathbb{R}$, it follows by an application of mean-value theorem that

$$|\exp(g(x)) - \exp(g(y))| \leq \exp(\|g\|_\infty)|g(x) - g(y)|, \quad \text{for all } x,y \in \mathscr{S}. \tag{24}$$

Therefore by an application of (24), we have for all $s \geq 0$, and $x \in \mathscr{S}$, that

$$\begin{aligned}
\mathscr{T}_{21}^n(s,x) &= \int_{\mathscr{S}} |\exp(\varphi_y^n(\mu^n(s)))||\exp(\varphi_x(\mu^n(s))) - \exp(\varphi_x^n(\mu^n(s)))|dy \\
&\leq \exp\|f_n\|_\infty \exp\|f\|_\infty \vee \|f_n\|_\infty \int_{\mathscr{S}} |\varphi_x(\mu^n(s)) - \varphi_x^n(\mu^n(s))|dy \\
&\leq \exp(\|f_n\|_\infty + \|f\|_\infty \vee \|f_n\|_\infty)\|f - f_n\|_\infty \\
&\leq \exp(2\|f\|_\infty \vee \|f_n\|_\infty)\|f - f_n\|_\infty.
\end{aligned}$$

Proceeding similarly as above we observe that

$$\mathscr{T}_{22}^n(s,x) \le \exp(2\|f\|_\infty \vee \|f_n\|_\infty)\|f - f_n\|_\infty, \quad \text{for all } x \in \mathscr{S}.$$

Combining the above inequalities we have,

$$\mathscr{T}_2^n(s,x) \le \mathscr{T}_{21}^n(s,x) + \mathscr{T}_{22}^n(s,x) \le 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty), \quad \text{for all } x \in \mathscr{S}.$$

This completes the proof of Lemma 6.1. ∎

We now proceed with the proof of the proposition. Using Lemma 6.1 in conjunction with (22), we now arrive at

$$\begin{aligned}
|\mu[t](A) - \mu^n[t](A)| &\le \int_0^t \mathscr{T}_1^n(s)ds + \int_0^t \mathscr{T}_2^n(s,x)ds \\
&\le (1+\kappa)\int_0^t \|\mu(s) - \mu^n(s)\|_{\mathbf{TV}} ds + 2\int_0^t \|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) ds \\
&\le (\kappa+3)\int_0^t \Big[\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) + \|\mu(s) - \mu^n(s)\|_{\mathbf{TV}}\Big] ds.
\end{aligned}$$

Taking supremum over all $A \in \mathscr{B}(\mathscr{S})$, we have

$$\|\mu(t) - \mu^n(t)\|_{\mathbf{TV}} \le 2(\kappa+3)\int_0^t \Big[\exp(2\|f\|_\infty \vee \|f_n\|_\infty)\|f_n - f\|_\infty + \|\mu(s) - \mu^n(s)\|_{\mathbf{TV}}\Big] ds.$$

Define $\Lambda_n(t) := \|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) + \|\mu(t) - \mu^n(t)\|_{\mathbf{TV}}$ for all $0 \le t \le T$. From the above inequality, it then follows that

$$\Lambda_n(t) \le 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) + 2(\kappa+3)\int_0^t \Lambda_n(s)ds, \text{ for all } 0 \le t \le T.$$

As a result, by Grönwall's inequality, we have for all $0 \le t \le T$ that

$$\begin{aligned}
\Lambda_n(t) &\le 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) \times \exp(2t\kappa + 6t) \\
&\le 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) \times \exp(2T\kappa + 6T).
\end{aligned}$$

Since $f_n \to f$ uniformly, $\|f_n\|_\infty$ is uniformly bounded, in which case it follows using the definition of Λ_n that

$$\sup_{0 \le t \le T} \|\mu(t) - \mu^n(t)\|_{\mathbf{TV}} \le 2\|f_n - f\|_\infty \exp(2\|f\|_\infty \vee \|f_n\|_\infty) \times \exp(2(\kappa+3)T - 1) \to 0, \quad \text{as } n \to \infty.$$

This completes the proof of Proposition 2.1.

6.2 Proof of Proposition 3.1

We begin the proof of the proposition along the lines of Benaïm and Weibull (2003). To get rid of cumbersome notations in the proof, we shall, for every $n \ge 1$, denote the vector field $\mathscr{P}(\mathscr{S}) \ni \mu \mapsto \mathfrak{L}^n(\mu) - \mu \in \mathscr{M}_0(\mathscr{S})$ by $\mu \mapsto \mathfrak{L}^n_{\mathsf{vect}}(\mu)$ for all $\mu \in \mathscr{P}(\mathscr{S})$. For $N, n \ge 1$, consider the stochastic evolutionary process as defined $(\mathbb{X}_n^N(k))_{k\geq 1}$ as defined in (16). Now, let $\mathscr{U}_{n,k}^N : [0,\infty) \to \mathscr{M}(\mathscr{S})$ be a mapping defined as

$$\mathscr{U}_n^N(k\delta_N) = \frac{1}{\delta_N}[\mathbb{X}_n^N((k+1)\delta_N) - \mathbb{X}_n^N(k\delta_N)] - \mathfrak{L}_{\text{vect}}^n(\mathbb{X}_n^N(k\delta_N)), \text{ for all } k \geq 1. \tag{25}$$

We now define the continuous time process as $\mathscr{U}_n^N(t) := \mathscr{U}_n^N(k\delta_N)$ for all $k\delta_N \leq t < (k+1)\delta_N$. For $k \geq 1$, let $\mathscr{F}_k$ denote the σ-algebra generated by $\{\mathbb{X}_n^N(0), \ldots, \mathbb{X}_n^N(k\delta_N)\}$. To proceed with the proof, we also require the following notation. Let $\bar{\mathbb{X}}_n^N(t) := \mathbb{X}_n^N(k\delta_N)$, for all $k\delta_N \leq t < (k+1)\delta_N$, be the step process corresponding to $(\mathbb{X}_n^N(k))_{k\geq 1}$. It now follows that

$$\begin{aligned}
\widehat{\mathbb{X}}_n^N(t) - \mu_0^n &= \int_0^t [\mathfrak{L}_{\text{vect}}^n(\bar{\mathbb{X}}_n^N(s)) + \mathscr{U}_n^N(s)]ds \\
&= \int_0^t [\mathfrak{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s)) + \mathfrak{L}_{\text{vect}}^n(\bar{\mathbb{X}}_n^N(s)) - \mathfrak{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s)) + \mathscr{U}_n^N(s)]ds.
\end{aligned}$$

By definition of the smoothing map $\mathfrak{s}_n$, it follows that

$$\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \mathfrak{s}_n(\mu_{\mathbf{x}}^{N,n}) = \int_0^t [\mathfrak{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s)) + \mathfrak{L}_{\text{vect}}^n(\bar{\mathbb{X}}_n^N(s)) - \mathfrak{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s)) + \mathfrak{s}_n(\mathscr{U}_n^N(s))]ds.$$

Also since we have $\mu^n(t) - \mu_0^n = \int_0^t \mathfrak{L}_{\text{vect}}^n(\mu^n(s))ds$, we have that

$$\begin{aligned}
\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \mu^n(t)\|_{\mathbf{TV}} \leq &\int_0^t \|\mathfrak{s}_n(\mathscr{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s))) - \mathfrak{s}_n(\mathscr{L}_{\text{vect}}^n(\mu^n(s)))\|_{\mathbf{TV}}ds \\
&\int_0^t \|\mathfrak{s}_n(\mathscr{L}_{\text{vect}}^n(\bar{\mathbb{X}}_n^N(s))) - \mathfrak{s}_n(\mathscr{L}_{\text{vect}}^n(\widehat{\mathbb{X}}_n^N(s)))\|_{\mathbf{TV}}ds + \Psi(T),
\end{aligned}$$

where $\Psi(T) = \sup_{0\leq t\leq T} \|\int_0^T \mathfrak{s}_n(\mathscr{U}_n^N(s))ds\|_{\mathbf{TV}}$. We now require the following two lemmas to complete the proof of Proposition 3.1.

