ASYMPTOTIC NASH EQUILIBRIA OF FINITE-STATE ERGODIC MARKOVIAN MEAN FIELD GAMES

ASAF COHEN AND ETHAN ZELL

ABSTRACT. Mean field games (MFGs) model equilibria in games with a continuum of weakly interacting players as limiting systems of symmetric n-player games. We consider the finite-state, infinite-horizon problem with ergodic cost. Assuming Markovian strategies, we first prove that any solution to the MFG system gives rise to a $(C/\sqrt{n})$-Nash equilibrium in the n-player game. We follow this result by proving the same is true for the strategy profile derived from the master equation. We conclude the main theoretical portion of the paper by establishing a large deviation principle for empirical measures associated with the asymptotic Nash equilibria. Then, we contrast the asymptotic Nash equilibria using an example. We solve the MFG system directly and numerically solve the ergodic master equation by adapting the deep Galerkin method of Sirignano and Spiliopoulos [33]. We use these results to derive the strategies of the asymptotic Nash equilibria and compare them. Finally, we derive an explicit form for the rate functions in dimension two.

1. INTRODUCTION

1.1. **Focus of the paper.** We consider a finite-state, ergodic mean field game (MFG) to model a large interacting system of symmetric agents using Markovian strategies. In such games, a *representative player* moves according to her jump process on a finite state-space $[d] := \{1, \ldots, d\}$ while subject to an ergodic cost that depends on a flow of measures. The representative player chooses a control to minimize her cost, and the control she chooses prescribes her transition rate between states, and hence the distributional law of her state-dynamics as well. Her distributional law, in turn, is *another* flow of measures over the states $[d]$. When the representative player's law matches the original measure flow she responded to, we refer to the control and the flow as a *mean field equilibrium* (MFE). The representative player's ergodic cost is given by:

$$J_0(\alpha, \mu) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \Big[\int_0^T \{ f(\mathcal{X}_t^\alpha, \alpha(t, \mathcal{X}_t^\alpha)) + F(\mathcal{X}_t^\alpha, \mu(t)) \} \, dt \Big], \tag{1.1}$$

where $(\mathcal{X}_t^\alpha)_{t \geq 0}$ is the position of the representative player using her chosen control α, $f + F$ is the running cost, and $(\mu(t))_{t \geq 0}$ is a flow of distributions over $[d]$. In these terms, we have an MFE when: (1) the distributional law of the player under control α, $\mathcal{L}(\mathcal{X}_t^\alpha)$, equals $\mu(t)$ and (2) α minimizes $J_0(\cdot, \mu)$.

In a typical n-player game, by contrast, players react to the dynamics of others, and when no player has a cost incentive to change their strategy, we say the profile of all strategies constitutes a *Nash equilibrium*. In a *symmetric* Nash equilibrium, a given player's strategy will mirror that of any other player, rendering players indistinguishable from a population perspective. While explicitly solving for Nash equilibria in finite-player games can be computationally intractable, the relationship between MFGs and n-player games provides one avenue to approximate Nash equilibria for large, finite-player games. Specifically, when the number of players n is large, we expect that the dynamics of the representative player in the MFG are close to the dynamics of any indistinguishable player from the n-player game. Denote the set of all players $[n] := \{1, \ldots, n\}$. The cost to player $i \in [n]$ is:

$$J_0^{n,i}(\boldsymbol{x}, \boldsymbol{\alpha}) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \Big[\int_0^T \Big\{ f(X_t^{\boldsymbol{\alpha},i}, \alpha^i(t, \boldsymbol{X}_t^{\boldsymbol{\alpha}})) + F(X_t^{\boldsymbol{\alpha},i}, \mu_t^{\boldsymbol{\alpha},i}) \Big\} \, dt \Big], \tag{1.2}$$

1

where now $\boldsymbol{\alpha} := (\alpha^1, \ldots, \alpha^n)$ is the collection of strategies for all players, $(\boldsymbol{X}_t^{\boldsymbol{\alpha}} := (X_t^{\boldsymbol{\alpha},1}, \ldots, X_t^{\boldsymbol{\alpha},n}))_{t \geq 0}$ are the players' positions, $(\mu_t^{\boldsymbol{\alpha},i})_{t \geq 0}$ is the empirical distribution of all players $\boldsymbol{X}^{\boldsymbol{\alpha}}$ without the i-th player, and $\boldsymbol{x}$ is the initial state of the players, $\boldsymbol{X}_0^{\boldsymbol{\alpha}} = \boldsymbol{x} \in [d]^n$. With respect to a Nash equilibrium, note that the i-th player minimizing $J_0^{n,i}$ can only select her own strategy α^i.

One way to characterize the MFG and n-player relationship then, and a focus of this paper, is whether the controls from the MFE result in an approximate Nash equilibrium when applied in the finite-player game. That is, when all n players are using the same strategies derived by the MFE, a single player who switches strategies receives, at most, a marginal benefit that vanishes as the number of players increases. Since the potential benefit of switching strategies decreases as the number of players increases, we refer to such strategy profiles as *asymptotic Nash equilibria*.

Typically, an ergodic MFE is characterized by a *MFG system* that is, in general, a coupled system of differential equations. One equation is Kolmogorov's equation that, evolving forward in time, describes the distribution of the representative player under the MFE's control. The second differential equation is a Bellman equation that includes the *value of the MFG* (optimal cost) and a *potential function*, that characterizes the state-wise deviations from the value. The *time-dependent ergodic MFG system* is written:

$$
\begin{aligned}
&-\frac{d}{dt}\check{u}(t,x) + \check{\varrho} = H(x, \Delta_x \check{u}(t, \cdot)) + F(x, \check{\mu}(t)), \\
&\frac{d}{dt}\check{\mu}(t,x) = \sum_{y \in [d]} \check{\mu}(t,y) \gamma_x^*(y, \Delta_y \check{u}(t, \cdot)), \\
&\check{\mu}(0) = \eta, \quad t \in \mathbb{R}_+ := (0, \infty), \quad x \in [d],
\end{aligned}
\tag{1.3}
$$

where $\mathcal{P}([d])$ is the set of probability distributions over $[d]$, H is the Hamiltonian, Δ_x is the finite difference operator $\Delta_x b := (b_y - b_x)_{y \in [d]}$, and $\gamma_x^*(y, \Delta_y \check{u}(t, \cdot))$ is the minimizer of the Hamiltonian, representing the rate to move from state x to state y. The running cost component f is part of the Hamiltonian H. A solution to (1.3) is a triple $(\check{\varrho}, \check{u}, \check{\mu})$ where $\check{\varrho} \in \mathbb{R}$ is the value, the potential function is $\check{u} : \mathbb{R}_+ \times [d] \to \mathbb{R}$, and the equilibrium measure flow is $\check{\mu} : \mathbb{R}_+ \to \mathcal{P}([d])$. *Any solution to the MFG system is linked to an MFE*: when the representative player uses the rates $[\gamma_y^*(x, \Delta_x \check{u}(t, \cdot))]_{x,y \in [d]}$ from the MFG system, then: (1) the controls and $\check{\mu}$ constitute an MFE, (2) the player pays the cost $\check{\varrho}$, and (3) the representative player's law equals $\check{\mu}$.

