Dynamics in Congestion Games

Devavrat Shah
Department of EECS
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
devavrat@mit.edu

Jinwoo Shin[*]
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
jinwoos@mit.edu

ABSTRACT

Game theoretic modeling and equilibrium analysis of congestion games have provided insights in the performance of Internet congestion control, road transportation networks, etc. Despite the long history, very little is known about their transient (non equilibrium) performance.

In this paper, we are motivated to seek answers to questions such as how long does it take to reach equilibrium, when the system does operate near equilibrium in the presence of dynamics, e.g. nodes join or leave. In this pursuit, we provide three contributions in this paper. First, a novel probabilistic model to capture realistic behaviors of agents allowing for the possibility of arbitrariness in conjunction with rationality. Second, evaluation of (a) time to converge to equilibrium under this behavior model and (b) distance to Nash equilibrium. Finally, determination of trade-off between the rate of dynamics and quality of performance (distance to equilibrium) which leads to an interesting *uncertainty* principle. The novel technical ingredients involve analysis of logarithmic Sobolov constant of Markov process with time varying state space and methodically this should be of broader interest in the context of dynamical systems.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Reliability, availability, and serviceability; G.3 [**Probability and Statistics**]: Markov processes

General Terms

Algorithms, Performance, Reliability

Keywords

Logit-response, Congestion game, Logarithmic Sobolov constant

1. INTRODUCTION

In the recent years, game theoretic frameworks have provided sound models for analyzing the performance of large networks formed out of independent, autonomous or non-engineered agents. A successful example is the study of equilibrium behavior of flow level, bandwidth sharing model of the Internet (cf. Kelly, Maullo and Tan [13]) under selfish behaviors of agents, users, computers or network nodes. Specifically, the result of Johari and Tsitsiklis [11] (also see Roughgarden and Tardos [20]) suggests that despite the selfish behaviors of agents or users, the performance loss (compared to optimal allocation) incurred under the equilibrium is *small*. This result about limited performance loss, popularly known as the small *price of anarchy* (cf. [14]), holds for the broader class of *congestion games* [2][20][6][19].

An interesting feature (and hence usefulness) of congestion games is that the network reaches the Nash equilibrium under simple, myopic *best response* strategy : each agent updates its action selfishly at every available opportunity. Under more realistic modeling, one expects agents to make *partly* rational or selfish decisions. For example, each agent updates its action selfishly with some probability and arbitrarily otherwise. For a wide class of such partly myopic, selfish behavior models, including the popular *logit-response*, it is well understood that equilibrium is reached in an asymptotic sense.

In summary, under a reasonable behavioral model of autonomous agents (cf. partly selfish and myopic), Nash equilibrium is reached and the performance loss under the equilibrium state is limited. However, in reality we expect the network to be in transience continually, e.g. agents join or leave the network. Therefore, it is important to understand the transience properties of the network evolving under reasonable behavioral models of agents. And very little (or nothing) is known in the literature about them.

In this paper, we shall undertake the study of transient network properties. Specifically, we shall examine the rate of convergence to equilibrium under a class of behavioral model for agents and study the effect of network dynamics, in terms of players joining or leaving, on the network performance.

1.1 Related Work

Here we briefly describe relevant prior works. To this end, the congestion game was introduced in [2]. The congestion

game is an instance of the symmetric potential games. A game is symmetric if an agent's utility (or payoff) depends only on other agents' aggregate actions, but not on their identities. And a game is potential if a (marginal) payoff due to change in an action for any agent can be described by the marginal payoff due to the same action change in a single global function. Indeed, congestion games satisfy both of these properties (see Example 1 in Section 2.1).

The potential games were introduced by Monderer and Shapley [17]. The potential games are widely practically applicable including congestion games. The best response mechanism reaches Nash equilibrium for any potential game (cf. the result by Cournot [7] for duopoly games). In this mechanism, agents update their strategies sequentially, by choosing the best possible strategy against the other agents' choices. While the best response mechanism is simple and myopic, in reality one does not expect agents to be fully rational. This led to the study of the best response with *error*: agents behave as per the best response with probability $1-\varepsilon$ and respond arbitrarily (commit mistake or error) with probability ε. Various results about the long-term behavior of this mechanism has been obtained by Friedlin [15], Kandori, Mailath and Rob [12] and Young [22]. However such results are often criticized for their extreme sensitivity to the underlying error model (cf. [3]).

In response to this criticism as well as to provide a more realistic behavioral model, Blume introduced the logit-response mechanism [4]. Here, each agent chooses strategy probabilistically with larger probabilities for actions with larger payoffs. More precisely, the probability is chosen as per the *logit* form and hence the name logit-response. A parameter, usually denoted by $\beta > 0$, governs the *intensity of rationality*: larger β, higher chance of agent choosing the best response and as $\beta \to \infty$ the logit-response mechanism becomes the standard best response mechanism.

Blume [5] observed that for any potential game the logit-response mechanism leads to a reversible Markov process with the product form stationary distribution. And as $\beta \to \infty$, this stationary distribution concentrates on a Nash equilibrium. We take a note of utilization of the logit-response mechanism in the context of design of control for networked systems [16].

Clearly, many natural and important questions remain unanswered. To begin with, we wish to understand the probabilistic distance to the Nash equilibrium, under the stationary distribution, for a given intensity of rationality $\beta > 0$. Next, determination of the time it takes to converge (close) to the stationary distribution under the logit-response mechanism. This will suggest that if the system changes at a slower time scale than the convergence time, then it will remain close to the stationary distribution. Finally, we wish to characterize the effect of dynamics on the performance for the entire spectrum of dynamics. But most importantly, we would like to come up with a more realistic behavioral model of agents that can capture aspects that are missing in the standard logit-response model.

1.2 Our Contributions

As the main result of this paper, we answer all of the above questions in the context of symmetric potential games, which includes the congestion game as a special instance.

To begin with, we define the notion of *universal* symmetric potential games so as to allow us formally study the effect of

dynamics in terms of agents (or players) joining or leaving. Indeed, the congestion game naturally extends to become an instance of universal symmetric potential games.

First, we study the stationary distribution under the logit-response mechanism. This suggests that, for any finite $\varepsilon > 0$, in order to be ε-close to the Nash equilibrium in any sense, β must scale as $\Omega(n)$, for games of n agents. That is, for the logit-response to be effective, it ought to be close to the best response. An immediate implication of this is that for a arbitrary symmetric potential function, the convergence time under the logit-response may need to be exponential in n for $\beta = \Omega(n)$ (see Example 2 in Section 2.2). In summary, Logit-response is rather undesirable from the perspective of error in the performance as well as the convergence rate.

The logit-response mechanism misses the following aspect – if an agent finds that she is playing a strategy that is played by a small fraction of other agents then she is likely to be more anxious to verify whether her current strategy is indeed a good choice. We explicitly model this aspect and provide a very minor modification of the standard logit-response mechanism. Essentially, in contrast to the standard logit-response in which each player updates her strategy at uniform rate, in the modified mechanism each player updates her strategy in a non-uniform manner.

Under this modified logit-response, we characterize the stationary distribution. We find it to be ε-close to the Nash equilibrium for β scaling as $\varepsilon \log 1/\varepsilon$, in contrast to $\Omega(n)$ as per the standard logit-response. Further, the convergence to the stationary distribution happens in essentially linear (in n) time, exponentially faster compared to the standard logit-response.

Finally, we study the effect of dynamics, in terms of agents joining or leaving, on the performance of our modified logit-response mechanism. We consider the scenario where the number of agents can change arbitrarily but at a bounded rate. We find a precise relation between the performance error (distance to the Nash equilibrium) and this rate of dynamics.

To establish our results, especially under the dynamic setup of agents, we develop a novel technique to analyze the *mixing time* of Markov process of which state space is changing over time. To obtain sharp results, we study the evolution of entropy distance between the empirical distribution of Markov process and its stationary distribution. This requires to evaluate a logarithmic Sobolov constant of Markov process. Evaluating a logarithmic Sobolov constant is harder in general than the more popular spectral analysis (e.g. spectral gap, conductance, canonical path, etc.), but is crucial for our success. Our evaluation is building upon the work of Frieze and Kannan [10].

In the context of congestion games, our results can be interpreted as follows. The modified logit-response mechanism provides a more realistic model for agents' behaviors, such as drivers on the road or computers which uses the Internet. The quick convergence to near Nash equilibrium even with small intensity of rationality and robustness of the long-term behavior with small price of anarchy suggests that in reality even though players are only partly rational and network is highly dynamic, the network operates near optimal. That is, for congestion games the *dynamic price of anarchy* is *small* !

Organization. Section 2 provides the necessary notions, the symmetric potential game and the logit-response learn-

ing mechanism. We also introduce our new notion of the *universal symmetric potential game* in Section 2.1. In Section 3, we present our main results, the modified logit-response and its robustness (or uncertainty principle) in the dynamics setup i.e. players join or leave over time. Section 4, 5 and 6 are dedicated to prove our theoretical claims.

2. SETUP

In this section, we start with description of symmetric potential game, the concept of Nash equilibrium and relation to the optima of potential function. We recall the congestion game and explain it as an instance of symmetric potential games. We describe a popular behavioral model of the logit-response mechanism and state its property for our setup.

2.1 Symmetric Potential Game

A game $G = (n, s, \{u_p\})$ consists of n agents, players or nodes[1]. Each player can play one of the s strategies ($s \geq 2$) denoted as $[s] = \{1, \ldots, s\}$. Let $u_p : [s]^n \to \mathbb{R}$ be the utility or payoff function of player p, for $1 \leq p \leq n$. That is, the payoff (profit/utility) obtained by player p is $u_p(s_1, \ldots, s_n)$ when the strategy played by all n players are $s_1, \ldots, s_n$ respectively. Throughout the paper, our interest will be when s is *fixed* and *small* while n is *large* and *dynamic*.