Lemma 6.2. *Fix $n \geq 1$. Let $\mu, \nu \in \mathscr{P}(\mathscr{S}_n)$ be two arbitrary purely atomic probability measures. Then*

$$\|\mathfrak{s}_n(\mu) - \mathfrak{s}_n(\nu)\|_{\boldsymbol{TV}} \leq \frac{1}{4}\|\mu - \nu\|_{\boldsymbol{TV}}. \tag{26}$$

Proof. By definition of smoothing of a probability measure, we have that $\mathfrak{s}_n(\mu), \mathfrak{s}_n(\mu)$ are both absolutely continuous with respect to the Lebesgue measure. Now by definition of total variation distance, we have

$$\begin{aligned}
\|\mathfrak{s}_n(\mu) - \mathfrak{s}_n(\nu)\|_{\mathbf{TV}} &= \frac{1}{2}\int_S \Big|\frac{d\mathfrak{s}_n(\mu)}{d\lambda}(x) - \frac{d\mathfrak{s}_n(\nu)}{d\lambda}(x)\Big|dx \\
&= \frac{1}{2}\int_S \Big|\sum_{k=0}^{2^n-1} 2^n\mu(\{\alpha_{n,k}\})\mathbb{I}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) - \sum_{k=0}^{2^n-1} 2^n\nu(\{\alpha_{n,k}\})\mathbb{I}_{[\alpha_{n,k},\alpha_{n,k+1})}(x)\Big|dx \\
&\leq \frac{1}{2}\sum_{k=0}^{2^n-1} |\mu(\{\alpha_{n,k}\}) - \nu(\{\alpha_{n,k}\})| \\
&= \frac{1}{4}\|\mu - \nu\|_{\mathbf{TV}}.
\end{aligned}$$

This completes the proof of Lemma 6.2. ■

Lemma 6.3. *Fix $\varepsilon > 0$. Then for every $\mu, \nu \in \mathscr{P}(\mathscr{S})$, we have*

$$\|\mathfrak{L}^n_{\mathit{vect}}(\mu) - \mathfrak{L}^n_{\mathit{vect}}(\nu)\|_{\mathbf{\mathit{TV}}} \leq 4\varepsilon \exp(2\|f_n\|_\infty) + 2\|f\|_\infty \exp(2\|f_n\|_\infty)\|\mu - \nu\|_{\mathbf{\mathit{TV}}}. \tag{27}$$

for sufficiently large n.

Proof. Fix $\mu, \nu \in \mathscr{P}(\mathscr{S})$. By definition of total-variation norm, we have that

$$\begin{aligned}
\|\mathfrak{L}^n_{\text{vect}}(\mu) - \mathfrak{L}^n_{\text{vect}}(\nu)\|_{\mathbf{TV}} &= \int_{\mathscr{S}} \Big| \frac{\exp(\varphi^n_x(\mu))}{\int_{\mathscr{S}} \exp(\varphi^n_y(\mu))dy} - \frac{\exp(\varphi^n_x(\nu))}{\int_{\mathscr{S}} \exp(\varphi^n_y(\nu))dy} \Big| dx \\
&= \int_{\mathscr{S}} \Big| \frac{\int_{\mathscr{S}} \exp(\varphi^n_x(\mu))\exp(\varphi^n_y(\nu))dy - \int_{\mathscr{S}} \exp(\varphi^n_x(\nu))\exp(\varphi^n_y(\mu))dy}{(\int_{\mathscr{S}} \exp(\varphi^n_y(\mu))dy)(\int_{\mathscr{S}} \exp(\varphi^n_y(\nu))dy)} \Big| dx \\
&\leq \int_{\mathscr{S}} \Big| \int_{\mathscr{S}} \exp(\varphi^n_x(\mu) + \varphi^n_y(\nu))dy - \int_{\mathscr{S}} \exp(\varphi^n_x(\mu) + \varphi^n_y(\nu))dy \\
&\qquad + \int_{\mathscr{S}} \exp(\varphi^n_x(\mu) + \varphi^n_y(\nu))dy - \int_{\mathscr{S}} \exp(\varphi^n_x(\nu) + \varphi^n_y(\mu))dy \Big| \\
&\leq \int_{\mathscr{S}} |\exp(\varphi^n_x(\mu))(\exp(\varphi^n_y(\nu)) - \exp(\varphi^n_y(\mu)))|dy \\
&\qquad \int_{\mathscr{S}} |\exp(\varphi^n_y(\mu))(\exp(\varphi^n_x(\mu)) - \exp(\varphi^n_x(\nu)))|dy \\
&\leq 2\exp(\|\varphi^n(\mu)\|_\infty + \|\varphi^n(\mu)\|_\infty \vee \|\varphi^n(\nu)\|_\infty)\|\varphi^n(\mu) - \varphi^n(\nu)\|_\infty \\
&\leq 2\exp(2\|f_n\|_\infty)\|\varphi^n(\mu) - \varphi^n(\nu)\|_\infty \\
&\leq 2\exp(2\|f_n\|_\infty)(\|\varphi^n(\mu) - \varphi(\mu)\|_\infty + \|\varphi(\mu) - \varphi(\nu)\|_\infty + \|\varphi^n(\nu) - \varphi(\nu)\|_\infty) \\
&\leq 2\exp(2\|f_n\|_\infty)(2\varepsilon + \|f\|_\infty\|\mu - \nu\|_{\mathbf{TV}}) \\
&= 4\varepsilon \exp(2\|f_n\|_\infty) + 2\|f\|_\infty \exp(2\|f_n\|_\infty)\|\mu - \nu\|_{\mathbf{TV}}.
\end{aligned}$$

As a result, we have that

$$\|\mathfrak{L}^n_{\text{vect}}(\mu) - \mathfrak{L}^n_{\text{vect}}(\nu)\|_{\mathbf{TV}} \leq 2\exp(2\|f_n\|_\infty)(2\varepsilon + (1 + \|f\|_\infty)\|\mu - \nu\|_{\mathbf{TV}}), \quad \text{for all } \mu, \nu \in \mathscr{P}(\mathscr{S}).$$

This concludes the proof of Lemma 6.3. ■

We now proceed with the proof of the proposition. Notice that

$$\begin{aligned}
\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)) - \mu^n(t)\|_{\mathbf{TV}} &\leq 8\varepsilon t \exp(2\|f_n\|_\infty) + 2(1 + \|f\|_\infty)\exp(2\|f_n\|_\infty)\delta T \\
&\qquad + 2(1 + \|f\|_\infty)\exp(2\|f_n\|_\infty)\int_0^t \|\mathfrak{s}(\widehat{\mathbb{X}}^N_n(s)) - \mu^n(s)\|_{\mathbf{TV}} ds + \Psi(T).
\end{aligned}$$

This implies that

$$\begin{aligned}
\sup_{0 \leq t \leq T} \|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)) - \mu^n(t)\|_{\mathbf{TV}} &\leq \Big[8\varepsilon T \exp(2\|f_n\|_\infty) + 2(1 + \|f\|_\infty)\exp(2\|f_n\|_\infty)\delta T \\
&\qquad + \Psi(T)\Big] + 2(1 + \|f\|_\infty)\exp(2\|f_n\|_\infty)\int_0^t \|\mathfrak{s}(\widehat{\mathbb{X}}^N_n(s)) - \mu^n(s)\|_{\mathbf{TV}} ds.
\end{aligned}$$

Therefore by Grönwall's inequality we have the following bound:

$$\sup_{0 \leq t \leq T} \|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)) - \mu^n(t)\|_{\mathbf{TV}} \leq \Big[8\varepsilon T \exp(2\|f_n\|_\infty) + 2(1 + \|f\|_\infty)\exp(2\|f_n\|_\infty)\delta T$$

$$+\Psi(T)\Big]\exp(2(1+\|f\|_\infty)\exp(2\|f_n\|_\infty)T). \tag{28}$$

Letting n,N to be sufficiently large enough such that

$$\varepsilon_n \vee \delta_N \le \frac{\varepsilon}{8T(1+\kappa)\exp(2\|f_n\|_\infty)}\exp(-2(1+\|f\|_\infty)\exp(2\|f_n\|_\infty)T), \tag{29}$$

in which case we have

$$\left\{\sup_{0\le t\le T}\|\mathfrak{s}(\widehat{\mathbb{X}}_n^N(t))-\mu^n(t)\|_{\mathbf{TV}}\ge\varepsilon\right\} \implies \left\{\Psi(T)\ge\varepsilon\Big(\frac{5+\|f\|_\infty}{4+4\|f\|_\infty}\Big)\exp(-2(1+\kappa)\exp(2\|f_n\|_\infty)T)\right\} \tag{30}$$

As a result, we have

$$\mathbb{P}_{\mu_0^n}\left(\sup_{0\le t\le T}\|\mathfrak{s}(\widehat{\mathbb{X}}_n^N(t))-\mu^n(t)\|_{\mathbf{TV}}\ge\varepsilon\right)\le\mathbb{P}_{\mu_0^n}\left(\Psi(T)\ge\varepsilon\Big(\frac{5+\|f\|_\infty}{4+4\|f\|_\infty}\Big)\exp(-2(1+\|f\|_\infty)\exp(2\|f_n\|_\infty)T)\right).$$

We now proceed to bound the probability on the right hand side of the above inequality. Let us define the following collection of functions:

$$\bar{\mathscr{L}}^\infty(\mathscr{S}):\left\{\bar g\in\mathscr{L}^\infty(\mathscr{S}):\exists\,(\beta_k)_{k=0}^{2^n-1}\subseteq\mathbb{R},\text{ such that }\bar g(x):=\sum_{k=0}^{2^n-1}\beta_k\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x),\text{ for all }x\in\mathscr{S}\right\}. \tag{31}$$

Now fix $\bar g\in\bar{\mathscr{L}}^\infty(S)$. For $k\ge1$, let $\mathbf{Z}_k(\bar g):=\exp\Big(\sum_{i=0}^k\langle\bar g,\delta\mathfrak{s}(\mathscr{U}_{n,i}^N)\rangle-\gamma k\delta^2\|\bar g\|_\infty^2\Big)$.