Our ergodic systems consider a long-time average cost and so admit a stationary system whose optimal control, measure flow, and potential function are all independent of time. The *stationary ergodic MFG system*, is:

$$
\begin{aligned}
&\bar{\varrho} = H(x, \Delta_x \bar{u}) + F(x, \bar{\mu}), \\
&0 = \sum_{y \in [d]} \bar{\mu}_y \gamma_y^*(x, \Delta_x \bar{u}), \quad x \in [d],
\end{aligned}
\tag{1.4}
$$

where any solution is a triple $(\bar{\varrho}, \bar{u}, \bar{\mu}) \in \mathbb{R} \times \mathbb{R}^d \times \mathcal{P}([d])$. When the initial data from (1.3), $\eta = \bar{\mu}$, we note that (1.4) is a special case of (1.3).

The MFG system is usefully characterized by a single PDE, the master equation. On one hand, the MFG system fixes the initial distribution of the players beforehand and then solves the system to find the MFE evolving from that distribution. More advantageously, the master equation solves for every initial distribution simultaneously. The master equation is particularly useful to study the convergence problem of (exact) Nash equilibria to the MFE. In this paper, we take the reversed direction, aiming to construct asymptotic Nash equilibria in the large population game from the MFE. For $(x, \eta) \in [d] \times \mathcal{P}([d])$, the finite-state *ergodic master equation* was studied in [12] and is given by:

$$
\varrho = H(x, \Delta_x U_0(\cdot, \eta)) + F(x, \eta) + \sum_{y,z \in [d]} \eta_y D_{yz}^\eta U_0(x, \eta) \gamma_z^*(y, \Delta_y U_0(\cdot, \eta)).
\tag{1.5}
$$

A solution to (1.5) is a pair (ϱ, U_0) where $\varrho \in \mathbb{R}$ is the *value of the MFG* and $U_0 : [d] \times \mathcal{P}([d]) \to \mathbb{R}$ is again referred to as the master equation's *potential function*.

Analogously to finite-horizon models, there are two natural ways to construct asymptotic Nash equilibria for the ergodic game: from the MFG system and from the master equation. By using controls from the MFG system, we mean players compute their optimal strategy through the potential function $\check{u}$ and the minimizer of the Hamiltonian, γ^*, and we expect the

$$\text{MFG system-derived strategy} \qquad (\gamma_y^*(X_t^{n,i}, \Delta_{X_t^{n,i}}\check{u}(t)))_{t \geq 0} \quad i \in [n], \qquad (1.6)$$

would be approximately optimal in the n-player game. Note that the potential function $\check{u}$ is coupled with the MFE flow $\check{\mu}$ through the MFG system (1.3) and using the MFG system-derived strategy results in a player's law equal to the MFE flow. When instead considering the stationary ergodic MFG system (1.4), the corresponding strategy is $\gamma_y^*(x, \Delta_x \bar{u})$.

We also consider the

$$\text{master equation-derived strategy} \qquad (\gamma_y^*(X_t^{n,i}, \Delta_{X_t^{n,i}}U_0(\cdot, \mu_t^{n,i})))_{t \geq 0} \quad i \in [n]. \qquad (1.7)$$

Here, $\mu_t^{n,i}$ denotes the current empirical distribution of the other players when all players are employing this profile. A property of the master equation assures us that replacing the empirical distribution in the above by the MFE flow just recovers the MFG system-derived strategy (namely, $\gamma_y^*(x, \Delta_x U_0(\cdot, \check{\mu}(t))) = \gamma_y^*(x, \Delta_x \check{u}(t))$) and so does not offer anything new. These two methods differ since, in the former, players respond to the deterministic flow of measures and their own states, hence can be thought of as *controls* while the latter lets players respond to the empirical distribution of the other players. In either case, players plug in a potential function different from the n-player game's potential function, which is harder to calculate.

At first blush, using the MFG system offers an advantage because (1.3) and (1.4) might have solutions even when (1.5) might not. Moreover, the MFG system-derived strategy is simpler because players only observe a theoretical MFE flow and their own state rather than the realized empirical distribution. However, through a numerical example provided in Section 7, we demonstrate the advantages of using the master equation-derived strategies; specifically, they result in a stationary distribution that is more concentrated around the stationary MFE, the law of the empirical distribution has fewer deviations to extreme values, and the realized cost in the n-player game is lower, as opposed to the MFG system-derived strategies.

1.2. **Relevant work.** The study of MFGs dates back to the pioneering work by Lasry and Lions [27], and Huang, Caines and Malhamé, e.g., [20, 19]. Ergodic MFGs on continuous state-spaces consider diffusions instead of jump processes and trace back to the origins of MFGs by Lasry and Lions [26]. Later, Cardaliaguet and Porretta [6] developed the theory of ergodic MFGs as limits of both finite-horizon MFGs and of infinite-horizon, discounted MFGs, again on the torus, and treated the ergodic master equation. The first reference to the *finite-state*, ergodic MFG comes from Gomes et. al. [17]. Later, Cohen and Zell [12] studied the finite-state ergodic MFG without the contractive assumptions of [17], and further proved the well-posedness of the finite-state, ergodic master equation. Other finite-state, ergodic MFG frameworks consider the discrete-time case [16] and with finite action space by Neumann [32].