Naturally, the selfish goal for each player is to maximize her own profit. However the profit of player, in addition to her own strategy, depends on the strategy of other players. Therefore, in a totally rational world, if a player can improve her own profit by changing her strategy, she will do it. Therefore, an equilibrium is the state in which no player can improve her payoff by changing her strategy unilaterally. This leads to the well known notion of the *pure* Nash equilibrium.

DEFINITION 1 (PURE NASH EQUILIBRIUM). *Strategies* **s** $= (s_1, s_2, \ldots, s_n)$ *is a pure Nash equilibrium of* $G = (n, s, \{u_p\})$ *if for each player* $p \in [n]$,

$$u_p(\mathbf{s}) \geq \max_{i \in [s]} u_p(s_1, \ldots, s_{p-1}, i, s_{p+1}, \ldots, s_n).$$

In general, a game may not have a pure Nash equilibrium. For the class of games of interest in this paper, the symmetric potential games do posses pure Nash equilibrium. To this end, we introduce definitions of potential and symmetric games.

DEFINITION 2 (EXACT POTENTIAL GAME [17]). *A game* G *is called an exact potential game if there exists a potential function* $P : [s]^n \to \mathbb{R}$ *such that*

$$u_p(i, \mathbf{s}_{-p}) - u_p(j, \mathbf{s}_{-p}) = P(i, \mathbf{s}_{-p}) - P(j, \mathbf{s}_{-p}). \quad (1)$$

For a potential game, it is well known that $\mathbf{s}^*$ is a pure Nash equilibrium if

$$\mathbf{s}^* \in \arg \max_{\mathbf{s} \in [s]^n} P(\mathbf{s}).$$

It is also known [21, 9] that G is an exact potential game if and only if there exists a potential function $P : [s]^n \to \mathbb{R}$ and auxiliary function $H : [s]^{n-1} \to \mathbb{R}$ such that

$$u_p(s_1, s_2, \ldots, s_n) = P(s_1, s_2, \ldots, s_n) + H(\mathbf{s}_{-p}), \quad (2)$$

where $\mathbf{s}_{-p} := (s_1, s_2, \ldots, s_{p-1}, s_{p+1}, \ldots, s_n)$.

[1] Throughout this paper, we shall use terms *agent, player* and *node* interchangeably.

DEFINITION 3 (SYMMETRIC GAME). *A game* G *is called a symmetric game if for any permutation* π *of* $\{1, \ldots, n\}$,

$$u_p(s_1, s_2, \ldots, s_n) = u_{\pi(p)}(s_{\pi(1)}, s_{\pi(2)}, \ldots, s_{\pi(n)}).$$

An important property in symmetric games is that the payoff of a player p for any given strategy s_p depends on the other players strategy only through their aggregate behaviors – i.e. how many other players are playing strategy 1, ..., strategy s matters, and identities of players do not.

We call a game G *symmetric potential* if it is both symmetric and exact potential. Specifically, as per (2) for such a game, it must be that P and H are symmetric. That is for any permutation π of n and $\mathbf{s} = (s_1, \ldots, s_n) \in [s]^n$,

$$P(s_1, \ldots, s_n) = P(s_{\pi(1)}, \ldots, s_{\pi(n)}),$$

and similarly for H. Therefore, value of P (resp. H) does not depend on which player exactly plays what strategy, but depends only on the aggregate information of how many players play a particular strategy. Therefore, in a symmetric potential game, the potential function P (resp. H) can be re-defined in terms of a lower-dimensional function $\mathcal{P} : \Psi_n^s \to \mathbb{R}$ (resp. $\mathcal{H} : \Psi_{n-1}^s \to \mathbb{R}$), so that for any $\mathbf{s} = (s_1, \ldots, s_n) \in [s]^n$,

$$P(s_1, \ldots, s_n) = \mathcal{P}(x_1(\mathbf{s}), \ldots, x_s(\mathbf{s})),$$

where $x_j(\mathbf{s}) = \frac{1}{n}|\{p \in [n] : s_p = j\}|$. Throughout, we shall use this notation

$$x(\mathbf{s}) = (x_1(\mathbf{s}), \ldots, x_s(\mathbf{s})) \quad \text{for } \mathbf{s} \in [s]^n.$$

Also we shall use Ψ_n^s to denote

$$\Psi_n^s = \left\{ \left(\frac{v_1}{n}, \ldots, \frac{v_s}{n} \right) : v_i \in \mathbb{Z}_+ \text{ for all } i \in [s], \sum_{i=1}^s v_i = n \right\}.$$

And, hence the payoff or utility function of each player p will have form

$$u_p(\mathbf{s}) = \mathcal{P}(x(\mathbf{s})) + \mathcal{H}(x(\mathbf{s}_{-p})), \quad \text{for any } \mathbf{s} \in [s]^n. \quad (3)$$

We shall use P and $\mathcal{P}$ interchangeably throughout. Therefore, in what follows, one may think of P as a function $\Psi_n^s \to \mathbb{R}$. Next we present the congestion game as an instance of symmetric potential games.

EXAMPLE 1 (CONGESTION GAME). *A congestion game is an n player game in which each player's strategy consists of a set of resources, and the cost of the set of strategy, say $[s] = \{1, \ldots, s\}$, depends only on the number of players using each resource, i.e. the cost takes form $\sum_e d_e(x_e(\mathbf{s}))$, where $x_e(\mathbf{s})$ is the number of players using resource e and d_e is a non-negative increasing function. A standard example is a* network *congestion game on a directed graph (e.g. road or transportation network), in which each player must select a path from some source to some destination, and each edge has an associated "delay" function that increases with the number of players using the edge. It is well known that this game admits potential function*

$$\phi(\mathbf{s}) = -\sum_e \int_0^{x_e(\mathbf{s})} d_e(z)dz.$$

Clearly, this is a symmetric potential function and hence an instance of symmetric potential games. And the maximizer of $\phi(\mathbf{s})$ (or minimizer of $\sum_e \int_0^{x_e(\mathbf{s})} d_e(z)dz$) is a Nash equilibrium. The relation between the "delay" induced by such

an equilibrium state and the socially optimal solution, minimizing total "delay" $\sum_e x_e(\mathbf{s}) \cdot d_e(x_e(\mathbf{s}))$, has been well studied under the *popular* price of anarchy *literature in the past decade, cf. [14][20][19][11].*

In this paper, we shall be interested in studying the setup where the number of players are changing over time. To be able to consistently define the notion of symmetric potential game for such a dynamic setup without *cumbersomeness*, we introduce the notion of *universal symmetric potential game* defined as follows.

DEFINITION 4 (UNIVERSAL SYMMETRIC POTENTIAL). *For a given* $s \in \mathbb{N}$, $\mathcal{P} : \Psi_\infty^s \to \mathbb{R}$ *and* $\mathcal{H} : \Psi_\infty^s \to \mathbb{R}$, *a sequence of games* $G_n = (n, s, u_p^n)$ *for any number of players* $n \geq 2$, *is called a universal symmetric potential game* $(s, \mathcal{P}, \mathcal{H})$ *if the payoff functions of* G_n *is given by*

$$u_p^n(\mathbf{s}) = \mathcal{P}(x(\mathbf{s})) + \mathcal{H}(x(\mathbf{s}_{-p})),$$

for any $\mathbf{s} \in [s]^n$.

In above, Ψ_∞^s is defined as:

$$\Psi_\infty^s := \left\{ (x_1, \ldots, x_s) : x_i \in [0,1] \text{ for all } i \in [s], \sum_{i=1}^{s} x_i = 1 \right\}.$$

Clearly, if the number of players is fixed, this definition reduces to the standard non-dynamic setup.

2.2 Learning Mechanism : Logit-Response

Our interest is in understanding the transience properties of universal symmetric potential games under a natural behavioral setup. As discussed earlier, it is only reasonable to expect players or agents to utilize simple, myopic learning rules to choose their strategies over time. For example, as a car driver using a road network everyday, she will update her route selection daily by reacting to delays observed over the recent past using a simple, myopic selfish rule. The logit-response learning rule or mechanism provides a reasonable model for it. We briefly recall its precise definition in our setup.

We shall consider an asynchronous version of the logit-response learning mechanism. Let us consider it for a symmetric potential game $G = (n, s, \mathcal{P}, \mathcal{H})$. In Logit-response [1], every player p has an independent Exponential clock of rate 1: that is, the times between two consecutive clock-ticks are independent and distributed as the exponential distribution of mean 1. When the clock of player p ticks, she obtains an opportunity to revise her strategy. And, she chooses to play strategy $i \in [s]$, till the next clock ticks, with probability

$$\frac{e^{\beta u_p(i, \mathbf{s}_{-p})}}{\sum_{j \in [s]} e^{\beta u_p(j, \mathbf{s}_{-p})}}. \tag{4}$$

In above, recall that $\mathbf{s}_{-p} \in [s]^{n-1}$ is the current strategy profile of other players and $\beta > 0$ is some constant. As β becomes larger, the player chooses strategy with the best payoff given $\mathbf{s}_{-p}$ with higher probability. In that sense, parameter β serves as the index of *rationality*. And, the finite value of β models the possible non-rational behaviors of players as one may expect in real scenarios. Using (1) and (3), in symmetric potential games, the updating probability can

be simplified to

$$\frac{e^{\beta u_p(i, \mathbf{s}_{-p})}}{\sum_{j \in [s]} e^{\beta u_p(j, \mathbf{s}_{-p})}} = \frac{e^{\beta \mathcal{P}(x(i, \mathbf{s}_{-p}))}}{\sum_{j \in [s]} e^{\beta \mathcal{P}(x(j, \mathbf{s}_{-p}))}}.$$

The above Logit-response induces a continuous time, reversible and irreducible Markov chain on the (finite) state space Ψ_n^s. Then, the following characterization of its unique stationary (invariant) distribution $\boldsymbol{\pi}$ follows from the standard arguments using the reversibility.