Lemma 6.4. *For every $\bar g\in\bar{\mathscr{L}}^\infty(S)$, the process $(\mathbf{Z}_k(\bar g))_{k\ge1}$ is a supermartingale with respect to the filtration $(\mathscr{F}_k)_{k\ge1}$.*

Proof. From definition of (U_k), we observe that

$$\begin{aligned}\|\mathfrak{s}(\mathscr{U}_k^N)\|_{\mathbf{TV}}&\le\frac{1}{\delta_N}\|\mathbb{X}_n^N((k+1)\delta_N)-\mathbb{X}_n^N(k\delta_N)\|_{\mathbf{TV}}+\|\mathfrak{L}^n_{\mathsf{vect}}(\mathbb{X}_n^N(k\delta_N))\|_{\mathbf{TV}}\\&\le\sqrt2+\|\mathfrak{L}^n_{\mathbf{vect}}\|_{\mathbf{TV}}=\sqrt\gamma\ \ (\text{say}).\end{aligned}$$

Now it follows along the lines of Benaïm and Weibull (2003), that $\mathbb{E}(\exp(\langle\bar g,\mathscr{U}_{n,k}^N\rangle))\le\exp(\frac{\gamma\|\bar g\|_\infty^2}{2})$, and thus completes the proof of Lemma 6.4. ■

Thus for any $\theta>0$, using Lemma 6.4, we have that

$$\begin{aligned}\mathbb{P}_{\mu_0^n}\Big(\max_{0\le k\le n}\Big\langle\bar g,\sum_{i=0}^{k-1}\delta_N\mathfrak{s}_n(\mathscr{U}_{n,i}^N)\Big\rangle\ge\theta\Big)&\le\mathbb{P}_{\mu_0^n}\Big(\max_{0\le k\le n}\mathbb{Z}_k(\bar g)\ge\exp(\theta-\frac{\gamma}{2}\|g\|_\infty^2n\delta_N^2)\Big)\\&\le\exp\Big(\frac{\gamma}{2}\|g\|_\infty^2n\delta_N^2-\theta\Big)\end{aligned}$$

For $k\ge1$, let $h_k:S\to\mathbb{R}$ be defined as $h_k(x):=\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x)$, for $x\in S$. It then follows that $(h_k)_{0\le k\le2^n-1}$ forms a basis of $\bar{\mathscr{L}}^\infty(S)$. Let $h=h_k$ or $h=-h_k$ for some k. Set $\theta=\frac{\varepsilon^2}{\gamma n\delta^2}$. Then there exists $\mathbf{h}=(h(\alpha_{n,0}),\dots,h(\alpha_{n,2^n-1}))\in\mathbb{R}^{2^n}$ and $\mathbf{U}_n^N\in\mathbb{R}^{2^n}$ such that $\Big\langle h,\sum_{i=0}^{k-1}\delta\mathfrak{s}_n(\mathscr{U}_{n,i}^N)\Big\rangle=\Big\langle\mathbf{h},\sum_{i=0}^{k-1}\delta\mathbf{U}_{n,i}^N\Big\rangle$. Now set

$g = \left(\frac{\theta}{\varepsilon}\right)h$ and define $\mathbf{g} := \left(\frac{\theta}{\varepsilon}\right)\mathbf{h}$. It then follows from Benaïm and Weibull (2003) that

$$\begin{aligned}\mathbb{P}_{\mu_0^n}\Big(\max_{0\le k\le m}\Big\langle h, \sum_{i=0}^{k-1}\delta_N \mathfrak{s}(U_{n,i}^N)\Big\rangle \ge \varepsilon\Big) &= \mathbb{P}_{\mu_0^n}\Big(\max_{0\le k\le m}\Big\langle \mathbf{h}, \sum_{i=0}^{k-1}\delta_N \mathbf{U}_{n,i}^N\Big\rangle \ge \varepsilon\Big)\\ &\le \mathbb{P}_{\mu_0^n}\Big(\max_{0\le k\le m}\Big\langle \mathbf{g}, \sum_{i=0}^{k-1}\delta_N \mathbf{U}_{n,i}^N\Big\rangle \ge \theta\Big)\\ &\le \exp\Big(\frac{-\varepsilon^2}{2m\gamma\delta_N^2}\Big).\end{aligned}$$

It now follows from Lemma 6.2 that

$$\mathbb{P}_{\mu_0^n}(\Psi(T)\ge\varepsilon) \le 2^{n+1}\exp\Big(\frac{-\varepsilon^2}{8\delta_N\gamma T}\Big).$$

Define $C = \Big(\frac{5+\|f\|_\infty}{4+4\|f\|_\infty}\Big)\exp(-2(1+\|f\|_\infty)\exp(2(1+\|f\|_\infty)T)$. Note that by definition of C, it is independent of μ_0^n. As a result, we then arrive the inequality:

$$\begin{aligned}\mathbb{P}_{\mu_0^n}\Big(\sup_{0\le t\le T}\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \mu^n(t)\|_{\mathbf{TV}} \ge \varepsilon\Big) &\le \mathbb{P}_{\mu_0^n}\Big(\Psi(T)\ge C\varepsilon\Big)\\ &\le 2^{n+1}\exp\left(\frac{-C^2\varepsilon^2}{8\delta_N\gamma T}\right).\end{aligned}$$

Let $\bar{C} := C^2/8\gamma T$. As $\delta_N = 1/N$, we have that

$$\mathbb{P}_{\mu_0^n}\Big(\sup_{0\le t\le T}\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \mu^n(t)\|_{\mathbf{TV}} \ge \varepsilon\Big) \le 2^{n+1}\exp(-N\bar{C}\varepsilon^2).$$

This concludes the proof of Proposition 3.1.

6.3 Proof of Theorem 3.1

The proof Theorem 3.1 follows from Propositions 2.1 and 3.1. To make the proof mathematically rigorous, consider the fixed probability space $(\Omega,\mathscr{F},\mathbb{P})$ on which the stochastic evolutionary process $(\widehat{\mathbb{X}}_n^N(t))_{t\ge0}$ is defined. For $n\ge1$, let $\widehat{\mathbb{X}}_n : \Omega \to \mathscr{P}(\mathscr{S})$ be the degenerate stochastic process such that for al $\omega\in\Omega$, we have $\widehat{\mathbb{X}}_n(\omega)(t) := \mu^n(t)$, for all $t\ge0$. Also define the process $\widehat{\mathbb{X}} : \Omega\to\mathscr{P}(\mathscr{S})$ such that for all $\omega\in\Omega$, $\widehat{\mathbb{X}}(\omega)(t) := \mu(t)$, for all $t\ge0$. For $\varepsilon>0$, we have that

$$\begin{aligned}\left\{\sup_{0\le t\le T}\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \widehat{\mathbb{X}}(t)\|_{\mathbf{TV}}\ge\varepsilon\right\} &\subseteq \left\{\sup_{0\le t\le T}\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \widehat{\mathbb{X}}_n(t)\|_{\mathbf{TV}}\ge\varepsilon/2\right\}\\ &\quad\cup\left\{\sup_{0\le t\le T}\|\widehat{\mathbb{X}}_n(t) - \widehat{\mathbb{X}}(t)\|_{\mathbf{TV}}\ge\varepsilon/2\right\}.\end{aligned}$$

As a result we have that

$$\mathbb{P}_{\mu_{\mathbf{p}^n}^N}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon\right) \le \mathbb{P}_{\mu_{\mathbf{p}^n}^N}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)) - \widehat{\mathbb{X}}_n(t)\right\|_{\mathbf{TV}}\ge\varepsilon/2\right)$$

$$+\mathbb{P}_{\mu^N_{\mathbf{p}^n}}\left(\sup_{0\le t\le T}\left\|\widehat{\mathbb{X}}_n(t)-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon/2\right).$$

By assumption of the theorem, we have that the initial conditions $\mu^N_{\mathbf{p}^n}=\mu_0$, for all $N,n\ge 1$, implies that

$$\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon\right)\le\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))-\widehat{\mathbb{X}}_n(t)\right\|_{\mathbf{TV}}\ge\varepsilon/2\right)$$
$$+\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\widehat{\mathbb{X}}_n(t)-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon/2\right).$$

Fix $n\ge 1$. Then by taking limits as $N\to\infty$ we have by Proposition 3.1 that

$$\lim_{N\to\infty}\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon\right)\le\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\widehat{\mathbb{X}}_n(t)-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon/2\right).\tag{32}$$

Now, allowing $n\to\infty$, in (32) we have by an application of Proposition 2.1 that

$$\lim_{n\to\infty}\lim_{N\to\infty}\mathbb{P}_{\mu_0}\left(\sup_{0\le t\le T}\left\|\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))-\widehat{\mathbb{X}}(t)\right\|_{\mathbf{TV}}\ge\varepsilon\right)=0.$$

Finally, using the fact that $\widehat{\mathbb{X}}(t)=\mu(t)$, for all $t\ge 0$, $\mathbb{P}_{\mu_0}$-a.s concludes the proof of Theorem 3.1.