In addition to investigating asymptotic Nash equilibria, we can directly compare the optimal cost functions of the MFG and the n-player game (under exact Nash equilibria). For open-loop controls, Lacker [22] and Fischer [14] established convergence of both asymptotic and exact Nash equilibria in the n-player finite-horizon game to the MFE in the corresponding MFG. The convergence in the continuous state-space, finite-horizon, with Markovian controls was established in the seminal work of Cardaliaguet, Delarue, Lasry, and Lions [5] using the master equation, and subsequently in the finite state-space by Cecchin and Pelino [9] and Bayraktar and Cohen [2]. Recently, Lacker [24, 25] showed that the convergence holds in the continuous state-space, finite horizon with general closed-loop controls (with and without common noise). The ergodic model in continuous state-space was studied by Arapostathis et al. [1], establishing the convergence of exact Nash equilibria to the MFE for open-loop controls. In sharp contrast to the finite-horizon work with closed-loop controls [24] and to the infinite-horizon with open-loop controls [1], Cardaliaguet and Rainer [7] provided an example of multiple mean field limits in ergodic differential

games with closed-loop controls, while the corresponding MFG has a unique MFE. The convergence problem in the ergodic Markovian setting is still an open question.

We emphasize that we study the question of asymptotic Nash equilibria and not convergence of the optimal value functions, as discussed in the preceding paragraph. In the following section, we precisely outline this work's novel contributions.

1.3. **Contributions.** Section 3 provides the complete presentation of our results, and we describe them heuristically here. We introduce the MFG system and master equation-derived strategy profiles (1.6) and (1.7), and prove that they constitute asymptotic Markovian Nash equilibria in Theorem 3.1 and Theorem 3.2, respectively. Commensurate results for the stationary ergodic MFG system (1.4) follow from Theorem 3.1 and are provided in Corollary 3.1. In each asymptotic Nash equilibria result, we also assert that the associated empirical measures converge to the MFE flow as the number of players grows and refer to this as *propagation of chaos*. Theorem 3.3 presents corresponding large deviation results for the empirical measures. We begin by describing two lemmas that assist the proofs of both Theorem 3.1 and Theorem 3.2, describe the differences between the proofs of each, and afterward speak to Theorem 3.3.

Since we assume all transition rates are bounded away from zero, and since the state-space $[d]$ is finite and fully-connected, any jump process on $[d]$ or $[d]^n$ must be exponentially ergodic. Lemma 3.1 uses this fact to prove that, when such a jump process is plugged into a function under a long-time average, the process' initial distribution is irrelevant. In particular, Lemma 3.1 removes the ergodic cost's dependence on its initial condition. The other intermediary, Lemma 4.2, estimates the difference between an empirical distribution of players under an arbitrary Markovian strategy profile and the measure flow $\check{\mu}$ from (1.3). The bound we prove depends on three things: the difference in the initial conditions, along with an exponential decay in time; the number of players, n; and, the difference between the arbitrary profile strategies and the MFG system-derived strategies (1.6).

Section 4 applies these lemmas to prove Theorem 3.1. To prove Theorem 3.1 in the most generality we can, we do not require that the mean-field cost F is Lasry–Lions monotone nor that the Hamiltonian H is concave. We use that F is Lipschitz and Lemma 3.1 to compare the costs (1.1) and (1.2), and use that the strategy is optimal in the MFG to obtain approximate optimality in the n-player game. At one point, Lemma 4.1, a concentration inequality for Bernoulli variables helps to estimate the error. The propagation of chaos, in turn, follows from Lemma 4.2, simplified by the fact that the strategies selected for the n-player game are exactly the controls from the MFG.

Theorem 3.2 establishes that the master equation-derived strategies have propagation of chaos and constitute an asymptotic Nash equilibrium too. While in Theorem 3.1 players respond to a theoretical MFE flow and so do not respond to their own system, in Theorem 3.2, players respond to their own empirical distribution. The main difficulty is the *feedback problem* arising from the Markovian strategies:

- a player deviates by selecting a different strategy than the others;
- this gives the deviating player her individual dynamics;
- the deviating player's position is plugged into the strategies of the other players;
- the other players' positions are plugged into the strategy of the deviating player.

One might already see the issue. If we compare empirical measures with and without a deviating player, the feedback problem ensures that any naive argument relying only on Lipschitz estimation becomes stuck in a loop. For the finite-horizon case, Gronwall's inequality can escape such loops [23], but it is not applicable for the infinite-horizon. In fact, as mentioned in the introduction, [7] showed that, for non-Markovian strategies, there may even be multiple limits to the exact Nash equilibria, while the corresponding MFG has only one MFE.

We address the feedback problem by combining Lemma 4.2 with Lemma 5.2. Now, recall that in Lemma 4.2, one of the terms in the upper bound for the difference between an empirical distribution of players under an arbitrary Markovian strategy profile and the measure flow $\check{\mu}$, is an integral over the difference between the arbitrary profile strategies and the MFG system-derived strategies (1.6). In the case of the MFG system-derived strategies, this integral term in the upper bound vanishes. This term does not vanish for the master equation-derived strategy profile. In the MFG setting, a typical way to deal with

an integral term like that one is to use the *duality* of the MFG system. That is, integrating a potential function against its measure flow as a test function results in estimates that overcome the feedback problem. However, the MFG system is an ODE system, while what we need to estimates concerning a random, empirical measure. So for Lemma 5.2, we try to proceed like a standard duality argument, but encounter random terms that we will need to estimate or that vanish under expectation as martingales. Consequently, we accumulate n-dependent error that we drag through the proof. The part of the proof of Theorem 3.2 that concerns Lemma 5.2 uses expectations taken under a stationary distribution. Hence, we deal with *stationary* Markovian strategies to start. Toward the end of the proof, we upgrade our analysis so it handles any Markovian strategy by Lemma 5.1. Lemma 5.1 relies on an auxiliary ergodic system that we construct through (another) auxiliary discounted system and by sending the discount factor to zero. We then prove the optimal strategy for a deviating player is *stationary* Markovian and use this fact to finish the proof of Theorem 3.2.

Lastly, Theorem 3.3 obtains large deviations results: (1) for the different empirical measures' laws at a fixed time, (2) for the laws of the measure's paths, and (3) for the invariant distributions. Like Theorem 3.2, the proof of the large deviation principle requires that F is monotone and H concave, but this time much of the heavy lifting is done by [4]. We only need to verify assumptions concerning ODE stability for the Kolmogorov equation from (1.3) and the details are presented in Section 6.