LEMMA 1. *The stationary distribution* $\boldsymbol{\pi} = [\pi_{\mathbf{x}}]_{\mathbf{x} \in \Psi_n^s}$ *of the (asynchronous) logit-response of symmetric potential game* $G = (n, s, \mathcal{P}, \mathcal{H})$ *is*

$$\pi_{\mathbf{x}} \propto e^{\beta \mathcal{P}(\mathbf{x})} \begin{pmatrix} n \\ nx_1 \ \ldots \ nx_s \end{pmatrix} \approx e^{\beta \mathcal{P}(\mathbf{x}) + nH(\mathbf{x})}, \quad (5)$$

with $H(\mathbf{x}) = -\sum_i x_i \ln x_i$, *for any* $\mathbf{x} = (x_1, \ldots, x_s) \in \Psi_n^s$.

Now if $\beta = o(n)$, i.e. a player's "rationality" scales slower than n, then

$$\beta \mathcal{P}(\mathbf{x}) + nH(\mathbf{x}) = (1 \pm o(1))nH(\mathbf{x}),$$

where we assume $\mathcal{P}$ is bounded below and above (independent of n). Hence essentially $\pi_{\mathbf{x}} \propto e^{nH(\mathbf{x})}$. Therefore, for n large enough the distribution $\boldsymbol{\pi}$ concentrates on the uniform strategy profile, i.e. $\mathbf{x} \approx [1/s]$ with high probability. And, hence the players' payoff function (or preferences) become irrelevant. That is, to have any reasonable equilibrium (i.e. the stationary distribution $\boldsymbol{\pi}$) under Logit-response in our setup with large number of players, n, it is essential that β scales at least proportional to n, i.e. $\beta = \Omega(n)$. In that case, as we explain in the following example, the worst convergence time of Logit-response to equilibrium is exponential in n.

EXAMPLE 2. *Suppose* $\beta = \Omega(n)$ *i.e.* $\beta \geq cn$ *for some* $c > 0$. *Consider the example when* $s = 2$ *and*

$$\mathcal{P}(x_1, x_2) = |2x_1 - x_2| - \frac{H(x_1, x_2)}{c}.$$

First observe that $\pi_{\mathbf{x}} \propto e^{\beta|2x_1 - x_2| - \zeta H(x_1, x_2)}$ *from* (5), *where* $\zeta \geq 0$. *Then, it is easy[2] to check that it takes* $\Omega(e^\beta)(= e^{\Omega(n)})$ *time for Logit-response starting from the initial state* $\mathbf{x}_1 = (0,1) \in \Psi_n^s$ *to reach* $\mathbf{x}_2 = (1/3, 2/3) \in \Psi_n^s$. *Hence, it also takes exponential time to reach the maximizer* $(1,0) \in \Psi_n^s$ *of* $\mathcal{P}$.

In summary, under Logit-response for symmetric potential games, (a) to reach reasonable equilibrium, the "rationality" of players must scale with the number of players n, and (b) for such a setting, the time to reach equilibrium is quite large. Due to this undesirable transience property, it is reasonable to expect that under Logit-response, the network state is very *fragile* to dynamics in terms of players.

3. MAIN RESULTS

Two key results of this paper are state here. The poor transient properties (or fragility) of the Logit-response mechanism raises a natural question : is there a simple, Logit-response *like* mechanism that has much nicer transient or convergence properties and subsequently is *robust* to dynamics in the network ?

---

[2]Use the fact that $\pi_{\mathbf{x}_1}/\pi_{\mathbf{x}_2} \geq e^\beta$.

3.1 Efficient Learning Mechanism

We propose a novel Logit-response like mechanism with the desired transient and robustness properties. It is exactly same as Logit-response – every player has an Exponential clock and when it ticks she updates its strategy probabilistically as per (4). The only *minor* change is that the clock rate of each player is time varying, unlike the fixed unit rate as in the standard Logit-response. And it is merely a function of number of players playing the same strategy at that time.

Specifically, let $\mathbf{s}(t) = (s_1(t), \ldots, s_n(t)) \in [s]^n$ be the strategies that n players are playing at time t. Then, the Exponential clock of any player p is $\alpha/z_p(t)$, where $z_p(t)$ is the fraction of players (including p) that are playing the same strategy as p at time t, i.e. $z_p(t) = \frac{1}{n}|\{q \in [n] : s_q(t) = s_p(t)\}|$. Here $\alpha > 0$ is a parameter. When the clock of player p ticks, she chooses her strategy probabilistically from $[s]$ as per (4). We shall call this the *modified* Logit-response learning mechanism with parameters α, β.

This mechanism induces a reversible and irreducible Markov process on Ψ_n^s. Somewhat surprisingly, we find that this minimal change leads to the following *exponentially twisted* distribution that removes the dependence on the 'entropy term', $nH(\mathbf{x})$, that was present in the stationary distribution of the standard Logit-response. As we shall see, this leads to the desired properties we have listed above.

LEMMA 2. *Given symmetric potential game $G = (n, s, \mathcal{P}, \mathcal{H})$, the stationary distribution under the modified Logit-response with parameters α, β, is*

$$\pi_{\mathbf{x}} \quad \propto \quad e^{\beta \mathcal{P}(\mathbf{x})}, \quad \mathbf{x} \in \Psi_n^s. \tag{6}$$

The proof of Lemma 2 is explained in Section 4.2. The parameter $\alpha > 0$ does not play a role in characterizing the stationary distribution but in the time to reach equilibrium. Next, we compare the total rate of changes or the average number of updates per unit time between the standard and modified versions of the Logit-response. Under the standard Logit-response, it is n. And, under the modification it becomes $\sum_{i=1}^s n x_i \frac{\alpha}{x_i} = \alpha s n$ since $n x_i$ players have clock rate α/x_i for $i \in [s]$. Thus, if $\alpha = 1/s$, then both version of the learning mechanisms have exactly the same effect update rate. However, as we state next, the time to reach near equilibrium for a good choice of β (i.e. the stationary distribution has near Nash equilibrium properties) under our modified Logit-response is essentially linear in n – this is in sharp contrast to what we would expect, i.e. exponentially in n, for the standard Logit-response.

THEOREM 3. *Given a symmetric potential game $G = (n, s, \mathcal{P}, \mathcal{H})$ with $\mathcal{P} : \Psi_n^s \to [0, 1]$. Let the potential function $\mathcal{P}$ be λ-Lipschitz, i.e.*

$$|\mathcal{P}(\mathbf{x}_1) - \mathcal{P}(\mathbf{x}_2)| \leq \lambda \|\mathbf{x}_1 - \mathbf{x}_2\|_1, \qquad \forall \, \mathbf{x}_1, \mathbf{x}_2 \in \Psi_n^s. \tag{7}$$

For any given $\varepsilon \in (0, 1)$, starting with any initial strategy state at time 0, under the modified Logit-response with parameters α, β such that

$$\beta \geq \max\left\{ \frac{4(s-1)}{\varepsilon} \log 2s, \frac{4(s-1)}{\varepsilon} \log \frac{8s\lambda}{\varepsilon} \right\} \approx \Theta\left(\frac{1}{\varepsilon} \log \frac{1}{\varepsilon} \right), \tag{8}$$

then

$$\mathbb{E}\left[\mathcal{P}(x(\mathbf{s}(t)))\right] \geq \sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) - \varepsilon, \tag{9}$$

for

$$t \geq \frac{ne^{3\beta}}{\alpha c}\left(\log \log n + \log \beta + \log \frac{1}{\varepsilon} \right) \approx \Theta\left(e^{3\beta} n \log \log n \right). \tag{10}$$

Here, the constant $c = c(s) > 0$ depends only on s, the number of distinct strategies.

Few remarks about the result as well as the interpretation of the modified Logit-response are in order. Due to removal of dependency on the entropy term in the stationary distribution, we find that even for little rationality, i.e. β scaling essentially as $\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$, under the stationary distribution the strategy profile is ε-close to Nash equilibrium in the sense of (9). For such a choice of β and $\alpha = 1/s$, from (10) it follows that the time to reach near such a good state is $O(n \log \log n)$ which is essentially the best one can expect. Thus with the same total update rate, the modified Logit-response appears exponentially (in n) faster than the standard Logit-response. Now it is worth pondering whether the minor modification we have suggested is reasonable. To this end, first observe that the time varying rate requires each player to know the aggregate information of strategies of other players, which is needed anyways for the player to even evaluate her payoff. Next, the modification captures the intuition that if a player finds herself playing a strategy that is played by too few other players, she gets 'alarmed' and checks rationality of playing her current action quickly – of course, if she finds her current strategy reasonable, then she does not change as captured by (4). Finally, we shall use the definition that a potential game is ε-predictable, equivalently ε-close to Nash equilibrium, under a learning mechanism if it satisfies inequality (9) with respect to its long-term or stationary distribution of the strategy profile of players.

3.2 Robustness of Modified Logit-Response

As the second key result of the paper, we study the robustness of our modified Logit-response learning mechanism with respect to dynamics in the number of players. To this end, let $n(t)$ be the number of players at time $t \in \mathbb{R}_+$. At any time t, if a new player joins than $n(t)$ increases by one, i.e. $n(t) = n(t^-) + 1$; if an existing player leaves, then $n(t) = n(t^-) - 1$. Under this dynamic setup, for a given universal symmetric potential game $G = (s, \mathcal{P}, \mathcal{H})$, the modified Logit-response learning mechanism naturally extends. That is, at time t each player $p \in [n(t)]$ has an Exponential clock with rate $\alpha/z_p(t)$ where as before $z_p(t)$ is the fraction of players playing the same strategy as player p at time t. When a clock of player p ticks, she updates her strategy as per (4).