6.4 Proof of Proposition 4.1

For this proof, we topologize the space of probability measures $\mathscr{P}(\mathscr{S})$ with the $\mathbf{BL}^*$ which we define below: let $g:\mathscr{S}\to\mathbb{R}$ be a bounded Lipschitz map. Then the bounded Lipschitz norm is defined as

$$\|g\|_{\mathbf{BL}}:=\sup_{x\in\mathscr{S}}|g(x)|+\sup_{x\neq y}\frac{|g(x)-g(y)|}{|x-y|},\quad\text{for all } g\in\mathsf{Lip}(\mathscr{S}),\tag{33}$$

where $\mathsf{Lip}(\mathscr{S})$ denotes the collection of all bounded Lipschitz functions on $\mathscr{S}$. Now let

$$\mathsf{Bounded\text{-}Lip}(\mathscr{S}):=\{g:\mathscr{S}\to\mathbb{R}\mid g\text{ is bounded and Lipschitz with }\|g\|_{\mathbf{BL}}\le 1\}$$

be the collection of bounded Lipschitz continuous functions with $\mathbf{BL}$-norm bounded by 1. Denote by $\mathscr{M}(\mathscr{S})$, the space of all signed measures on $\mathscr{S}$. Then the dual $\mathbf{BL}^*$-norm on $\mathscr{M}(\mathscr{S})$ is defined as

$$\|\mu-\nu\|_{\mathbf{BL}^*}:=\sup_{g\in\mathsf{Bounded\text{-}Lip}(\mathscr{S})}\left(\int_{\mathscr{S}}f\,d\mu-\int_{\mathscr{S}}f\,d\nu\right),\quad\text{for all }\mu,\nu\in\mathscr{M}(\mathscr{S}).\tag{34}$$

It is well-known (see Billingsley (2013), Perkins and Leslie (2014)) that the weak topology on $\mathscr{P}(\mathscr{S})$ coincides with the topology generated by the $\mathbf{BL}^*$-norm on $\mathscr{P}(\mathscr{S})$. Therefore, weak convergence is equivalent to convergence with respect to the $\mathbf{BL}^*$-norm.

We now proceed with the proof of the proposition. Suppose that φ is supermodular and let $n\ge 1$. We show that the level-n discretization $\mathscr{D}^n_\varphi$ of φ as defined in Definition 3.1 is a supermodular game. Fix $\mathbf{x},\mathbf{y}\in\Delta(\mathscr{S}_n)$ with $\mathbf{x}\succeq\mathbf{y}$. It then follows that $\mu_{\mathbf{x}}:=\sum_{0\le j\le 2^n}\mathbf{x}_j\delta_{\alpha_{n,j}}\succeq\sum_{0\le j\le 2^n}\mathbf{y}_j\delta_{\alpha_{n,j}}=:\mu_{\mathbf{y}}$. Now fix $i<j$.

It then follows from the definition of $\mathscr{D}_\varphi^n$ that,

$$\begin{aligned}
\mathscr{D}_\varphi^n[\mathbf{x}](\alpha_{n,j}) - \mathscr{D}_\varphi^n[\mathbf{x}](\alpha_{n,i}) &= \langle f_n(\alpha_{n,j},\cdot),\mathbf{x}\rangle - \langle f_n(\alpha_{n,i},\cdot),\mathbf{x}\rangle \\
&= \sum_{0\le k\le 2^n} f_n(\alpha_{n,j},\alpha_{n,k})\mathbf{x}_k - \sum_{0\le k\le 2^n} f_n(\alpha_{n,i},\alpha_{n,k})\mathbf{x}_k \\
&= \int_S f_n(\alpha_{n,j},t)\mu_{\mathbf{x}}(dt) - \int_S f_n(\alpha_{n,i},t)\mu_{\mathbf{x}}(dt) \\
&= \varphi_{\alpha_{n,j}}(\mu_{\mathbf{x}}) - \varphi_{\alpha_{n,i}}(\mu_{\mathbf{x}}) \\
&\ge \varphi_{\alpha_{n,j}}(\mu_{\mathbf{y}}) - \varphi_{\alpha_{n,i}}(\mu_{\mathbf{y}}) \\
&= \int_S f_n(\alpha_{n,j},t)\mu_{\mathbf{y}}(dt) - \int_S f_n(\alpha_{n,i},t)\mu_{\mathbf{y}}(dt) \\
&= \sum_{0\le k\le 2^n} f_n(\alpha_{n,j},\alpha_{n,k})\mathbf{y}_k - \sum_{0\le k\le 2^n} f_n(\alpha_{n,i},\alpha_{n,k})\mathbf{y}_k \\
&= \langle f_n(\alpha_{n,j},\cdot),\mathbf{y}\rangle - \langle f_n(\alpha_{n,i},\cdot),\mathbf{y}\rangle \\
&= \mathscr{D}_\varphi^n[\mathbf{y}](\alpha_{n,j}) - \mathscr{D}_\varphi^n[\mathbf{y}](\alpha_{n,i}).
\end{aligned}$$

This completes the proof that $\mathscr{D}_\varphi^n$ is a supermodular game. Since $n \ge 1$ is arbitrary, this concludes the proof of "only-if part" of Proposition 4.1.

We now proceed to prove the "if-part" of the theorem. Now suppose that $\mathscr{D}_\varphi^n$ is a supermodular game for every $n \ge 1$. We need to show that φ is a supermodular game. To this end, first $x > y$. Let $\mathscr{S}_\infty := \cup_{n\ge1}\mathscr{S}_n$. Since $\mathscr{S}_\infty$ is dense in $\mathscr{S}$, there exists two subsequences $\{n_k\}_{k\ge1}, \{m_k\}_{k\ge1} \subseteq \mathbb{N}$ and sequences $\{\alpha^x_{n_k,j_{n_k}}\} \subseteq \mathscr{S}_{n_k}$ and $\{\alpha^y_{m_k,j_{m_k}}\} \subseteq \mathscr{S}_{m_k}$ such that $\alpha^x_{n_k,j_{n_k}} \downarrow x$ and $\alpha^y_{m_k,j_{m_k}} \uparrow y$ as $k \to \infty$. Again, since $\mathscr{S}_\infty$ is dense in $\mathscr{S}$, it follows (see Billingsley (2013)) that that space of atomic probability measures $\mathscr{P}(\mathscr{S}_\infty)$ on $\mathscr{S}_\infty$ is dense in $\mathscr{P}(\mathscr{S})$ relative to the topology of weak convergence. Since weak topology in equivalent to the $\mathbf{BL}^*$-norm on $\mathscr{P}(\mathscr{S})$, we therefore have that for every $\mu \in \mathscr{P}(\mathscr{S})$, there exists $(\mu_n)_{n\ge1} \subseteq \mathscr{P}(\mathscr{S}_\infty)$ such that $\mu_n \xrightarrow{\mathbf{BL}^*} \mu$. By definition since φ is a weakly continuous population game, we have $\varphi_x(\mu_n)$ converges to $\varphi_x(\mu)$ for all $x \in \mathscr{S}$. As a result we have for all $x \in \mathscr{S}$ that,

$$\begin{aligned}
\varphi_x(\mu) &= \lim_{n\to\infty} \varphi_x(\mu_n) \\
&= \lim_{n\to\infty} \int_S f(x,t)\mu_n(dt).
\end{aligned}$$

By the definition in (6), the sequence $(f_n)_{n\ge1}$ converges uniformly to f. This, in conjuction with Bounded Convergence Theorem (Billingsley (2013)) implies that

$$\begin{aligned}
\varphi_x(\mu) &= \lim_{n\to\infty} \int_S f(x,t)\mu_n(dt). \\
&= \lim_{n\to\infty}\lim_{k\to\infty} \int_S f_k(\alpha^x_{n_k,j_{n_k}},t)\mu_n(dt) \\
&= \lim_{n\to\infty}\lim_{k\to\infty} \mathscr{D}_\varphi^n[\mathbf{p}_{\mu_n}](\alpha^x_{n_k,j_{n_k}}).
\end{aligned}$$

We are now ready to conclude the proof of the theorem. Fix $\mu, \nu \in \mathscr{P}(\mathscr{S})$ with $\mu \succeq \nu$. Then there exists two sequences of atomic probability measures $(\mu_n)_{n\ge1} \subseteq \mathscr{P}(\mathscr{S}_\infty)$ and $(\nu_n)_{n\ge1} \subseteq \mathscr{P}(\mathscr{S}_\infty)$ such that $\mu_n \succeq \nu_n$ for all $n \ge 1$, such that $\mu_n \xrightarrow{\mathbf{BL}^*} \mu$ and $\nu_n \xrightarrow{\mathbf{BL}^*} \nu$. By hypothesis of the theorem, the level-n discretization $\mathscr{D}_\varphi^n$ is a supermodular game. Also by construction, we have that $\alpha^x_{n_k,j_{n_k}} > \alpha^y_{m_k,j_{m_k}}$ for every