In the final part of the paper, Sections 7 and 8, we study a numerical example to elucidate differences between MFG system-derived and master equation-derived asymptotic Nash equilibria. We contrast the strategies, the stationary distributions of the empirical distributions, the large deviation rate functions, and the players' realized costs. As we discuss, the master equation-derived strategies result in a lower realized cost to the players than those of the MFG.

Computing the master equation-derived strategies requires computing U_0 and, since the master equation is a fully nonlinear PDE, it is not clear how to analytically do this, even for a concrete example. To approximate the solution to the master equation, we adapt the deep Galerkin method (DGM) of [33] to the ergodic master equation in Algorithm 1. The DGM samples the state-space of the PDE, and using the PDE as part of the loss function of a neural network, trains the network to approximate the solution to the PDE. For the last leg of the numerics, we return to the large deviation principles established in Theorem 3.3. In the case the dimension $d = 2$, we recognize a way to interpret the empirical measures as a birth-death process and introduce a function that we can use to prove an explicit formula for the good rate functions holds; we present this special case in an ancillary result, Theorem 8.1.

1.4. **Outline of the paper.** In Section 2 we introduce the n-player ergodic game and the assumptions we consider. After, Section 3 presents the main theorems of the paper. Section 4 gives the proof for Theorem 3.1 through its several lemmas and Section 5 does the same for Theorem 3.2. Briefly, Section 6 gives the proofs for the large deviation result Theorem 3.3. In Section 7, we begin to numerically investigate the different aspects of the asymptotic Nash equilibria, where one profile is derived from the MFG system and one from the master equation. Finally, Section 8 concludes the paper with the numerics pertaining to large deviations. We conclude the current section by introducing some commonly used notation.

1.5. **Notation and basic definitions.** We denote the ray $[0, \infty)$ by $\mathbb{R}_+$, the set of positive integers by $\mathbb{N}$, and the set of positive integers larger than one by $\mathbb{N}_{>1}$. Let $\vec{1} \in \mathbb{R}^d$ be the vector of all ones, so $k\vec{1}$ will be an additive constant vector appearing in the context of uniqueness of potential functions. We also denote the sets $[n] := \{1, 2, \ldots, n\}$ and $[d] := \{1, 2, \ldots, d\}$, where n denotes a number of players and d the number of states. Let $\mathbb{1}_A$ be the indicator function corresponding to the event A; that is, the function that is 1 on A and 0 on its complement. The parameters i, j, and k (resp., x, y, and z) represent a player (reps., state) and belong to $[n]$ (resp., $[d]$). For any random variable Y with distributional law $\mathcal{L}(Y)$, we may write $Y \sim \mathcal{L}(Y)$. Fix two positive real numbers $0 < \mathfrak{a}_l < \mathfrak{a}_u < +\infty$. For any $d \in \mathbb{N}_{>1}$ and $m \in \mathbb{N}_{>1}$, we set:

$$\mathcal{P}([d]) := \Big\{ \eta \in \mathbb{R}_+^d : \sum_{x \in [d]} \eta_x = 1 \Big\}, \quad \text{a } (d-1)\text{-dimensional simplex,}$$

$$\mathcal{P}^m([d]) := \left\{ \eta \in \mathcal{P}([d]) : m\eta \in (\mathbb{N} \cup \{0\})^d \right\}, \quad \text{valid configurations of } m \text{ players},$$

$$\mathbb{A}^d_{-x} := \left\{ a \in \mathbb{R}^d : \forall y \neq x, \quad a_y \in [\mathfrak{a}_l, \mathfrak{a}_u], \quad a_x = - \sum_{y, y \neq x} a_y \right\}, \text{ allowable rates leaving state } x,$$

$$\mathbb{A}^d_{[d]} := \bigcup_{x \in [d]} \mathbb{A}^d_{-x}, \quad \text{all allowable rates.}$$

Moreover, let $\mathcal{Q}$ be the set of $d \times d$ rate matrices; that is, $\alpha \in \mathcal{Q}$ if for every distinct $x, y \in [d]$, $\alpha_{xy} \geq 0$ and $\sum_{z \in [d]} \alpha_{xz} = 0$. When necessary, we denote the Lipschitz constant of a given function h by $C_{L,h}$ and a bound by C_h. For any $x \in [d]$ and $p \in \mathbb{R}^d$, set $\Delta_x p := (p_y - p_x)_{y \in [d]}$. For a function K mapping $\mathbb{R}^d \ni p \mapsto K(p) \in \mathbb{R}$, we denote its gradient by K_p or $D_p K$ and Hessian matrix by K_{pp}. We also use derivatives on the simplex as follows. When $K : \mathcal{P}([d]) \to \mathbb{R}$ is a measurable function, we say that K is differentiable if there exists a function $D^\eta_{yz} K : \mathcal{P}([d]) \to \mathbb{R}^{d \times d}$ such that

$$D^\eta_{yz} K(\eta) = \lim_{h \to 0^+} h^{-1}(K(\eta + h(e_z - e_y)) - K(\eta))$$

for all $z \in [d]$, where e_y is the y-th standard basis vector in $\mathbb{R}^d$. Denote the vector $D^\eta_y K := (D^\eta_{yz} K)_{z \in [d]}$ and the matrix $D^\eta K := (D^\eta_{yz} K)_{y,z \in [d]}$. The set of continuous functions with at least one continuous derivative, which are defined on $\mathbb{R}_+$ and with values in some Banach space X, is denoted $\mathcal{C}^1(\mathbb{R}_+, X)$. Lastly, for $K : [d] \times \mathcal{P}([d]) \to \mathbb{R}$ we say K is *Lasry–Lions monotone* if for all $\eta, \hat{\eta} \in \mathcal{P}([d])$,

$$\sum_{x \in [d]} (K(x, \eta) - K(x, \hat{\eta}))(\eta_x - \hat{\eta}_x) \geq 0.$$

2. The model

In Section 2.1, we describe the n-player, finite-state ergodic game model. The assumptions about the models are listed in Section 2.2. Next in Section 2.3, we introduce the corresponding ergodic MFG via its dynamics. Then, we recall some results about the ergodic MFG systems (1.3) and (1.4) as well as their connection to the master equation (1.5).

2.1. **The n-player model.** Throughout the paper, we consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}, \mathbb{P})$, satisfying the usual conditions, supporting Poisson random measures and jump processes as we detail below.