Here, $n(t)$ (and the state space $\Psi_{n(t)}^s$ as well) is changing with t. The state $\mathbf{s}(t)$, strategy profile of $n(t)$ players, is evolving as per the modified Logit-response as explained above. Since this is a symmetric potential game, the aggregate strategy profile $\mathbf{x}^*(t) \in \arg\max_{\mathbf{y} \in \Psi_{n(t)}^s} \mathcal{P}(\mathbf{y})$ is a Nash equilibrium. Ideally, we would like the aggregate state $\mathbf{x}(\mathbf{s}(t))$ so that $\mathcal{P}(\mathbf{x}(\mathbf{s}(t))) \approx \mathcal{P}(\mathbf{x}^*(t))$. If $n(t)$ is fixed, then as stated in Theorem 3, this is indeed true for all $t = \Omega(n \log \log n)$. If $n(t)$ is changing very *wildly*, clearly $\mathbf{x}(\mathbf{s}(t))$ should be arbitrary. Therefore, what one can hope for is the characterization of tradeoff between $|\mathcal{P}(\mathbf{x}(\mathbf{s}(t))) - \mathcal{P}(\mathbf{x}^*(t))|$ and the rate of change in $n(t)$. Indeed, we obtain such a characterization suggesting that there is an inherent *uncertainty* in the nearness to Nash equilibrium (ε-predictability)

and the (rate $(1/\Lambda)$ of) dynamics as stated below. Specifically, as reader will notice, (11) and (16)-(17) suggest that to be ε-predictable at all times, the rate of change $1/\Lambda$ should be slower than $O(\varepsilon^2 \exp(-4\beta))$ for β scaling as $\Omega\left(\frac{1}{\varepsilon}\log\frac{\lambda}{\varepsilon}\right)$.

THEOREM 4. *Given a universal symmetric potential game* $G = (s, \mathcal{P}, \mathcal{H})$ *with* $\mathcal{P} : \Psi_\infty^s \to [0, 1]$. *Let* $\mathcal{P}$ *be* λ-*Lipschitz, i.e.*

$$|\mathcal{P}(\mathbf{x}_1) - \mathcal{P}(\mathbf{x}_2)| \leq \lambda \|\mathbf{x}_1 - \mathbf{x}_2\|_1, \qquad \forall\, \mathbf{x}_1, \mathbf{x}_2 \in \Psi_\infty^s, \ (11)$$

Let the number of players $n(\cdot)$ *evolves so that for a given* $\Lambda > 0$

$$|n(t + \Lambda) - n(t)| \leq 1, \quad \forall\, t \geq 0. \tag{12}$$

Here $1/\Lambda \in (0, \infty)$ *represents the rate of change in* $n(t)$. *The players evolve as per the modified Logit-response with parameters* α, β. *We shall assume that* $n(t)$ *is large enough so that for all* $t \geq 0$,

$$n(t) \geq \max\left\{4s\alpha c_0 e^{-3\beta}\Lambda, 2\beta\lambda\right\}. \tag{13}$$

Then, for any given $\varepsilon \in (0, 1)$

$$\mathbb{E}\left[\mathcal{P}\left(x(\mathbf{s}(t))\right)\right] \geq \sup_{\mathbf{y} \in \Psi_{n(t)}^s} \mathcal{P}(\mathbf{y}) - \varepsilon, \tag{14}$$

for

$$t \geq \frac{2n(0)e^{3\beta}}{\alpha c_1 \varepsilon^2}\left(\log n(0) + \beta\right) \approx \Theta\left(e^{3\beta}n(0)\log n(0)\right), \tag{15}$$

as long as

$$\beta \geq \max\left\{\frac{4(s-1)}{\varepsilon}\log 2s, \frac{4(s-1)}{\varepsilon}\log\frac{8s\lambda}{\varepsilon}\right\} \approx \Theta\left(\frac{1}{\varepsilon}\log\frac{\lambda}{\varepsilon}\right) \tag{16}$$

$$\Lambda \geq 2\varepsilon^{-2}e^{3\beta}\left(\frac{6\beta\lambda + e^\beta(s-1)}{sc_0\alpha}\right) \approx \Theta\left(\varepsilon^{-2}e^{4\beta}\lambda\right). \tag{17}$$

In above, constants c_0, c_1 *are strictly positive and dependent only on* s.

4. PRELIMINARIES

4.1 Markov Chain & Mixing Time

Consider a discrete-time Markov chain $\{X_\tau\}_{\tau \in \mathbb{Z}_+}$ over a finite state space Ω. Let an $|\Omega| \times |\Omega|$ matrix M be its transition probability matrix:

$$\boldsymbol{\mu}(\tau) = \boldsymbol{\mu}(\tau - 1)M = \boldsymbol{\mu}(0)M^\tau,$$

where $\boldsymbol{\mu}(\tau)$ is the distribution of $X_\tau \in \Omega$. If M is irreducible and aperiodic, then the Markov chain has a unique stationary distribution $\boldsymbol{\pi}$ and it is ergodic in the sense that $\lim_{\tau \to \infty} \boldsymbol{\mu}(\tau) = \boldsymbol{\pi}$. The adjoint of the transition matrix M, also called the time-reversal of M, is denoted by M^* and defined as: for any $i, j \in \Omega$, $\pi_i M^*(i,j) = \pi_j M(j,i)$.[3] By definition, M^* has $\boldsymbol{\pi}$ as its stationary distribution as well. If $M = M^*$ then M is called *reversible*.

The continuous-time Markov process $\{X_t\}_{t \in \mathbb{R}_+}$ over a finite state space Ω can be characterized using a discrete-time Markov chain M. For $t \geq 0$, $e^{t(M-I)}$ represents the transition matrix of the process:

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}(0)e^{t(M-I)}.$$

---

[3]Throughout this paper, bold letters (e.g. $\mathbf{u}$) are reserved for vectors or distributions.

We call M as the kernel of the Markov process.

The distribution $\boldsymbol{\mu}(t)$ of the continuous-time Markov process with its irreducible and aperiodic kernel M converges to the stationary distribution $\boldsymbol{\pi}$ of M starting from any initial condition $\boldsymbol{\mu}(0)$. To establish our results, we will need quantifiable bounds on the time it takes for the process to reach close to its stationary distribution – popularly known as *mixing time*. To make this notion precise and recall known bounds on mixing time, we start with definitions of the distance between probability distributions.

DEFINITION 5 (DISTANCES OF MEASURES). *Given two probability distributions* $\boldsymbol{\mu}$ *and* $\boldsymbol{\nu}$ *on a finite space* Ω, *we define the following two distances. The total variation distance, denoted as* $\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{TV}$, *is*

$$\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{TV} = \frac{1}{2}\sum_{i \in \Omega}|\mu_i - \nu_i|.$$

The relative entropy (or Kullback-Leibler divergence), denoted as $D(\boldsymbol{\mu} : \boldsymbol{\nu})$, *is*

$$D(\boldsymbol{\mu} : \boldsymbol{\nu}) = \sum_{i \in \Omega}\mu_i \log\frac{\mu_i}{\nu_i}.$$

We make note of the following relation between the above distances:

$$\|\boldsymbol{\nu} - \boldsymbol{\mu}\|_{TV} \leq \sqrt{\frac{D(\boldsymbol{\mu} : \boldsymbol{\nu})}{2}}. \tag{18}$$

The mixing time can be quantified using these distances. For $\varepsilon > 0$ and given initial distribution $\mu(0)$,

$$T_{TV}(\varepsilon) = \min_t \left\{\|\boldsymbol{\mu}(t) - \boldsymbol{\pi}\|_{TV} \leq \varepsilon\right\}.$$
$$T_D(\varepsilon) = \min_t \left\{D(\boldsymbol{\mu}(t) : \boldsymbol{\pi}) \leq \varepsilon\right\}.$$

If the kernel M of the process is irreducible, it is known [8] that $D(\boldsymbol{\mu}(t) : \boldsymbol{\pi})$ is exponentially decaying:

$$D(\boldsymbol{\mu}(t) : \boldsymbol{\pi}) \leq e^{-4t\rho(M)}D(\boldsymbol{\mu}(0) : \boldsymbol{\pi}), \tag{19}$$

where $\rho(M) > 0$ denotes the logarithmic Sobolev constant of M defined as

$$\rho(M) := \min_{\phi:\Omega\to\mathbb{R}}\frac{\mathcal{E}(\phi, \phi)}{\mathcal{L}(\phi)}$$

where

$$\mathcal{E}(\phi, \phi) = \frac{1}{2}\sum_{i,j\in\Omega}(\phi(i) - \phi(j))^2\pi_i M(i,j)$$

$$\mathcal{L}(\phi) = \sum_{i\in\Omega}\phi(i)^2\log\left(\frac{\phi(i)^2}{\sum_{j\in\Omega}\phi(j)^2\pi(j)}\right)\pi_i.$$

Therefore, from (18) and (19), it follows that

$$T_D(\varepsilon) \leq \frac{1}{4\rho(M)}\left(\log\log\frac{1}{\pi_{\min}} + \log\frac{1}{\varepsilon}\right)$$

$$T_{TV}(\varepsilon) \leq T_D(2\varepsilon^2) \leq \frac{1}{4\rho(M)}\left(\log\log\frac{1}{\pi_{\min}} + 2\log\frac{1}{\varepsilon}\right), \tag{20}$$

where $\pi_{\min} = \min_i \pi_i$; it can be verified that $D(\boldsymbol{\mu}(0) : \boldsymbol{\pi}) \leq \log\frac{1}{\pi_{\min}}$ for any $\mu(0)$.