$k \geq 1$. Therefore we have

$$\begin{aligned}
\varphi_x(\mu) - \varphi_y(\mu) &= \lim_{n\to\infty}\lim_{k\to\infty} \mathscr{D}^n_\varphi[\mathbf{p}_{\mu_n}](\alpha^x_{n_k,j_{n_k}}) - \lim_{n\to\infty}\lim_{k\to\infty} \mathscr{D}^n_\varphi[\mathbf{p}_{\mu_n}](\alpha^y_{m_k,j_{m_k}}) \\
&= \lim_{n\to\infty}\lim_{k\to\infty} \left[\mathscr{D}^n_\varphi[\mathbf{p}_{\mu_n}](\alpha^x_{n_k,j_{n_k}}) - \mathscr{D}^n_\varphi[\mathbf{p}_{\mu_n}](\alpha^y_{m_k,j_{m_k}})\right] \\
&\geq \lim_{n\to\infty}\lim_{k\to\infty} \left[\mathscr{D}^n_\varphi[\mathbf{p}_{\nu_n}](\alpha^x_{n_k,j_{n_k}}) - \mathscr{D}^n_\varphi[\mathbf{p}_{\nu_n}](\alpha^y_{m_k,j_{m_k}})\right] \\
&= \lim_{n\to\infty}\lim_{k\to\infty} \mathscr{D}^n_\varphi[\mathbf{p}_{\nu_n}](\alpha^x_{n_k,j_{n_k}}) - \lim_{n\to\infty}\lim_{k\to\infty} \mathscr{D}^n_\varphi[\mathbf{p}_{\nu_n}](\alpha^y_{m_k,j_{m_k}}) \\
&= \varphi_x(\nu) - \varphi_y(\nu).
\end{aligned}$$

Since $\mu \succeq \nu$ is arbitrary, this proves that φ is a supermodular game and hence concludes the proof of Proposition 4.1.

6.5 Proof of Proposition 4.2

Recall that the two-player game f which induces φ in Example 2.1 is defined as $f(x,y) = m(x)y - c(x)$, for all $x, y \in \mathscr{S}$, where m is an increasing function of x and c is an arbitrary function of x. Using (6), we define the step–wise approximation of f as

$$f_n(x,y) = \begin{cases} m(\alpha_{n,j})\alpha_{n,k} - c(\alpha_{n,j}), \text{ if } (x,y) \in [\alpha_{n,j}, \alpha_{n,j+1}) \times [\alpha_{n,k}, \alpha_{n.k+1}) \text{ and } j,k = 0,\dots,2^n-2 \\ m(\alpha_{n,2^n-1})\alpha_{n,2^n-1} - c(\alpha_{n,2^n-1}) \text{ if } (x,y) \in [\alpha_{n,2^n-1}, 1] \times [\alpha_{n,2^n-1}, 1]. \end{cases} \tag{35}$$

This implies that for every $n \geq 1$, we have from (13),

$$\mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j}) = \sum_{0\leq k\leq 2^n-1} f_n(\alpha_{n,j}, \alpha_{n,k})\mathbf{p}^n_k, \quad \text{for all } j = 0,1,\dots,2^n-1.$$

We now proceed to show that $\mathscr{D}^n_\varphi$ is irreducible. Fix $\mathbf{p}^n \in \Delta(S_n)$ and $K \subset \widehat{S}_n$ arbitrary as in Definition 4.3. Now let $j \in K$ and $i \in \widehat{S}_n \setminus K$. Since m is an increasing function of x, we have

$$\begin{aligned}
\frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j+1})}{\partial \mathbf{p}^n_{i+1}} - \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j+1})}{\partial \mathbf{p}^n_i} &= f_n(\alpha_{n,j+1}, \alpha_{n,i+1}) - f_n(\alpha_{n,j+1}, \alpha_{n,i}) \\
&= m(\alpha_{n,j+1})(\alpha_{n,i+1} - \alpha_{n,i}) \\
&> m(\alpha_{n,j})(\alpha_{n,i+1} - \alpha_{n,i}) \\
&= f_n(\alpha_{n,j}, \alpha_{n,i+1}) - f_n(\alpha_{n,j}, \alpha_{n,i}) \\
&= \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j})}{\partial \mathbf{p}^n_{i+1}} - \frac{\partial \mathscr{D}^n_\varphi[\mathbf{p}^n](\alpha_{n,j})}{\partial \mathbf{p}^n_i}
\end{aligned}$$

This proves that $\mathscr{D}^n_\varphi$ is irreducible and hence concludes the proof of Proposition 4.2.

6.6 Proof of Lemma 4.1

To prove the lemma we shall first exploit the properties of $\mathbf{BL}^*$-norm, or equivalently weak topology on the space $\mathscr{P}(\mathscr{S})$. Observe that the sequence of level-n step logit equilibria $\mathfrak{s}_n(\mu_{\mathbf{p}^*_n}) \in \mathscr{P}(\mathscr{S})$ for every $n \geq 1$. Since the weak topology, and hence the $\mathbf{BL}^*$-norm renders the space $\mathscr{P}(\mathscr{S})$ compact, there exists a subsequence $(n_k)_{k\geq 1}$ such that $\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})$ converges in the $\mathbf{BL}^*$-norm to some $\mu^* \in \mathscr{P}(\mathscr{S})$. We now show that μ^* is a logit equilibrium of the original game φ. To this end, we need to show that

$$\mu^*(A) = \mathfrak{L}_\eta[\mu^*](A), \quad \text{for all } A \in \mathscr{B}(\mathscr{S}).$$

Note that since $\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})$ is a level-n_k step logit equilibrium of the game φ^{n_k} for every $k \geq 1$, we have that

$$\mathfrak{s}_{n_k}[\mu_{\mathbf{p}^*_{n_k}}](A) = \mathfrak{L}^{n_k}_\eta[\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})](A), \quad \text{for all } A \in \mathscr{B}(\mathscr{S}) \text{ such that } \mu^*(\partial A) = 0. \tag{36}$$

Now, we show that $\varphi^{n_k}(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})) \to \varphi(\mu^*)$ as $k \to \infty$. Fix $x \in \mathscr{S}$. Since f is Lipschitz continuous, we have

$$\begin{aligned}
|\varphi^{n_k}_x(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})) - \varphi_x(\mu^*)| &= \left| \int_{\mathscr{S}} f_{n_k}(x,y)\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})(dy) - \int_{\mathscr{S}} f(x,y)\mu^*(dy) \right| \\
&\leq \int_{\mathscr{S}} |f_{n_k}(x,y) - f(x,y)|\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})(dy) + \left| \int_{\mathscr{S}} f(x,y)(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}) - \mu^*)(dy) \right| \\
&\leq 2\|f_{n_k} - f\|_\infty + \|f\|_\infty \|\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}) - \mu^*\|_{\mathbf{BL}^*}.
\end{aligned}$$

As a result, using the fact that $f_n \to f$ uniformly, and that $\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})$ converges in the $\mathbf{BL}^*$-norm to μ^*, we have

$$\begin{aligned}
\|\varphi^{n_k}(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})) - \varphi(\mu^*)\|_\infty &= \sup_{x \in \mathscr{S}} |\varphi^{n_k}_x(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})) - \varphi_x(\mu^*)| \\
&\leq 2\|f_{n_k} - f\|_\infty + \|f\|_\infty \|\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}) - \mu^*\|_{\mathbf{BL}^*} \\
&\to 0, \quad \text{as } k \to \infty.
\end{aligned}$$

Also since

$$\exp(\varphi^{n_k}_x(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))) \leq \exp(2(1 + \|f\|_\infty)), \quad \text{for all } x \in \mathscr{S},$$

we have using dominated convergence theorem that $\int_{\mathscr{S}} \exp(\varphi^{n_k}_y(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})))dy \to \int_{\mathscr{S}} \exp(\varphi_y(\mu_*))dy$. Now, in order to prove that $\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))$ converges in the variational norm to $\mathfrak{L}_\eta(\mu_*)$, it is enough to show that

$$\frac{d\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))}{d\lambda} \to \frac{d\mathfrak{L}_\eta(\mu_*)}{d\lambda} \quad \text{in } \mathscr{L}^1(\mathscr{S}) \text{ as } k \to \infty.$$[17]

Note that the density function of $\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))$ satisfies

$$\frac{d\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))}{d\lambda}(x) \leq \exp(2(1 + \|f\|_\infty)), \quad \text{for all } x \in \mathscr{S}.$$

We also have that for all $x \in \mathscr{S}$,

$$\frac{d\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))}{d\lambda}(x) \to \frac{d\mathfrak{L}_\eta(\mu_*)}{d\lambda}(x) \quad \text{as } k \to \infty.$$