We consider a finite-state, continuous-time, symmetric n-player game with weak interactions. We begin by describing the n-player game's dynamics. The state-space is the finite set $[d]$, where d is an integer larger than one. Players choose their rates of transition from their own state to the other states and each player aims to minimize a personal cost. Both the transition rates and the cost depend on the players' current states.

Denote the players' positions via the vector $\boldsymbol{x} := (x_1, \ldots, x_n) \in [d]^n$. Players move at rates from the compact set $\mathbb{A} = [\mathfrak{a}_l, \mathfrak{a}_u] \subset \mathbb{R}_+$. A measurable function $\alpha : \mathbb{R}_+ \times [d]^n \to \mathbb{R}^d$ satisfying $\alpha_y(t, \boldsymbol{x}) \in \mathbb{A}$ for any $y \neq x_i$, and $\alpha_{x_i}(t, \boldsymbol{x}) = - \sum_{y \in [d], y \neq x_i} \alpha_y(t, \boldsymbol{x})$ is called a *Markovian strategy* for player $i \in [n]$. Whenever α is independent of the time component we refer to it as a *stationary Markovian strategy*. For $y \neq x_i$, the magnitude $\alpha^i_y(t, \boldsymbol{x})$ represents the transition rate of player i from its own state x_i to state y, when the rest of the players are in states $\boldsymbol{x}^{-i} := (x_j : j \in [n] \setminus \{i\})$. Denote by $\mathcal{A}^{n,i}$ (resp., $\mathcal{A}^{n,i}_S$) the set of all Markovian strategies (resp. stationary Markovian strategies) for each one of the players. Then $\boldsymbol{\alpha} := (\alpha^i : i \in [n]) \in \mathcal{A}^{(n)} := \prod_{i=1}^n \mathcal{A}^{n,i}$ (resp., $\boldsymbol{\alpha} = (\alpha^i : i \in [n]) \in \mathcal{A}^{(n)}_S := \prod_{i=1}^n \mathcal{A}^{n,i}_S$) is called a *Markovian strategy profile* (or respectively, *stationary Markovian strategy profile*).

For the following description of player dynamics, fix a number $n \in \mathbb{N}_{>1}$, an initial state for the players $\boldsymbol{x} \in [d]^n$, and a Markovian strategy profile $\boldsymbol{\alpha} \in \mathcal{A}^{(n)}$. Often, we will refer to the n-players' initial state as a *configuration* of the players. Given the fixed Markovian strategy profile $\boldsymbol{\alpha}$, the process standing for the dynamics of the n-players on $[d]^n$ is denoted by the bold vector $(\boldsymbol{X}_t)_{t \geq 0}$. On occasions when we

want to emphasize the dependence of $\boldsymbol{X}$ on the Markovian strategy profile $\boldsymbol{\alpha}$, we denote it by $(\boldsymbol{X}_t^{\boldsymbol{\alpha}})_{t \geq 0}$. Rigorously, the dynamics of player i are given by the jump process:

$$X_t^{\boldsymbol{\alpha},i} = X_0^{\boldsymbol{\alpha},i} + \int_0^t \int_{\mathbb{A}^d} \sum_{y \in [d]} (y - X_{s^-}^{\boldsymbol{\alpha},i}) \mathbb{1}_{\{\xi_y \in (0, \alpha_y^i(s^-, \boldsymbol{X}_{s^-}^{\boldsymbol{\alpha}}))\}} \mathcal{N}_i(ds, d\xi),$$

where $\{\mathcal{N}_i\}_{i \in [n]}$ are i.i.d. Poisson random measures with common intensity measure ν given by:

$$\nu(E) := \sum_{y \in [d]} \mathrm{Leb}(E \cap \mathbb{A}_y^d), \tag{2.1}$$

where $\mathbb{A}_y^d := \{u \in \mathbb{A}^d \mid u_x = 0 \text{ for all } x \neq y\}$ and Leb is the Lebesgue measure on $\mathbb{R}$.

Recall that each player attempts to minimize the cost (1.2) that is a function of the initial configuration $\boldsymbol{x} \in [d]^n$ and the Markovian strategy profile $\boldsymbol{\alpha}$. We have that $f : [d] \times \mathbb{A}_{[d]}^d \to \mathbb{R}$ is measurable, $F : [d] \times \mathcal{P}([d]) \to \mathbb{R}$ is measurable, and

$$\mu_t^{\boldsymbol{\alpha},i} := \frac{1}{n-1} \sum_{j \in [n], j \neq i} \delta_{X_t^{\boldsymbol{\alpha},j}}.$$

As a technicality, we also assume that $f(x, \alpha(t, \boldsymbol{x}))$ should be independent of the x-th coordinate of $\alpha(t, \boldsymbol{x})$. This technical assumption is justified since the x-th coordinate of the function $\alpha(t, \boldsymbol{x})$ for a player in a state x does not represent a rate, but rather the negative sum of the player's other rates. In that sense, f already accounts for the "payment" for the player's rates. With this game setup, we have a notion of ε-Nash equilibrium:

Definition 2.1. The Markovian strategy profile $\boldsymbol{\alpha} \in \mathcal{A}^{(n)}$ is called an ε-*Nash equilibrium profile* for the n-player ergodic game if for any $i \in [n]$ and any $\beta^i \in \mathcal{A}^{n,i}$, one has

$$J_0^{n,i}(\boldsymbol{x}, \boldsymbol{\alpha}) \leq J_0^{n,i}(\boldsymbol{x}, [\boldsymbol{\alpha}^{-i}; \beta^i]) + \varepsilon,$$

where

$$[\boldsymbol{\alpha}^{-i}; \beta^i] := (\alpha^1, \dots, \alpha^{i-1}, \beta^i, \alpha^{i+1}, \dots, \alpha^n).$$

For all $n \in \mathbb{N}_{>1}$, let $\boldsymbol{\alpha}^{(n)} \in \mathcal{A}^{(n)}$ be a ε_n-Nash Markovian equilibrium. We say that $(\boldsymbol{\alpha}^{(n)})_{n \in \mathbb{N}_{>1}}$ is an *asymptotic Nash equilibrium* whenever $\varepsilon_n \to 0$ as $n \to \infty$.