4.2 Application to Modified Logit-Response

Recall the modified Logit-response learning mechanism with parameters α, β described in Section 3.1. It is a continuous-time Markov process $\{X_t\}_{t \in \mathbb{R}_+}$ with $X_t = x(\mathbf{s}(t))$ over the state space Ψ_n^s. Consider the following Markov chain M, which is essentially its *kernel*, over state space Ψ_n^s. From a current state $\mathbf{x} \in \Psi_n^s$, it transits to the next state $\mathbf{y}$ as follows:

○ Choose a strategy $i \in [s]$ uniformly at random.

○ If $\mathbf{x}_i > 0$, $\mathbf{y} = \mathbf{x} + \frac{1}{n}(e_j - e_i)$ with probability

$$\frac{e^{\beta \, \mathcal{P}\left(\mathbf{x} + \frac{1}{n}(e_j - e_i)\right)}}{\sum_{j \in [s]} e^{\beta \, \mathcal{P}\left(\mathbf{x} + \frac{1}{n}(e_j - e_i)\right)}},$$

where $\beta > 0$ is some (fixed) constant and e_i is a s-dimensional unit vector whose coordinate values are 0 except for the ith one.

○ Otherwise, $\mathbf{y} = \mathbf{x}$.

It can be verified that M is reversible and the stationary distribution $\boldsymbol{\pi}$ of M is

$$\pi_{\mathbf{x}} \quad \propto \quad e^{\beta \mathcal{P}(\mathbf{x})} \qquad \text{for} \quad \mathbf{x} \in \Psi_n^s. \tag{21}$$

More importantly, we relate M to the modified Logit-response described in Section 3.1. In the modified Logit-response, recall that the total clock-rate of players of that are playing strategy $i \in [s]$ is always αn for all $i \in [s]$. Hence, due to the memoryless property of Exponential distribution and the independence between clocks, the modified Logit-response is equivalent to having a global exponential clock of rate $\alpha s n$, and the transition happening according to M when the clock ticks. Let $\boldsymbol{\mu}(t)$ be the distribution of strategies $x(\mathbf{s}(t))$ under this modified Logit-response. Then, clearly

$$\boldsymbol{\mu}(t) \; = \; \boldsymbol{\mu}(0) \sum_{k=0}^{\infty} \Pr(\zeta = k) M^k \; = \; \boldsymbol{\mu}(0) e^{s \alpha n t (M - I)}, \tag{22}$$

where ζ is an exponential random variable with mean $s \alpha n t$. From this relation, Lemma 2 naturally follows since its stationary distribution has to be the same as that of M. Further, the mixing time of the modified Logit-respose can be obtained from (20) in term of $\rho(M)$ as follows:

$$T_{TV}(\varepsilon)$$
$$\leq \; \frac{1}{4 s \alpha n \rho(M)} \left(\log \log \frac{1}{\pi_{\min}} + 2 \log \frac{1}{\varepsilon} \right)$$
$$\overset{(a)}{\leq} \; \frac{1}{4 s \alpha n \rho(M)} \left(\log \log \left(|\Psi_n^s| e^{\beta} \right) + 2 \log \frac{1}{\varepsilon} \right)$$
$$\overset{(b)}{\leq} \; \frac{1}{4 s \alpha n \rho(M)} \left(\log \left((s-1) \log(n+1) + \beta \right) + 2 \log \frac{1}{\varepsilon} \right), \tag{23}$$

where (a) is from the characterization of $\boldsymbol{\pi}$ in Lemma 2 with $\mathcal{P}(\cdot) \in [0, 1]$ and (b) is due to $|\Psi_n^s| \leq (n+1)^{s-1}$. Therefore, the following lemma implies that $T_{TV}(\varepsilon) = O(n \log \log n + n \log 1/\varepsilon)$.

LEMMA 5. *If $\mathcal{P} : \Psi_n^s \to [0, 1]$, there exists a constant $c_0 = c_0(s)$ such that*

$$\rho(M) \; \geq \; \frac{c_0 e^{-3\beta}}{n^2}.$$

We note that the above bound is independent from the Lipschitz property of $\mathcal{P}$. The proof of Lemma 5 follows by somewhat direct adaptation of arguments in the paper by Frieze and Kannan [10]. The main difference is that they study continuous convex sets, while we consider lattice points in some simplex. At the first glance, this difference looks not significant since simplexes are also convex. But, one needs to be careful to deal with discrete objects (i.e. lattices) instead of continuous ones. The proof of Lemma 5 is omitted due to space constraints.

5. PROOF OF THEOREM 3

Now, we present proof of Theorem 3. We wish to establish that for choice of large enough β as per (8) and large enough time t as per (10), the aggregate state of strategy profile $\mathbf{x}(\mathbf{s}(t))$ is such that $\mathcal{P}(\mathbf{x}(\mathbf{s}(t)))$ is ε-close to $\max_{\mathbf{y} \in \Psi_n^s} \mathcal{P}(\mathbf{y})$, in expectation. This is established by using two key results. First, Lemma 6 implies that for β large enough as per (8), the expectation of the strategy profile with respect to the stationary distribution $\boldsymbol{\pi}$ of the modified Logit-response is $\varepsilon/2$-close to $\max_{\mathbf{y} \in \Psi_n^s} \mathcal{P}(\mathbf{y})$. Second, from (23) and Lemma 5 for t large enough as per (10), the distribution of strategy profile is $\varepsilon/2$-close to $\boldsymbol{\pi}$ starting from any initial state. Putting these together, the desired conclusion follows. To this end, we state and prove the lemma required for the first step.

LEMMA 6. *For β large enough so that*

$$\beta \geq \max \left\{ \frac{4(s-1)}{\varepsilon} \log 2s, \frac{4(s-1)}{\varepsilon} \log \frac{8s\lambda}{\varepsilon} \right\},$$

then

$$\mathbb{E}_{\boldsymbol{\pi}} \left[\mathcal{P}(\mathbf{x}) \right] \; \geq \; \sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) - \varepsilon/2.$$

In the above, $\mathcal{P}$ as in Theorem 3, $\boldsymbol{\pi}$ being the stationary distribution as defined in Lemma 2.

PROOF. We start by defining the following notations.

$$L_\beta = \mathbb{E}_{\boldsymbol{\pi}} \left[\mathcal{P}(\mathbf{x}) \right], \qquad C_\beta = \sum_{\mathbf{x} \in \Psi_n^s} e^{\beta \mathcal{P}(\mathbf{x})}$$

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}), \quad B(\mathbf{x}^*, \delta) = \{ \mathbf{x} \in \Psi_n^s : \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \delta \},$$

where $\delta \in [0, 1]$ is some small constant which will be decided later. The distribution $\boldsymbol{\pi}$ is of *exponential* form with its normalization constant C_β. Therefore, it can be verified (and very well known) that the derivative of $\log C_\beta$ with respect to exponential parameter β is the expectation L_β. Further, it is easy to observe that L_β is monotonically non-decreasing (increasing) in β. Therefore, by the standard Mean Value Theorem, it follows that

$$\begin{aligned} L_\beta \; &\geq \; \frac{1}{\beta} \left(\log C_\beta - \log C_0 \right) \\ &= \; \frac{1}{\beta} \log \frac{C_\beta}{|\Psi_n^s|} \\ &= \; \frac{1}{\beta} \log \frac{\sum_{\mathbf{x} \in \Phi_n^s} e^{\beta \mathcal{P}(\mathbf{x})}}{|\Psi_n^s|} \\ &= \; \mathcal{P}(\mathbf{x}^*) + \frac{1}{\beta} \log \frac{\sum_{\mathbf{x} \in \Psi_n^s} e^{\beta \left(\mathcal{P}(\mathbf{x}) - \mathcal{P}(\mathbf{x}^*) \right)}}{|\Psi_n^s|} \\ &\geq \; \mathcal{P}(\mathbf{x}^*) + \frac{1}{\beta} \log \frac{\sum_{\mathbf{x} \in B(\mathbf{x}^*, \delta)} e^{\beta \left(\mathcal{P}(\mathbf{x}) - \mathcal{P}(\mathbf{x}^*) \right)}}{|\Psi_n^s|} \end{aligned}$$

$$\overset{(a)}{\geq} \quad \mathcal{P}(\mathbf{x}^*) + \frac{1}{\beta} \log \frac{\sum_{\mathbf{x} \in B(\mathbf{x}^*, \delta)} e^{-\beta \delta \lambda}}{|\Psi_n^s|}$$

$$= \quad \mathcal{P}(\mathbf{x}^*) + \frac{1}{\beta} \log \frac{|B(\mathbf{x}^*, \delta)| e^{-\beta \delta \lambda}}{|\Psi_n^s|}$$

$$= \quad \mathcal{P}(\mathbf{x}^*) - \delta \lambda + \frac{1}{\beta} \log \frac{|B(\mathbf{x}^*, \delta)|}{|\Psi_n^s|},$$

where (a) is from the λ-Lipschitz property of $\mathcal{P}$ and the definition of $B(\mathbf{x}^*, \delta)$. Now $|B(\mathbf{x}^*, \delta)|$ and $|\Psi_n^s|$ can be bounded as follows:

$$|B(\mathbf{x}^*, \delta)| \geq \left(\frac{\delta(n+1)}{2s}\right)^{s-1} \qquad \text{and} \qquad |\Psi_n^s| \leq (n+1)^{s-1}.$$

Therefore, we have

$$\begin{aligned} L_\beta &\geq \quad \mathcal{P}(\mathbf{x}^*) - \delta \lambda + \frac{1}{\beta} \log \frac{|B(\mathbf{x}^*, \delta)|}{|\Psi_n^s|} \\ &\geq \quad \mathcal{P}(\mathbf{x}^*) - \delta \lambda + \frac{1}{\beta} \log \frac{\left(\frac{\delta(n+1)}{2s}\right)^{s-1}}{(n+1)^{s-1}} \\ &= \quad \mathcal{P}(\mathbf{x}^*) - \delta \lambda + \frac{s-1}{\beta} \log \frac{\delta}{2s}. \end{aligned} \qquad (24)$$

To complete the proof, we consider two cases: (i) $\lambda \leq \varepsilon/4$ and (ii) $\lambda > \varepsilon/4$. First, let us consider case (i). For this, we choose $\delta = 1$ and $\beta \geq \frac{4(s-1)}{\varepsilon} \log 2s$. Then, from (24),

$$\begin{aligned} L_\beta &\geq \quad \mathcal{P}(\mathbf{x}^*) - \delta \lambda + \frac{s-1}{\beta} \log \frac{\delta}{2s} \\ &= \quad \mathcal{P}(\mathbf{x}^*) - \lambda - \frac{s-1}{\beta} \log 2s \\ &\geq \quad \mathcal{P}(\mathbf{x}^*) - \varepsilon/4 - \varepsilon/4 \\ &= \quad \mathcal{P}(\mathbf{x}^*) - \varepsilon/2, \end{aligned}$$

where each step follows from choice of δ, β and the fact that in the case (i) we have $\lambda \leq \varepsilon/4$.