Again, by an application of Dominated Convergence Theorem, we conclude that $\mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}))$ converges in the variational norm to $\mathfrak{L}_\eta(\mu_*)$. Therefore, for every $A \in \mathscr{B}(\mathscr{S})$, we have that $\mathfrak{L}^{n_k}_\eta[\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})](A) \to \mathfrak{L}_\eta[\mu^*](A)$ as $k \to \infty$. Using this observation in (36), we have that μ^* is a fixed point of $\mathfrak{L}_\eta$ and hence a

[17] The space $\mathscr{L}^1(\mathscr{S})$ is defined as the collection of all measurable functions $g : \mathscr{S} \to \mathbb{R}$ such that $\int_{\mathscr{S}} |f(x)|dx < \infty$. We say that $g_n \to g$ in $\mathscr{L}^1(\mathscr{S})$ if $\int_{\mathscr{S}} |g_n(x) - g(x)|dx \to 0$ as $n \to \infty$.

logit equilibrium of φ. In fact, we proved that that $\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}}) = \mathfrak{L}^{n_k}_\eta(\mathfrak{s}_{n_k}(\mu_{\mathbf{p}^*_{n_k}})) \to \mathfrak{L}_\eta(\mu^*) = \mu^*$ as $k \to \infty$ in with respect to the total variation norm. This concludes the proof of Lemma 4.1.

6.7 Proof of Theorem 4.1

The proof of Theorem 4.1 is subdivided into several steps.

Step 1: In this step, we show that the discretized version of the stochastic evolutionary process converges in the medium run to an element of $\mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$. Fix $\varepsilon > 0$. Pick $N, n \geq 1$ and consider the interpolated stochastic evolutionary process $(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$. We observe that though $(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$ is a stochastic process in taking values in $\mathscr{P}(\mathscr{S})$, however in principle, $(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$ takes values in the space of atomic measures on $\mathscr{S}$ with support S_n. Note that any probability measure $\mu \in \mathscr{P}(S_n)$ can be expressed as $\mu = \sum_{0\leq k\leq 2^n-1} \mathbf{p}^n_j \delta_{\alpha_{n,j}}$, for some $\mathbf{p}^n \in \Delta(S_n)$ and vice-versa. We therefore define the bijective mapping $\mathfrak{d}_n : \mathscr{P}(S_n) \to \Delta(S_n)$ such that $\mathfrak{d}_n(\sum_{0\leq k\leq 2^n-1} \mathbf{p}^n_j \delta_{\alpha_{n,j}}) = \mathbf{p}^n$ for all $\mathbf{p}^n \in \Delta(S_n)$. It then follows that $(\mathfrak{d}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ is the interpolated stochastic evolutionary process on $\Delta(S_n)$. Let $\mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$ denote the collection of discretized logit equilibria corresponding to the level-n discretized game $\mathscr{D}^n_\varphi$. More precisely, $\mathrm{LE}_\eta(\mathscr{D}^n_\varphi) := \{\mathbf{p}^{n,*}_\eta : L^n_\eta[\mathbf{p}^{n,*}_\eta] = \mathbf{p}^{n,*}_\eta\}$. By assumption of the theorem, we have that the original population game φ is supermodular. Therefore by Proposition 4.1, we have that the level-n discretization $\mathscr{D}^n_\varphi$ is supermodular. Also by hypothesis of the theorem, we have that $\mathscr{D}^n_\varphi$ is irreducible. Thus, by Hofbauer and Sandholm (2007), the process $(\mathfrak{d}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0} \subseteq \Delta(S_n)$ converges in the medium run to an element of $\mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$ as $N \to \infty$. Let us denote by $\mathbf{p}^*_n$ the element in $\mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$ to which $(\mathfrak{d}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ converges in the medium run as $N \to \infty$.[18] That is, there exists $T_\varepsilon > 0$ such that for all $T \geq T_\varepsilon$, such that

$$\lim_{N\to\infty} \mathbb{P}_{\mu_0}\Big(\mathfrak{d}_n(\widehat{\mathbb{X}}^N_n(t)) \in \mathscr{O}^\varepsilon(\mathbf{p}^*_n) \text{ for all } T_\varepsilon \leq t \leq T\Big) = 1.$$

We now proceed to show that the smoothed version of the stochastic evolutionary process $(\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ converges in the medium run to an element of $\mathrm{LE}_\eta(\varphi^n)$ as $N \to \infty$. In particular, we show that $(\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ converges in the medium run to $\mathfrak{s}_n(\mu_{\mathbf{p}^*_n})$. The fact that the processes $(\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ and $(\mathfrak{d}_n(\widehat{\mathbb{X}}^N_n(t)))_{t\geq 0}$ are in one-to-one correspondence with each other implies that $\mathfrak{s}_n(\widehat{\mathbb{X}}^N_n(t))_{t\geq 0}$ converges in the medium run to an element of $\mathfrak{s}_n(\mathrm{LE}_\eta(\mathscr{D}^n_\varphi))$.

Step 2: As a next step in the proof, we proceed to show that $\mathfrak{s}_n(\mathrm{LE}_\eta(\mathscr{D}^n_\varphi))$ coincides with $\mathrm{LE}_\eta(\varphi^n)$, thereby proving that $\mathfrak{s}_n(\mu_{\mathbf{p}^*_n}) \in \mathrm{LE}_\eta(\varphi^n)$. We prove this in the following lemma.

Lemma 6.5. *For $n \geq 1$, let $LE_\eta(\mathscr{D}^n_\varphi)$ and $LE_\eta(\varphi^n)$ be the collection of level-n discretized and step logit equilibria corresponding to $\mathscr{D}^n_\varphi$ and φ^n respectively. Then $\boldsymbol{p}^{n,*}_\eta \in LE_\eta(\mathscr{D}^n_\varphi)$ if and only if $\mathfrak{s}_n(\mu_{\boldsymbol{p}^{n,*}_\eta}) \in LE_\eta(\varphi^n)$.*

Proof. Again, as before, we shall use the notation $\mathbf{p}^*$ to denote a discretized level-n logit equilibrium to simplify the presentation of the proof. First observe that for any $\mathbf{p}^n \in \Delta(S_n)$, we have by definition of $\mathfrak{s}_n$ that

$$\mu_{\mathbf{p}^n}[\alpha_{n,k}, \alpha_{n,k+1}) = \mathfrak{s}_n(\mu_{\mathbf{p}^n})[\alpha_{n,k}, \alpha_{n,k+1}) = \mathbf{p}^n_k, \quad \text{for all } k = 0, 1, \ldots, 2^n - 1.$$

Suppose that $\mathbf{p}^* \in \mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$. Then we have $\mathbf{p}^* = L^n_\eta[\mathbf{p}^*]$. This, in particular implies that

$$\mu_{\mathbf{p}^*} = \sum_{0\leq k\leq 2^n-1} L^n_{\eta,k}[\mathbf{p}^*]\delta_{\alpha_{n,k}}.$$

[18] To keep the notations simple, we remove the dependence of η in $\mathbf{p}^*_n$.

Thus by definition of $\mathfrak{s}_n$ (Definition 3.2), above equality implies that for all $x \in \mathscr{S}$,

$$
\begin{aligned}
\frac{\mathfrak{s}_n(\mu_{\mathbf{p}^*})}{d\lambda}(x) &= \sum_{0\le k\le 2^n-1} 2^n L^n_{\eta,k}[\mathbf{p}^*]\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} 2^n \frac{\exp(\eta^{-1}\langle f_n(\alpha_{n,k},\cdot),\mathbf{p}^*\rangle)}{\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\langle f_n(\alpha_{n,j},\cdot),\mathbf{p}^*\rangle)}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\sum_{0\le j\le 2^n-1} f_n(\alpha_{n,k},\alpha_{n,j})\mathbf{p}^*_j)}{2^{-n}\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\sum_{0\le i\le 2^n-1} f_n(\alpha_{n,j},\alpha_{n,i})\mathbf{p}^*_i))}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\sum_{0\le j\le 2^n-1} f_n(\alpha_{n,k},\alpha_{n,j})\mu_{\mathbf{p}^*}[\alpha_{n,j},\alpha_{n,j+1}))}{2^{-n}\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\sum_{0\le i\le 2^n-1} f_n(\alpha_{n,j},\alpha_{n,i})\mu_{\mathbf{p}^*}[\alpha_{n,i},\alpha_{n,i+1}))}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\int_{\mathscr{S}} f_n(\alpha_{n,k},y)\mu_{\mathbf{p}^*}(dy))}{2^{-n}\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\int_{\mathscr{S}} f_n(\alpha_{n,j},y)\mu_{\mathbf{p}^*}(dy))}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\varphi^n_{\alpha_{n,k}}(\mu_{\mathbf{p}^*}))}{2^{-n}\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\varphi^n_{\alpha_{n,j}}(\mu_{\mathbf{p}^*}))}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\varphi^n_{\alpha_{n,k}}(\mu_{\mathbf{p}^*}))}{\sum_{j=0}^{2^n-1}\exp(\eta^{-1}\varphi^n_{\alpha_{n,j}}(\mu_{\mathbf{p}^*}))\lambda([\alpha_{n,j},\alpha_{n,j+1})}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x) \\
&= \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\varphi^n_{\alpha_{n,k}}(\mu_{\mathbf{p}^*}))}{\int_{\mathscr{S}}\exp(\eta^{-1}\varphi^n_{y}(\mu_{\mathbf{p}^*}))\lambda(dy)}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x). \qquad (37)
\end{aligned}
$$