Remark 2.1. While we define an ε-Nash equilibrium in terms of Markovian strategies and not stationary Markovian strategies, an ε-Nash equilibrium can be a stationary Markovian strategy profile. This means that the deviation of one player, even to a time-dependent strategy, will not benefit that player by more than ε.

Definition 2.2. A Markovian strategy profile $\boldsymbol{\alpha} := (\alpha^i)_{i \in [n]}$ is called *symmetric* if for all players $i, j \in [n]$, whenever $x^i = x^j$, for all $t \in \mathbb{R}_+$, and $\boldsymbol{x} \in [d]^n$:

$$\alpha^i(t, \boldsymbol{x}) = \alpha^j(t, \boldsymbol{x}).$$

For a stationary Markovian strategy profile to be called symmetric, the definition is the same, except there is no time component of α^i, α^j.

2.2. **Assumptions.** We require a few assumptions that we split into two sets. The first set of assumptions is labeled with an A. We refer to this set of assumptions as the *reduced assumptions*. The reduced assumptions are sufficient to establish the main result concerning asymptotic Nash equilibria, Theorem 3.1. For Theorem 3.2 and Theorem 3.3, we require additional regularity of the functions comprising the problem data. These assumptions are denoted with a B. The set of all assumptions (both A and B indexed) are referred to as *full assumptions*.

(A_1) For $0 < \mathfrak{a}_l < \mathfrak{a}_u < \infty$, the action space is $\mathbb{A} := [\mathfrak{a}_l, \mathfrak{a}_u]$.

Consider the Hamiltonian $H : [d] \times \mathbb{R}^d \to \mathbb{R}$ and the optimal selector function $\gamma^* : [d] \times \mathbb{R}^d \to \mathbb{A}_{[d]}^d$, given by:

$$H(x,p) := \min_{a \in \mathbb{A}_{-x}^d} \left\{ f(x,a) + a \cdot p \right\}, \qquad \gamma^*(x,p) := \arg\min_{a \in \mathbb{A}_{-x}^d} \left\{ f(x,a) + a \cdot p \right\}.$$

(A_2) We assume that the Hamiltonian H admits a unique optimal rate selector given by the measurable function γ^*. Moreover, γ^* is Lipschitz in its p argument.

The latter assumptions hold, for instance, when the running cost component f is strictly convex in the control/strategy α. In fact, the optimal rate selector γ^* is globally Lipschitz when f is uniformly convex.

(A_3) Given any state $x \in [d]$, the mean field cost $F(x, \cdot) \in \mathcal{C}^1(\mathcal{P}([d]))$ is Lipschitz continuous.

Consequently, $F(x, \cdot)$ is bounded since it is defined on a compact set. Lastly, we assume:

(B_1) Additionally, the mean field cost F is Lasry–Lions monotone and $D^\eta F(x, \cdot)$ is Lipschitz.

(B_2) The Hamiltonian H is twice continuously differentiable with respect to its p-argument. Moreover, on any compact set $[-K, K]$ we will assume that $D_p H$ and $D_{pp}^2 H$ are Lipschitz in p and that there exists $C_{2,H} > 0$ such that:

$$D_{pp}^2 H(x,p) \le -C_{2,H}.$$

Because any argument of H in the paper will be uniformly bounded, it is enough to assume these regularity properties of H only on compact sets. At any point when H is differentiable, we have by [17, Proposition 1] that:

$$\gamma^*(x,p) = D_p H(x,p). \tag{2.2}$$

Remark 2.2. The full assumptions are considered standard in MFG literature—for instance, [12, 2, 9] all consider finite-state MFGs with a concave (or if they maximize a profit instead of minimizing a cost, convex) Hamiltonian and a separable cost $f + F$, with Lasry–Lions monotone assumptions on F. These assumptions are typical for studying the well-posedness of the MFG and master equation over continuous state-spaces as well [5, 6, 3]. However, some papers consider alternate assumptions to deal with Hamiltonians of a different form. A few examples of alternate schemes include: [15] considers a displacement monotonicity condition allowing for analysis of non-separable Hamiltonians, [30] considers non-smooth Hamiltonians under technical assumptions, and [31] considers an anti-monotonicity condition. Each approach is able to handle data that the other approaches cannot. By considering the reduced assumptions when appropriate, we prove Theorem 3.1 in the greatest generality so that it might be widely applicable.

2.3. **The ergodic MFG, the MFG systems, and the master equation.** The finite-state stationary ergodic MFG system (1.4) was first studied in [17] under additional contractivity assumptions. Later, [12] studied the same MFG system without any contractivity assumptions and we recall the major points here.

In an ergodic MFG, the representative player chooses a measurable control function $\alpha : \mathbb{R}_+ \times [d] \to \mathbb{A}_{[d]}^d$ while playing against a measure flow $\mu : \mathbb{R}_+ \to \mathcal{P}([d])$ in order to minimize the long-run average cost (1.1).

The representative player's position evolves according to a controlled jump process, given by:

$$\mathcal{X}_t^\alpha = \mathcal{X}_0^\alpha + \int_0^t \int_{\mathbb{A}^d} \sum_{y \in [d]} (y - \mathcal{X}_{s^-}^\alpha) \mathbf{1}_{\{\xi_y \in (0, \alpha_y(s^-, \mathcal{X}_{s^-}^\alpha))\}} \mathcal{N}(ds, d\xi),$$

where $\mathcal{X}_0^\alpha$ has the distributional law $\mu(0)$ and $\mathcal{N}$ is a Poisson random measure with intensity measure ν as in (2.1). For the next definition, recall the ergodic cost function in the MFG, J_0, provided in (1.1).

Definition 2.3 (Ergodic MFE). Let $\alpha \in \mathcal{A}$ and $\mu : \mathbb{R}_+ \to \mathcal{P}([d])$. Let $\mathcal{X}^\alpha$ be the dynamics of the representative player using α. We say that (α, μ) is an ergodic MFE if the following hold:

- For all $t \in \mathbb{R}_+$, $\mathcal{L}(\mathcal{X}_t^\alpha) = \mu(t)$,

- and the control α satisfies:

$$J_0(\alpha, \mu) = \inf_{\beta \in \mathcal{A}} J_0(\beta, \mu).$$

Typically in ergodic problems, there is an optimal cost or *value* that is independent of the initial state. Moreover, the long-time averaging in the cost often yields a stationary Markovian optimal control and hence a stationary measure that describes a limiting law for the representative player. We saw this stationarity in the system (1.4) from before while the general system was given in (1.3). For any solution $(\bar{\varrho}, \bar{u}, \bar{\mu})$ to (1.4), and there could be more than one solution, note that the potential function $\bar{u}$ is modifiable up to addition of a constant vector—this comes from the fact that $\bar{u}$ only appears in (1.4) as the finite difference vector $\Delta_x \bar{u}$.