Now the case (ii), $\lambda > \varepsilon/4$. For this we choose $\delta = \frac{\varepsilon}{4\lambda}$. This is a valid choice since $\delta = \frac{\varepsilon}{4\lambda} < 1$ since we have $\lambda > \varepsilon/4$. Consider $\beta \geq \frac{4(s-1)}{\varepsilon} \log \frac{8s\lambda}{\varepsilon}$. Then

$$L_\beta \geq \mathcal{P}(\mathbf{x}^*) - \varepsilon/2.$$

In summary, for both cases (i) and (ii), the desired conclusion follows as long as we have β large enough so that

$$\beta \geq \max \left\{ \frac{4(s-1)}{\varepsilon} \log 2s, \frac{4(s-1)}{\varepsilon} \log \frac{8s\lambda}{\varepsilon} \right\}.$$

This completes the proof of Lemma 6. $\square$

Completing the proof of Theorem 3. As before, let $\boldsymbol{\mu}(t)$ denote the distribution of strategies $x(\mathbf{s}(t))$ under the modified Logit-response mechanism with parameters α, β as described in Section 3.1. From (23) and Lemma 5, if

$$\begin{aligned} t &\geq \quad \frac{ne^{3\beta}}{\alpha c} \left(\log \log n + \log \beta + \log \frac{1}{\varepsilon} \right) \qquad (25) \\ &\overset{(a)}{\geq} \quad \frac{n}{4s\alpha c_0 e^{-3\beta}} \left(\log \left((s-1) \log(n+1) + \beta \right) + 2 \log \frac{2}{\varepsilon} \right), \end{aligned}$$

then

$$\|\boldsymbol{\mu}(t) - \boldsymbol{\pi}\|_{TV} \leq \varepsilon/2.$$

In above, for (a), one can find an appropriate constant c which depends on s. Therefore, we have the desired bound

for t larger than (25):

$$\begin{aligned} \mathbb{E}\left[\mathcal{P}(x(\mathbf{s}(t)))\right] &= \quad \mathbb{E}_{\boldsymbol{\mu}(t)}\left[\mathcal{P}(\mathbf{x})\right] \\ &\geq \quad \mathbb{E}_{\boldsymbol{\pi}(t)}\left[\mathcal{P}(\mathbf{x})\right] - \|\boldsymbol{\mu}(t) - \boldsymbol{\pi}(t)\|_{TV} \cdot \sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) \\ &\overset{(a)}{\geq} \quad \sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) - \varepsilon/2 - \varepsilon/2 \\ &= \quad \sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) - \varepsilon, \end{aligned}$$

where (a) is from Lemma 6 and $\sup_{\mathbf{x} \in \Psi_n^s} \mathcal{P}(\mathbf{x}) \leq 1$. This completes the proof of Theorem 3.

6. PROOF OF THEOREM 4

We wish to establish that under condition that time t satisfies (15), rationality parameter of modified Logit-response β satisfies (16) and dynamics rate Λ satisfies (17), the ε-predictability of modified Logit-response holds for all time t as per (14).

Some formalism. To this end, we start by noting that the underlying state, the strategy profile $x(\mathbf{s}(t))$ of players, has time varying state space. Specifically, it is $\Psi_{n(t)}^s$ which changes with $n(t)$. To address the associated formalism, we introduce the natural 'projection' operator. Now since $\Lambda > 0$, it must be that at each time t one the three things can happen: no change in $n(t)$ (i.e. $n(t) = n(t^-)$); one new player joins (i.e. $n(t) = n(t^-) + 1$); or one existing player leaves (i.e. $n(t) = n(t^-) - 1$). For each of these three cases, we associate state $x(\mathbf{s}(t))$ with $x(\mathbf{s}(t^-))$ through the 'projection' operation $[\cdot]_t : \Psi_{n(t^-)}^s \to \Psi_{n(t)}^s$ as follows:

1. No change, i.e. $n(t^-) = n(t)$. Then, $[\mathbf{x}]_t = \mathbf{x}$ for any $\mathbf{x} \in \Psi_{n(t^-)}^s$.

2. A new player joins with initial strategy $i \in [s]$, i.e. $n(t) = n(t^-) + 1$ and $s_{n(t)} = i \in [s]$. Then, for any $\mathbf{x} \in \Psi_{n(t^-)}^s$, $[\mathbf{x}]_t = (n(t^-)\mathbf{x} + e_i)/n(t)$. That is, for $\mathbf{y} = [\mathbf{x}]_t$

$$y_j = \begin{cases} n(t^-)x_j/n(t) & \text{for } j \neq i \\ (n(t^-)x_j + 1)/n(t) & \text{for } j = i. \end{cases}$$

3. An existing player with strategy $i \in [s]$ departs, i.e. $n(t) = n(t^-) - 1$. Then, for any $\mathbf{x} \in \Psi_{n(t^-)}^s$, $[\mathbf{x}]_t = (n(t^-)\mathbf{x} - e_i)/n(t)$. That is, for $\mathbf{y} = [\mathbf{x}]_t$

$$y_j = \begin{cases} n(t^-)x_j/n(t) & \text{for } j \neq i \\ (n(t^-)x_j - 1)/n(t) & \text{for } j = i. \end{cases}$$

Let $\boldsymbol{\mu}(t) = [\mu(t)_{\mathbf{x}}]$ be the distribution over the space $\Psi_{n(t)}^s$ of strategy profile of players under the modified Logit-response at time t. Then as per the above notation,

$$\mu(t)_{[\mathbf{x}]_t} = \mu(t^-)_{\mathbf{x}}, \ \forall \ \mathbf{x} \in \Psi_{n(t^-)}^s.$$

Therefore, with an abuse of notation, we shall use

$$\boldsymbol{\mu}(t) = \left[\boldsymbol{\mu}(t^-)\right]_t. \qquad (26)$$

Proof of Theorem 4. For $t \geq 0$, let $M(t)$ denotes the Markov chain M associated with the modified Logit-response over

state space $\Psi^s_{n(t)}$ as described in Section 4.2. Its stationary distribution be denoted by $\boldsymbol{\pi}(t)$ which has the form

$$\pi(t)_{\mathbf{x}} \quad \propto \quad e^{\beta \mathcal{P}(\mathbf{x})} \qquad \text{for} \quad \mathbf{x} \in \Psi^s_{n(t)}. \tag{27}$$

Let the sequence $\{t_0 < t_1 < t_2 < t_3 < \dots\}$ be the collection of times when possibly changes in $n(t)$ happens. Without loss of generality, we shall assume that for $m \geq 0$,

$$t_{m+1} - t_m \ \in [\Lambda, 2\Lambda] \qquad \text{and} \quad t_0 \leq \Lambda. \tag{28}$$

This is because (12) implies $t_{m+1} - t_m \geq \Lambda$ and for $t_{m+1} - t_m \leq 2\Lambda$, we can insert additional times with $n(t) = n(t^-)$ for those particular time instances. Let $\boldsymbol{\mu}(t_m)$ is the distribution over $\Psi^s_{n(t_m)}$, just after the 'change' at time t_m. Let $\boldsymbol{\pi}(t_m)$ be the stationary distribution on $\Psi_{n(t_m)}$ corresponding to thus 'changed' system. We wish to study $D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m))$. Specifically, we will establish the following.

LEMMA 7. *For any $m \geq 0$, with t_m satisfying (28),*

$$D(\boldsymbol{\mu}(t_{m+1}) : \boldsymbol{\pi}(t_{m+1}))$$
$$\leq \quad \left(1 - \frac{A_1}{n(t_m)}\right) D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) + \frac{A_2}{n(t_m)}, \quad (29)$$

where $A_1 = 2s\alpha c_0 e^{-3\beta}\Lambda$ and $A_2 = 6\beta\lambda + e^\beta(s-1)$.

Using (29), we first complete the proof of Theorem 4 and then present the proof of Lemma 7. To this end, we claim that (29) implies that for any $m \geq m_0$,

$$D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) \quad \leq \quad \varepsilon^2/2, \tag{30}$$

with

$$m_0 \quad = \quad \Omega\left(\frac{n(0)}{A_1 \varepsilon^2} \left(s \log n(0) + \beta\right)\right)$$
$$= \quad \frac{n(0)}{\alpha c_1 e^{-3\beta}\Lambda\varepsilon^2} \left(\log n(0) + \beta\right). \tag{31}$$

The (30) essentially completes the proof of Theorem 4. This is because, for any $t \in [t_m, t_{m+1})$ since there is no change, we have

$$D(\boldsymbol{\mu}(t) : \boldsymbol{\pi}(t)) \quad \leq \quad e^{-4s\alpha n(t_m)\rho(M(t_m))(t-t_m)} D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m))$$
$$\leq \quad D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)),$$

where the first inequality follows from (19) and $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(t_m)$. Therefore, from (30) it follows that for $t \geq t_{m_0}$,

$$D(\boldsymbol{\mu}(t) : \boldsymbol{\pi}(t)) \leq \varepsilon^2/2.$$

And hence using relation (18) between entropy distance and total variation distance, we have

$$\|\boldsymbol{\mu}(t) - \boldsymbol{\pi}(t)\|_{TV} \leq \varepsilon/2.$$

Additionally, we can choose m_0 from (31) such that

$$t_{m_0} \quad \leq \quad 2\Lambda m_0 = \frac{2n(0)}{\alpha c_1 e^{-3\beta}\varepsilon^2} \left(\log n(0) + \beta\right),$$

where the first inequality is due to (28). In summary,

$$\|\boldsymbol{\mu}(t) - \boldsymbol{\pi}(t)\|_{TV} \leq \varepsilon/2, \qquad \forall \, t \geq \frac{2n(0)}{\alpha c_1 e^{-3\beta}\varepsilon^2} \left(\log n(0) + \beta\right).$$

From this and use of identical arguments as in the last part of the proof of Theorem 3, the desired statement of Theorem 4 follows.