The proof of Lemma 6.5 concludes if we are able to show that $\varphi^n_x(\mathfrak{s}_n(\mu_{\mathbf{p}^*})) = \varphi^n_x(\mu_{\mathbf{p}^*})$ for all $x \in \mathscr{S}$. In fact, a more general equality holds: for all $x \in \mathscr{S}$, we have

$$\varphi^n_x(\mathfrak{s}_n(\mu_{\mathbf{p}^n})) = \varphi^n_x(\mu_{\mathbf{p}^n}), \quad \text{for all } \mathbf{p}^n \in \Delta(S_n).$$

To this end, fix $x \in \mathscr{S}$. Then, using the definition of φ^n, we have that

$$
\begin{aligned}
\varphi^n_x(\mathfrak{s}_n(\mu_{\mathbf{p}^n})) &= \int_{\mathscr{S}} f_n(x,y)\mathfrak{s}_n(\mu_{\mathbf{p}^n})(dy) \\
&= \sum_{0\le k\le 2^n-1} f_n(x,\alpha_{n,k})\mathfrak{s}_n(\mu_{\mathbf{p}^n})[\alpha_{n,k},\alpha_{n,k+1}) \\
&= \sum_{0\le k\le 2^n-1} f_n(x,\alpha_{n,k})\mathbf{p}^n_k \\
&= \int_{\mathscr{S}} f_n(x,y)\mu_{\mathbf{p}^n}(dy) \\
&= \varphi^n_x(\mu_{\mathbf{p}^n}).
\end{aligned}
$$

Plugging in the above observation in (37), we have the following equality:

$$\frac{\mathfrak{s}_n(\mu_{\mathbf{p}^*})}{d\lambda}(x) = \sum_{0\le k\le 2^n-1} \frac{\exp(\eta^{-1}\varphi^n_{\alpha_{n,k}}(\mathfrak{s}_n(\mu_{\mathbf{p}^*})))}{\int_{\mathscr{S}}\exp(\eta^{-1}\varphi^n_{y}(\mathfrak{s}_n(\mu_{\mathbf{p}^*})))\lambda(dy)}\mathbb{1}_{[\alpha_{n,k},\alpha_{n,k+1})}(x), \quad \text{for all } x \in \mathscr{S}.$$

Thus, using the above equality, we have by definition of level-n step logit equilibrium that $\mathfrak{s}_n(\mu_{\mathbf{p}^*}) = \mathfrak{L}^n_\eta(\mathfrak{s}_n(\mu_{\mathbf{p}^*}))$. As a result, we have that $\mathfrak{s}_n(\mu_{\mathbf{p}^*}) \in \mathrm{LE}_\eta(\varphi^n)$. Similarly, suppose $\mathbf{p}^*$ be such that $\mathfrak{s}_n(\mu_{\mathbf{p}^*}) \in \mathrm{LE}_\eta(\varphi^n)$. It then follows using similar arguments that $\mathbf{p}^* \in \mathrm{LE}_\eta(\mathscr{D}^n_\varphi)$. This concludes the proof of Lemma 6.5. ■

Therefore by applying Lemma 6.5 along with Hofbauer and Sandholm (2007), we then have that for every $\varepsilon > 0$, interpolated smoothed version $\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t))_{t\geq 0}$ converges in the medium run to $\mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))$.

Step 3: In this step, we use the deterministic approximation result (Proposition 3.1) to obtain that for sufficiently large $T \geq T_\varepsilon$, the solution to the level-n step logit dynamic $(\mu^n(t))_{t\geq 0}$ is close to an element of $\mathrm{LE}_\eta(\varphi^n)$ in the medium run. We require the following lemma in this regard.

Lemma 6.6. *For every $\varepsilon > 0$, there exits $T_\varepsilon > 0$ such that for all $T \geq T_\varepsilon$, we have $\mu^n(t) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{p_n^*}))$ for all $T_\varepsilon \leq t \leq T$.*

Proof. Consider the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ on which the interpolated stochastic evolutionary process $(\widehat{\mathbb{X}}_n^N(t))_{t\geq 0}$ is defined. Let $(\widehat{\mathbb{X}}_n(t))_{t\geq 0}$ be a process such that $\widehat{\mathbb{X}}_n(t) : \Omega \to \mathscr{P}(\mathscr{S})$ is defined as $\widehat{\mathbb{X}}_n(t,\omega) = \mu^n(t)$ for all $\omega \in \Omega$. The fact that $(\widehat{\mathbb{X}}_n(t))_{t\geq 0}$ is a deterministic process implies that it is measurable. To prove the lemma, we need to show that

$$\mathbb{P}_{\mu_0}\Big(\widehat{\mathbb{X}}_n(t) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*})) \text{ for all } T_\varepsilon \leq t \leq T\Big) = 1.$$

To this end, we first show that the solution trajectory $\mu^n(t)$ is continuous in t. Note that by definition of φ^n, we have that for all $\mu, \nu \in \mathscr{P}(\mathscr{S})$ and all $x \in \mathscr{S}$,

$$\begin{aligned}|\varphi_x^n(\mu) - \varphi_n^n(\nu)| &= \Big|\int_{\mathscr{S}} f_n(x,y)\mu(dy) - \int_{\mathscr{S}} f_n(x,y)\nu(dy)\Big| \\ &\leq \|f_n\|_\infty \|\mu - \nu\|_{\mathbf{TV}}.\end{aligned}$$

As a result, we have $\|\varphi^n(\mu) - \varphi^n(\nu)\|_\infty \leq \|f_n\|_\infty \|\mu - \nu\|_{\mathbf{TV}}$ for all $\mu, \nu \in \mathscr{P}(\mathscr{S})$. Thus, φ is a Lipschitz population game, and hence by Lahkar and Riedel (2015), we have that the continuum logit dynamic admits a continuous solution $(\mu^n(t))_{t\geq 0}$. Therefore $\widehat{\mathbb{X}}_n$ has continuous paths $\mathbb{P}$-a.s. This implies that

$$\mathbb{P}_{\mu_0}\Big(\widehat{\mathbb{X}}_n(t) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*})) \text{ for all } T_\varepsilon \leq t \leq T\Big) = \mathbb{P}_{\mu_0}\Big(\widehat{\mathbb{X}}_n(t) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*})) \text{ for all } t \in [T_\varepsilon, T] \cap \mathbb{Q}^+\Big),$$

where $\mathbb{Q}^+$ denotes the collection of all positive rationals. Let $(r_k)_{k\geq 1}$ be an enumeration of rationals in $[T_\varepsilon, T]$. Thus, in order to prove the lemma, we need to show that $\mathbb{P}\big(\cap_{k=1}^\infty \widehat{\mathbb{X}}_n(r_k) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\big) = 1$. To this end, we show that $\mathbb{P}\big(\widehat{\mathbb{X}}_n(r_k) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\big) = 1$ for all $k \geq 1$. Fix $\bar{\varepsilon} > 0$. Then, for all $N \geq 1$, it holds that

$$\begin{aligned}\Big\{\|\widehat{\mathbb{X}}_n(r_k) - \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\|_{\mathbf{TV}} > \bar{\varepsilon}\Big\} \subseteq &\Big\{\|\widehat{\mathbb{X}}_n(r_k) - \mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(r_k))\|_{\mathbf{TV}} > \bar{\varepsilon}/2\Big\} \\ &\cup \Big\{\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(r_k) - \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\|_{\mathbf{TV}} > \bar{\varepsilon}/2\Big\}.\end{aligned}$$

This implies that

$$\begin{aligned}\mathbb{P}_{\mu_0}\Big(\|\widehat{\mathbb{X}}_n(r_k) - \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\|_{\mathbf{TV}} > \bar{\varepsilon}\Big) \leq &\,\mathbb{P}_{\mu_0}\Big(\|\widehat{\mathbb{X}}_n(r_k) - \mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(r_k))\|_{\mathbf{TV}} > \bar{\varepsilon}/2\Big) \\ &+ \mathbb{P}_{\mu_0}\Big(\|\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(r_k) - \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\|_{\mathbf{TV}} > \bar{\varepsilon}/2\Big).\end{aligned}$$