The result of [12, Proposition 2.1] proved that the stationary ergodic MFG system (1.4) is well-posed and that any solution to it naturally characterizes a mean field equilibrium, while any mean field equilbrium naturally characterizes a solution. The first main point is recalled here:

Proposition 2.1. With reduced assumptions, (1.4) admits a solution, and, with full assumptions, the solution to (1.4) is unique. We emphasize that for a unique solution $(\bar{\varrho}, \bar{u}, \bar{\mu})$ to (1.4), the potential function $\bar{u}$ is only unique up to an additive constant vector $k\vec{1}$, $k \in \mathbb{R}$. Moreover, under the reduced assumptions, $((\gamma_y(x, \Delta_x \bar{u}))_{x,y \in [d]}, \bar{\mu})$ is an ergodic MFE and so for all $t > 0$:

$$\mathcal{X}_t^{(\gamma_y(x, \Delta_x \bar{u}))_{x,y \in [d]}} \sim \bar{\mu}, \quad \text{with} \quad \bar{\varrho} = J_0((\gamma_y(x, \Delta_x \bar{u}))_{x,y \in [d]}, \bar{\mu})$$

when $\mathcal{X}_0^{(\gamma_y(x, \Delta_x \bar{u}))_{x,y \in [d]}} \sim \bar{\mu}$.

The proof of existence under reduced assumptions is not explicitly provided in [12], but a look at the proofs for existence allows us to recognize that only the reduced assumptions are used there. For the time-dependent system, [18] studied an example that fits under the reduced assumptions, where a solution to (1.3) exists only when η is in a neighborhood of $\bar{\mu}$. Moreover, (1.3) and the master equation (1.5) are only known to have a solution (for all initial data) under the full assumptions; in that case, the solutions to (1.3), (1.4), and (1.5) are also unique. The main points of [12, Theorem 2.1] and the relationship between (1.3) and (1.5) are recalled below.

Proposition 2.2. Under the full assumptions, (ϱ, U_0) has a unique solution, where U_0 is unique up to an additive constant vector $k\vec{1}$, $k \in \mathbb{R}$. Similarly, the solution $(\check{\varrho}, \check{u}, \check{\mu})$ to (1.3) is unique. The potential function U_0 is Lipschitz and Lasry–Lions monotone with a Lipschitz derivative $D^\eta U_0$ as well. The relationship between the MFG system and master equation is given by:

$$U_0(x, \check{\mu}(t)) = \check{u}_x(t), \text{ and } \varrho = \check{\varrho}.$$

Moreover, for any initial condition $\check{\mu}(0) = \eta \in \mathcal{P}([d])$, $((\gamma_y(x, \Delta_x \check{u}(\cdot)))_{x,y \in [d]}, \check{\mu})$ is an ergodic MFE and there exist $C, c > 0$ independent of η such that for all $t > 0$,

$$\mathcal{X}_t^{(\gamma_y^*(x, \Delta_x \check{u}(t)))_{x,y \in [d]}} \sim \check{\mu}(t), \quad \text{with } \check{\varrho} = J_0((\gamma_y(x, \Delta_x \check{u}(t)))_{x,y \in [d]}, \check{\mu}), \quad \text{and } |\check{\mu}(t) - \bar{\mu}| \leq Ce^{-ct},$$

when $\mathcal{X}_0^{(\gamma_y^*(x, \Delta_x \check{u}(t)))_{x,y \in [d]}} \sim \check{\mu}(0)$.

We now make some final remarks on the time-dependent ergodic MFG system (1.3).

Remark 2.3. Under the full assumptions, there is a way to construct the solution for (1.3) for any initial condition $\eta \in \mathcal{P}([d])$. Let (ϱ, U_0) satisfy (1.5), and define $\mu_0 : \mathbb{R}_+ \to \mathcal{P}([d])$ as the solution to:

$$\frac{d}{dt}\mu_0(t, x) = \sum_{y \in [x]} \mu_0(t, y)\gamma_x^*(y, \Delta_y U_0(\cdot, \mu_0(t))), \quad \mu_0(0) = \eta \in \mathcal{P}([d]), \tag{2.3}$$

where we abuse notation and denote the x-th coordinate of $\mu_0(t)$ by $\mu_0(t, x)$, hence referring to μ as a mapping from $\mathbb{R}_+ \times [d] \to [0,1]$. Note that under the full assumptions, $\check{\varrho} = \bar{\varrho} = \varrho$. Note that (2.3) is a forward Kolmogorov equation associated with a Markov jump process $\mathcal{X}^0$, where $\mathcal{X}^0$ jumps with rates $[\gamma_y^*(x, \Delta_x U_0(\cdot, \mu_0(t)))]_{x,y \in [d]}$. Next, set $u_0(t, x) := U_0(x, \mu_0(t))$. By a short computation, we can verify that (ϱ, u_0, μ_0) satisfies (1.3). Of course, for $\eta = \bar{\mu}$, the solution would be the stationary one $(\bar{\varrho}, \bar{u}, \bar{\mu})$ to (1.4).

Remark 2.4. At the same time, we note that any stationary solution $(\bar{\varrho}, \bar{u}, \bar{\mu})$ solving (1.4) solves (1.3) with $\check{u}(0) = \bar{\mu}$. Therefore Proposition 2.1 implies that, even under the reduced assumptions, there is a solution to (1.3) for at least one initial condition.

Remark 2.5. All the potential functions, $\check{u}$, $\bar{u}$, and U_0, are only unique up to addition of a constant vector, $k\vec{1}$, $k \in \mathbb{R}$. For (1.3) this means that $\check{u}$ can be unbounded as t changes. To avoid this we will say that $(\check{\varrho}, \check{u}, \check{\mu})$ is a solution to (1.3) only when $\check{u}$ is bounded. For both examples given in Remark 2.3 and Remark 2.4, $\check{u}$ is bounded because U_0 is bounded.