Now we justify (30) using the Lemma 7. To this end, first observe that suppose for any m_0,

$$D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0})) \leq \varepsilon^2/2. \tag{32}$$

Then, from (29) of Lemma 7,

$$D(\boldsymbol{\mu}(t_{m_0+1}) : \boldsymbol{\pi}(t_{m_0+1}))$$
$$\leq \quad \left(1 - \frac{A_1}{n(t_{m_0})}\right) D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0})) + \frac{A_2}{n(t_{m_0})}$$
$$\leq \quad \left(1 - \frac{A_1}{n(t_{m_0})}\right) \varepsilon^2/2 + \frac{A_2}{n(t_{m_0})}$$
$$\leq \quad \varepsilon^2/2,$$

where the last inequality follows from condition on Λ as per (17), since

$$\frac{A_2}{n(t_m)} \quad = \quad \frac{6\beta\lambda + e^\beta(s-1)}{n(t_m)}$$
$$\leq \quad \frac{s\alpha c_0 e^{-3\beta}\Lambda\varepsilon^2/2}{n(t_m)} \quad = \quad \frac{A_1\varepsilon^2}{4n(t_m)}. \tag{33}$$

Therefore, for $m \geq m_0$,

$$D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) \leq \varepsilon^2/2. \tag{34}$$

Hence, it suffices to show that there exists m_0 satisfying (31) so that $D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0})) \leq \varepsilon^2/2$. To this end, suppose that for a given m_0 for all $m \leq m_0$,

$$D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) > \varepsilon^2/2. \tag{35}$$

Then, from (29),

$$D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0}))$$
$$\leq \quad \left(1 - \frac{A_1}{n(t_{m_0-1})}\right) D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0-1})) + \frac{A_2}{n(t_{m_0-1})}$$

$$< \quad \left(1 - \frac{A_1}{n(t_{m_0-1})}\right) D(\boldsymbol{\mu}(t_{m_0-1}) : \boldsymbol{\pi}(t_{m_0-1}))$$
$$+ \frac{A_2}{n(t_{m_0-1})} \times \frac{D(\boldsymbol{\mu}(t_{m_0-1}) : \boldsymbol{\pi}(t_{m_0-1}))}{\varepsilon^2/2}$$
$$\overset{(a)}{\leq} \quad \left(1 - \frac{A_1}{2n(t_{m_0-1})}\right) D(\boldsymbol{\mu}(t_{m_0-1}) : \boldsymbol{\pi}(t_{m_0-1}))$$
$$\leq \quad D(\boldsymbol{\mu}(t_0) : \boldsymbol{\pi}(t_0)) \prod_{m=0}^{m_0-1} \left(1 - \frac{A_1}{2n(t_m)}\right)$$
$$\leq \quad D(\boldsymbol{\mu}(t_0) : \boldsymbol{\pi}(t_0)) \prod_{m=0}^{m_0-1} \left(1 - \frac{A_1}{2(n(0)+m)}\right)$$
$$\overset{(b)}{\leq} \quad ((s-1)\log(n(0)+2) + \beta) \prod_{m=0}^{m_0-1} \left(1 - \frac{A_1}{2(n(0)+m)}\right),$$

where (a) follows by using (33); (b) is due to

$$D(\boldsymbol{\mu}(t_0) : \boldsymbol{\pi}(t_0)) \quad \leq \quad \log \frac{1}{\pi(t_0)_{\min}}$$
$$\leq \quad (s-1)\log(n(t_0)+1) + \beta$$
$$\leq \quad (s-1)\log(n(0)+2) + \beta,$$

as we discussed in (23). Therefore, it follows immediately

that if

$$m_0 \geq \Omega\left(\frac{n(0)}{A_1\varepsilon^2}\left(s\log n(0) + \beta\right)\right)$$
$$= \frac{n(0)}{\alpha c_1 e^{-3\beta}\Lambda\varepsilon^2}\left(\log n(0) + \beta\right),$$

then we have

$$D(\boldsymbol{\mu}(t_{m_0}) : \boldsymbol{\pi}(t_{m_0})) < \varepsilon^2/2.$$

This completes the justification of (30) based on Lemma 7 and hence proof of Theorem 4.

6.1 Proof of Lemma 7

To simplify notations, define

$$\widehat{\boldsymbol{\mu}}(t_{m+1}^-) := \left[\boldsymbol{\mu}(t_{m+1}^-)\right]_{t_{m+1}}, \qquad \widehat{\boldsymbol{\pi}}(t_m) := \left[\boldsymbol{\pi}(t_m)\right]_{t_{m+1}}.$$

Note that $\boldsymbol{\mu}(t_{m+1}) = \widehat{\boldsymbol{\mu}}(t_{m+1}^-)$ from (26), but $\widehat{\boldsymbol{\pi}}(t_m) \neq \boldsymbol{\pi}(t_{m+1})$; $\widehat{\boldsymbol{\mu}}(t_{m+1}^-)$ is absolutely continuous with respect to $\widehat{\boldsymbol{\pi}}(t_m)$; and as a distribution $\widehat{\boldsymbol{\mu}}(t_{m+1}^-)$ (resp. $\widehat{\boldsymbol{\pi}}(t_m)$) is the same as $\boldsymbol{\mu}(t_{m+1}^-)$ (resp. $\boldsymbol{\pi}(t_{m+1}^-) = \boldsymbol{\pi}(t_m)$). With these observations, we have

$$D(\boldsymbol{\mu}(t_{m+1}) : \boldsymbol{\pi}(t_{m+1}))$$
$$= D(\widehat{\boldsymbol{\mu}}(t_{m+1}^-) : \boldsymbol{\pi}(t_{m+1}))$$
$$= \sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\mu}(t_{m+1}^-)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}$$
$$= \sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \left(\log \frac{\widehat{\mu}(t_{m+1}^-)_{\mathbf{x}}}{\widehat{\pi}(t_m)_{\mathbf{x}}} + \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}\right)$$
$$= D(\widehat{\boldsymbol{\mu}}(t_{m+1}^-) : \widehat{\boldsymbol{\pi}}(t_m)) + \sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}$$
$$= D(\boldsymbol{\mu}(t_{m+1}^-) : \boldsymbol{\pi}(t_m)) + \sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}.$$
$$(36)$$

For the first term in (36), using (19), (22) and Lemma 5, we obtain

$$D(\boldsymbol{\mu}(t_{m+1}^-) : \boldsymbol{\pi}(t_m))$$
$$\leq \exp\left(-4s\alpha n(t_m)\rho(M(t_m))(t_{m+1} - t_m)\right) D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m))$$
$$\leq \exp\left(-\frac{4s\alpha c_0 e^{-3\beta}\Lambda}{n(t_m)}\right) D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)), \qquad (37)$$

where the last inequality is from $t_{m+1} \geq t_m + \Lambda$ in (28). For the second term in (36), we state the following.

LEMMA 8. *For any* $m \geq 0$,

$$\sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}} \leq \frac{6\beta\lambda + e^\beta(s-1)}{n(t_m)}.$$
$$(38)$$

Before we prove Lemma 8, we complete the proof of Lemma 7 using it. To that end, using the bounds (37) and (38) in

(36), we obtain

$$D(\boldsymbol{\mu}(t_{m+1}) : \boldsymbol{\pi}(t_{m+1}))$$
$$\leq \exp\left(-\frac{4s\alpha c_0 e^{-3\beta}\Lambda}{n(t_m)}\right) D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) + \frac{6\beta\lambda + e^\beta(s-1)}{n(t_m)}$$
$$\stackrel{(a)}{\leq} \left(1 - \frac{A_1}{n(t_m)}\right) D(\boldsymbol{\mu}(t_m) : \boldsymbol{\pi}(t_m)) + \frac{A_2}{n(t_m)}, \qquad (39)$$

where $A_1 := 2s\alpha c_0 e^{-3\beta}\Lambda$ and $A_2 := 6\beta\lambda + e^\beta(s-1)$. In above, we have used that $x = \frac{2A_1}{n(t_m)} \leq 1$ from (13), and $e^{-x} \leq 1 - x/2$ for $x \in [0,1]$. This completes the proof of Lemma 7.

6.2 Proof of Lemma 8

There are three possible scenarios at time t_{m+1} as discussed while defining the projection operator $[\cdot]$: (i) no change, (ii) one player joins with some strategy, say $i \in [s]$, or (iii) an existing player playing some strategy, say $i \in [s]$, leaves.

The case (i) is trivial due to no change.