We now apply Theorem 3.1 and Theorem 4.1 of Hofbauer and Sandholm (2007) to conclude that there exists $N_{\bar{\varepsilon}} \geq 1$ sufficiently large such that for all $N \geq N_{\bar{\varepsilon}}$, we have

$$\begin{aligned}\mathbb{P}_{\mu_0}\Big(\|\widehat{\mathbb{X}}_n(r_k) - \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))\|_{\mathbf{TV}} > \bar{\varepsilon}\Big) &\leq 2^{n+1} e^{-N\bar{C}\bar{\varepsilon}^2} + \bar{\varepsilon} \\ &\leq 2\bar{\varepsilon}\end{aligned}$$

Since $\bar{\varepsilon}$ is arbitrary, we conclude that $\mathbb{P}_{\mu_0}(\widehat{\mathbb{X}}_n(r_k) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))) = 1$. Also, since $k \geq 1$ is arbitrary, we have that $\mathbb{P}(\widehat{\mathbb{X}}_n(r_k) \in \mathscr{O}^{\varepsilon}(\mathrm{LE}_\eta(\varphi^n))) = 1$ for all $k \geq 1$. Now, by an application of Bonferroni's inequality (see Billingsley (2008)), we have that $\mathbb{P}(\cap_{k=1}^{\infty} \widehat{\mathbb{X}}_n(r_k) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))) = 1$. This concludes the proof of Step 2. Since $(\widehat{\mathbb{X}}_n(t))_{t\geq 0}$ is a degenerate stochastic process, we have that $\mu^n(t) \in \mathscr{O}^{\varepsilon}(\mathfrak{s}_n(\mu_{\mathbf{p}_n^*}))$ for all $T_\varepsilon \leq t \leq T$. ∎

Step 4: As a final step of the proof, we now proceed to show that the solution to the continuum strategy logit dynamic $(\mu(t))_{t\geq 0}$ is close to the continuum strategy logit equilibria $\mathrm{LE}_\eta(\varphi)$ in the medium run. We again show this in the following lemma.

Lemma 6.7. *For $\varepsilon > 0$, there exists $T_\varepsilon > 0$ such that for all $T \geq T_\varepsilon$, we have $\mu(t) \in \mathscr{O}^{3\varepsilon}(LE_\eta(\varphi))$ for all $T_\varepsilon \leq t \leq T$.*

Proof. First note that since the original supermodular game φ is supermodular, we have by Theorem 4.1 that the level-n discretization $\mathscr{D}_\varphi^n$ is supermodular for every $n \geq 1$. By assumption of the theorem, $\mathscr{D}_\varphi^n$ is irreducible for every $n \geq 1$. Thus, it follows from Hofbauer and Sandholm (2007) that the discretized stochastic evolutionary process $(\mathfrak{d}_n(\widehat{\mathbb{X}}_n^N(t)))_{t\geq 0}$ converges in the medium run to an element of $\mathrm{LE}_\eta(\mathscr{D}_\varphi^n)$. Let us denote by $\mu_{\mathbf{p}_*^n}$ the medium run limit. Again, it follows from Lemma 6.5 that $(\mathfrak{s}_n(\widehat{\mathbb{X}}_n^N(t)))_{t\geq 0}$ converges in the medium run to $\mathfrak{s}_n(\mu_{\mathbf{p}_*^n})$, which is an element of $\mathrm{LE}_\eta(\varphi^n)$. By applying Lemma, we have that 6.6, $\mu^n(t)$ stays in the medium run near $\mathfrak{s}_n(\mu_{\mathbf{p}_*^n})$. By Lemma 4.1, there exists a subsequence of $\mathfrak{s}_n(\mu_{\mathbf{p}_*^n})$ which converges to μ^* under the variational norm for some $\mu^* \in \mathscr{P}(\mathscr{S})$. Again by Lemma 4.1, the limit μ^* is a continuum strategy logit equilibrium corresponding to the game φ, that is, $\mu^* \in \mathrm{LE}_\eta(\varphi)$. We now proceed show that the solution to the continuum strategy logit dynamic $(\mu(t))_{t\geq 0}$ stays near μ^* in the medium run. To this end we make use of the approximation result (Proposition 2.1). The fact that $\sup_{0\leq t\leq T} \|\mu^n(t) - \mu(t)\|_{\mathbf{TV}} \to 0$ as $n \to \infty$, implies that for every $\varepsilon > 0$, there exists $N_\varepsilon \geq 1$ such that $\sup_{0\leq t\leq T} \|\mu^n(t) - \mu(t)\|_{\mathbf{TV}} \leq \varepsilon$ for all $n \geq N_\varepsilon$. For $T \geq T_\varepsilon$, we have by Lemma 6.6 that $\mu^n(t)$ stays in the medium run near $\mathfrak{s}_n(\mu_{\mathbf{p}_*^n})$. Again by Lemma 4.1, we have that $\mathfrak{s}_n(\mu_{\mathbf{p}_*^n}) \to \mu^*$ in the total variation norm, with μ^* being a continuum strategy logit equilibrium. Combining all these observations we have that $\mu(t)$ stays in the 3ε neighborhood of μ^* in the medium run. This completes the proof of Lemma 6.7. ∎

This concludes the proof of Theorem 4.1.

References

[1] Michel Benaïm and Jörgen W Weibull. Deterministic approximation of stochastic evolution in games. *Econometrica*, 71(3):873–903, 2003.

[2] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

[3] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[4] Kenneth G Binmore, Larry Samuelson, and Richard Vaughan. Musical chairs: Modeling noisy evolution. *Games and economic behavior*, 11(1):1–35, 1995.

[5] Tilman Börgers and Rajiv Sarin. Learning through reinforcement and replicator dynamics. *Journal of economic theory*, 77(1):1–14, 1997.

[6] Richard T Boylan. Continuous approximation of dynamical systems with randomly matched individuals. *Journal of Economic Theory*, 66(2):615–625, 1995.

[7] Man-Wah Cheung. Pairwise comparison dynamics for games with continuous strategy space. *Journal of Economic Theory*, 153:344–375, 2014.

[8] Man-Wah Cheung. Imitative dynamics for games with continuous strategy space. *Games and Economic Behavior*, 99:206–223, 2016.

[9] Man-Wah Cheung and Ratul Lahkar. Nonatomic potential games: the continuous strategy case. *Games and Economic Behavior*, 108:341–362, 2018.

[10] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pages 1047–1071, 1993.

[11] Drew Fudenberg and David K. Levine. *The theory of learning in games*. MIT press, 1998.

[12] Josef Hofbauer and Ed Hopkins. Learning in perturbed asymmetric games. *Games and Economic Behavior*, 52(1):133–152, 2005.

[13] Josef Hofbauer and William H Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

[14] Josef Hofbauer and William H Sandholm. Evolution in games with randomly disturbed payoffs. *Journal of economic theory*, 132(1):47–69, 2007.

[15] Josef Hofbauer, Jörg Oechssler, and Frank Riedel. Brown–von neumann–nash dynamics: the continuous strategy case. *Games and Economic Behavior*, 65(2):406–429, 2009.

[16] Thomas G Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of applied Probability*, 7(1):49–58, 1970.

[17] Ratul Lahkar and Saptarshi Mukherjee. Evolutionary implementation in a public goods game. *Journal of Economic Theory*, 181:423–460, 2019.

[18] Ratul Lahkar and Saptarshi Mukherjee. Evolutionary implementation in aggregative games. *Mathematical Social Sciences*, 109:137 – 151, 2021. ISSN 0165-4896. doi: https://doi.org/10.1016/j.mathsocsci.2020.11.004.

[19] Ratul Lahkar and Vinay Ramani. An evolutionary approach to pollution control in competitive markets. *Dynamic Games and Applications*, pages 1–25, 2021.

[20] Ratul Lahkar and Frank Riedel. The logit dynamic for games with continuous strategy sets. *Games and Economic Behavior*, 91:268–282, 2015.

[21] Ratul Lahkar and William H Sandholm. The projection dynamic and the geometry of population games. *Games and Economic Behavior*, 64(2):565–590, 2008.

[22] Ratul Lahkar, Sayan Mukherjee, and Souvik Roy. Generalized perturbed best response dynamics with a continuum of strategies. *Journal of Econimic Theory*, 200(105398), 2022.

[23] Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, pages 1255–1277, 1990.

[24] Jörg Oechssler and Frank Riedel. Evolutionary dynamics on infinite strategy spaces. *Economic Theory*, 17(1):141–162, 2001.

[25] Jörg Oechssler and Frank Riedel. On the dynamic foundation of evolutionary stability in continuous models. *Journal of Economic Theory*, 107(2):223–252, 2002.

[26] Steven Perkins and David S Leslie. Stochastic fictitious play with continuous action sets. *Journal of Economic Theory*, 152:179–213, 2014.

[27] William H Sandholm. Evolution and equilibrium under inexact information. *Games and economic behavior*, 44(2):343–378, 2003.

[28] William H Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

[29] Karl H Schlag. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of economic theory*, 78(1):130–156, 1998.

[30] Donald M Topkis. *Supermodularity and complementarity*. Princeton university press, 1998.

[31] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.