Next, we consider case (ii). To this end, let $n = n(t_{m+1}) = n(t_m) + 1 = n(t_{m+1}^-) + 1$. Note that

$$\sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}} \leq \max_{\mathbf{x} \in \Psi_n^s : x_i > 0} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}},$$
$$(40)$$

since $\widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} = 0$ if $x_i = 0$ as the new player starts with strategy i. Recall definitions of $\widehat{\boldsymbol{\pi}}(t_m)$ and $\boldsymbol{\pi}(t_{m+1})$: for $\mathbf{x} \in \Psi_n^s$,

$$\pi(t_{m+1})_{\mathbf{x}} = \frac{1}{C_1} e^{\beta\mathcal{P}(\mathbf{x})}$$
$$\widehat{\pi}(t_m)_{\mathbf{x}} = \begin{cases} \frac{1}{C_2} e^{\beta\mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)} & \text{if } x_i > 0 \\ 0 & \text{otherwise} \end{cases},$$

where

$$C_1 := \sum_{\mathbf{x} \in \Psi_n^s} e^{\beta\mathcal{P}(\mathbf{x})}, \qquad C_2 := \sum_{\mathbf{x} \in \Psi_n^s : x_i > 0} e^{\beta\mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}.$$

Thus, from (40) we have

$$\sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}$$
$$\leq \max_{\mathbf{x} \in \Psi_n^s : x_i > 0} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}}$$
$$= \max_{\mathbf{x} \in \Psi_n^s : x_i > 0} \log \frac{\frac{1}{C_2} e^{\beta\mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}}{\frac{1}{C_1} e^{\beta\mathcal{P}(\mathbf{x})}}$$
$$= \log \frac{C_1}{C_2} + \max_{\mathbf{x} \in \Psi_n^s : x_i > 0} \beta\left(\mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right) - \mathcal{P}(\mathbf{x})\right)$$
$$\leq \log \frac{C_1}{C_2} + \frac{2\beta\lambda}{n-1}, \qquad (41)$$

where the last inequality is from the λ-Lipschitz property of $\mathcal{P}$. To derive the conclusion, it suffices to bound C_1/C_2 in

116

(41), which we obtain as follows:

$$
\begin{aligned}
\frac{C_1}{C_2} &= \frac{\sum_{\mathbf{x} \in \Psi_n^s} e^{\beta \mathcal{P}(\mathbf{x})}}{\sum_{\mathbf{x} \in \Psi_n^s : x_i > 0} e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}} \\
&= \frac{\sum_{\mathbf{x} \in \Psi_n^s : x_i > 0} e^{\beta \mathcal{P}(\mathbf{x})}}{\sum_{\mathbf{x} \in \Psi_n^s : x_i > 0} e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}} + \frac{\sum_{\mathbf{x} \in \Psi_n^s : x_i = 0} e^{\beta \mathcal{P}(\mathbf{x})}}{\sum_{\mathbf{x} \in \Psi_n^s : x_i > 0} e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}} \\
&\overset{(o)}{\leq} \max_{\mathbf{x} \in \Psi_n^s : x_i > 0} \frac{e^{\beta \mathcal{P}(\mathbf{x})}}{e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} - e_i}{n-1}\right)}} + \frac{e^{\beta} \left| \{ \mathbf{x} \in \Psi_n^s : x_i = 0 \} \right|}{\left| \{ \mathbf{x} \in \Psi_n^s : x_i > 0 \} \right|} \\
&\overset{(a)}{\leq} e^{\frac{2\beta}{n-1}} + e^{\beta} \frac{\left| \{ \mathbf{x} \in \Psi_n^s : x_i = 0 \} \right|}{\left| \Psi_n^s \right| - \left| \{ \mathbf{x} \in \Psi_n^s : x_i = 0 \} \right|} \\
&\leq e^{\frac{2\beta\lambda}{n-1}} + e^{\beta} \frac{\left| \Psi_n^{s-1} \right|}{\left| \Psi_n^s \right| - \left| \Psi_n^{s-1} \right|} \\
&\overset{(b)}{=} e^{\frac{2\beta\lambda}{n-1}} + e^{\beta} \frac{s-1}{n} \\
&\leq 1 + \frac{4\beta\lambda}{n-1} + e^{\beta} \frac{s-1}{n}. \quad (42)
\end{aligned}
$$

In above, (o) follows using $\mathcal{P}(\cdot) \in [0,1]$; (a) follows from the λ-Lipschitz property of $\mathcal{P}$; and (b) follows from $|\Psi_n^s| = \binom{n+s-1}{n}$. For the last inequality, we use $e^x \leq 1 + 2x$ for $x \in [0,1]$ and the condition (13) of $n - 1 = n(t_m)$ that $n(t_m) \geq 2\beta\lambda$. Therefore, from (41) and (42), we obtain the conclusion:

$$
\begin{aligned}
\sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} & \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}} \\
&\leq \log \frac{C_1}{C_2} + \frac{2\beta\lambda}{n-1} \\
&\leq \log \left(1 + \frac{4\beta\lambda}{n-1} + e^{\beta} \frac{s-1}{n} \right) + \frac{2\beta\lambda}{n-1} \\
&\overset{(a)}{\leq} \frac{4\beta\lambda}{n-1} + e^{\beta} \frac{s-1}{n} + \frac{2\beta\lambda}{n-1} \\
&\leq \left(6\beta\lambda + e^{\beta}(s-1) \right) \frac{1}{n-1} \\
&= \frac{6\beta\lambda + e^{\beta}(s-1)}{n(t_m)},
\end{aligned}
$$

where (a) follows from the fact that $\log(1+x) \leq x$ for $x \geq 0$. This completes the proof of Lemma 8 for case (ii).

Finally, consider case (iii) i.e. $n = n(t_{m+1}) = n(t_m) - 1 = n(t_{m+1}^-) - 1$. In this case, $\widehat{\boldsymbol{\pi}}(t_m)$ and $\boldsymbol{\pi}(t_{m+1})$ has the following form: for $\mathbf{x} \in \Psi_n^s$,

$$
\pi(t_{m+1})_{\mathbf{x}} = \frac{1}{C_1} e^{\beta \mathcal{P}(\mathbf{x})}, \qquad \widehat{\pi}(t_m)_{\mathbf{x}} = \frac{1}{C_2} e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} + e_i}{n+1}\right)},
$$

where

$$
C_1 := \sum_{\mathbf{x} \in \Psi_n^s} e^{\beta \mathcal{P}(\mathbf{x})}, \qquad C_2 := \sum_{\mathbf{x} \in \Psi_n^s} e^{\beta \mathcal{P}\left(\frac{n\mathbf{x} + e_i}{n+1}\right)}.
$$

Using the similar arguments as (41) and (42)[4], we obtain

$$
\begin{aligned}
\sum_{\mathbf{x} \in \Psi_{n(t_{m+1})}^s} \widehat{\mu}(t_{m+1}^-)_{\mathbf{x}} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}} &\leq \max_{\mathbf{x} \in \Psi_n^s} \log \frac{\widehat{\pi}(t_m)_{\mathbf{x}}}{\pi(t_{m+1})_{\mathbf{x}}} \\
&\leq \frac{4\beta\lambda}{n+1} = \frac{4\beta\lambda}{n(t_m)},
\end{aligned}
$$

which implies the Lemma 8 for case (iii). This concludes the proof of Lemma 8.

7. CONCLUSION

In this paper, we studied transient properties of a simple learning mechanism on the context of symmetric potential games, which include the congestion game. We obtain a precise relation between the performance error and the rate of anarchical dynamics in the number of players, which shows that the dynamic price of anarchy is small in the congestion game. Our novel techniques to analyze "space-varying" Markov processes using the entropy distance and logarithmic Sobolev constants were crucial for obtaining the desired results. We believe that the method of this paper should be of broad interest in understanding the reliability and controllability of dynamical systems.

8. REFERENCES

[1] C. Alos-Ferrer and N. Netzer. The logit-response dynamics. TWI Research Paper Series 28, Thurgauer Wirtschaftsinstitut, Universita"t Konstanz, 2008.

[2] M. Beckmann, C. B. McGuire, and C. B. Winston. *Studies in the Economics of Transportation.* Yale University Press, 1956.

[3] J. Bergine and B. Lipman. Evolution with state-dependent mutations. *Econometrica*, 64:943–956, 1996.

[4] L. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387–424, 1993.

[5] L. Blume. Population games. Game theory and information, EconWPA, July 1996.

[6] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. On the inefficiency of equilibria in congestion games. *Lecture Notes in Computer Science*, 3509:167–181, 2005.

[7] A. Cournot. Recherches sur les principes mathematiques de la theories des richesse. *Paris:Hachette*, 1, 1838.

[8] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, 1996.

[9] G. Facchini, F. van Megen, P. Borm, and S. Tijs. Congestion models and weighted Bayesian potential games. *Theory and Decision*, 42(2):193–206, 1997.

[10] A. Frieze and R. Kannan. Log-sobolev inequalities and sampling from log-concave distributions. *Annals of Applied Probability*, 9:14–26, 1998.

[11] R. Johari and J. N. Tsitsiklis. Efficiency loss in a network resource allocation game. *Mathematics of Operations Research*, 29(3):407–435, 2004.

---

[4]This case is a lot easier to analyze compared to case (ii) since C_1 and C_2 are summation of quantities over the same space.

[12] M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

[13] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, March 1998.

[14] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *In Proc. 16th STACS*, 1999.

[15] A. D. Wentzell M. I. Freidlin. *Random perturbations of dynamical systems*. Springer, 1984.

[16] J. R. Marden and J. S. Shamma. Revisiting log-linear learning: asynchrony, completeness, and payoff-based implementation. *Submitted for journal publication*, September 2008.

[17] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[18] R. Montenegro and P. Tetali. *Mathematical aspects of mixing times in Markov chains*. Now Pub, 2006.

[19] T. Roughgarden. *Selfish routing and the price of anarchy*. The MIT Press, 2005.

[20] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):259, 2002.

[21] M. E. Slade. What does an oligopoly maximize? *Journal of Industrial Economics*, 42(1):45–61, 1994.

[22] P. Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.