3. Asymptotic Results

In this section we provide the results for the three kinds of asymptotic Nash equilibria derived from: the time-dependent MFG (1.3), the stationary MFG system (1.4), and the master equation (1.5). The following subsection introduces the large deviation results for the empirical distributions of each of these asymptotic equilibria, formulated generally.

3.1. **Asymptotic Markovian Nash-equilibria.** Before presenting the asymptotic Markovian Nash equilibria, we introduce a lemma that eliminates the dependence on the initialization for long-time averaging under expectation, particularly for $J_0^{n,i}$. For the rest of the paper, we write $\mathbb{E}_\pi[g(\boldsymbol{X}_t)]$ to denote the conditional expectation $\mathbb{E}[g(\boldsymbol{X}_t) \mid \boldsymbol{X}_0 \sim \pi]$, for g a measurable function. For example, $g(\boldsymbol{X}_t)$ might be the empirical distribution of $\boldsymbol{X}_t$.

Lemma 3.1. Fix $n \in \mathbb{N}_{>1}$. Let $(\beta^i)_{i \in [n]} = \boldsymbol{\beta} \in \mathcal{A}^{(n)}$, and let $g : [d]^n \to \mathbb{R}$ be measurable and bounded. If $(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{t \geq 0}$ is the jump process of players using the profile $\boldsymbol{\beta}$, then for all $\pi_0, \pi_1 \in \mathcal{P}([d]^n)$:

$$\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_0} \int_0^T g(\boldsymbol{X}_t^{\boldsymbol{\beta}}) dt = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_1} \int_0^T g(\boldsymbol{X}_t^{\boldsymbol{\beta}}) dt.$$

The proof of this lemma is standard and follows typical ergodicity techniques. For the sake of completeness, its proof is provided in the Appendix.

Remark 3.1. By Lemma 3.1, we obtain that the n-player ergodic cost $J_0^{n,i}$ is independent of the initial configuration of players. So for a Markovian strategy profile $\boldsymbol{\beta} \in \mathcal{A}^{(n)}$ and configuration $\boldsymbol{x} \in [d]^n$, we may write $J_0^{n,i}(\boldsymbol{x}, \boldsymbol{\beta})$ as $J_0^{n,i}(\boldsymbol{\beta})$ without confusion. When $\boldsymbol{\beta}$ is symmetric, the costs for all the players are the same and we may further simplify the notation and write the latter as $J_0^n(\boldsymbol{\beta})$.

For any Markovian strategy profile $\boldsymbol{\beta} := (\beta^i)_{i \in [n]} \in \mathcal{A}^{(n)}$, denote by $(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{t \geq 0}$ the jump process of all players using $\boldsymbol{\beta}$ and $(X_t^{\boldsymbol{\beta},i})_{t \geq 0}$ the position of the i-th player. Denote $(\mu_t^{\boldsymbol{\beta}})_{t \geq 0}$ as the empirical measure of $(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{t \geq 0}$ and, after removing the i-th player as $(\mu_t^{\boldsymbol{\beta},i})_{t \geq 0}$. The empirical distribution $\mu^{\boldsymbol{\beta}}$ has the dynamics:

$$\mu_t^{\boldsymbol{\beta}} = \mu_0^{\boldsymbol{\beta}} + \int_0^t \sum_{x,y \in [d]} \frac{e_y - e_x}{n} dK_{xy}^{\boldsymbol{\beta}}(s), \tag{3.1}$$

where:

$$K_{xy}^{\boldsymbol{\beta}}(t) = \int_0^t \int_{\mathbb{A}^{d \times d}} \mathbb{1}\left\{\left[0, \sum_{i \in [n]} \mathbb{1}_{\{X_{s^-}^{\boldsymbol{\beta},i}=x\}} \beta_{xy}^i(s^-, \boldsymbol{X}_{s^-}^{\boldsymbol{\beta}})\right]\right\} \mathcal{K}_{xy}^{\boldsymbol{\beta}}(ds, dr), \tag{3.2}$$

and $\mathcal{K}_{xy}^{\boldsymbol{\beta}}$ is a Poisson random measure with jump intensity given by:

$$\tilde{\nu}(A) := \sum_{x,y \in [d]} \text{Leb}(A \cap \mathbb{A}_{xy}), \quad \mathbb{A}_{xy} := \{u \in \mathbb{A}^{d \times d} \mid u_{wz} = 0 \quad \forall (w,z) \neq (x,y)\}.$$

We note that the compensator for this random measure is:

$$\int_0^t \sum_{i \in [n]} \mathbb{1}_{\{X_s^{\boldsymbol{\beta},i}=x\}} \beta_{xy}^i(s, \boldsymbol{X}_s^{\boldsymbol{\beta}}) ds.$$

Now with the requisite background, we introduce the main theoretical results of the paper. The first, Theorem 3.1, asserts that any solution to the time-dependent ergodic MFG system gives an approximate Nash equilibrium when applied in the n-player game; moreover, for the empirical measure associated with this profile, we have the propagation of chaos.

Theorem 3.1. Assume the reduced assumptions hold. Let $\eta \in \mathcal{P}([d])$ and assume (1.3) has a solution $(\check{\varrho}, \check{u}, \check{\mu})$ with initial condition $\check{\mu}(0) = \eta$, as with the examples given in Remarks 2.3 and 2.4. Then there exists $C > 0$, independent of n and η, such that the Markovian strategy profile:

$$\check{\Gamma} := \big(\check{\gamma}\big)_{i \in [n]}, \quad \check{\gamma}(t) := [\gamma_y^*(x, \Delta_x \check{u}(t))]_{x,y \in [d]},$$

is a Markovian $(C/\sqrt{n})$-Nash equilibrium. Moreover, the cost for each player in the game, $J_0^n(\check{\Gamma})$, satisfies

$$|J_0^n(\check{\Gamma}) - \check{\varrho}| \leq \frac{C}{\sqrt{n}}. \tag{3.3}$$

Also, there exists $c > 0$ such that for any $\pi \in \mathcal{P}([d]^n)$ and any $n \in \mathbb{N}_{>1}$,

$$\mathbb{E}_\pi |\mu_t^{\check{\Gamma}} - \check{\mu}(t)| \leq Ce^{-ct}\mathbb{E}_\pi |\mu_0^{\check{\Gamma}} - \check{\mu}(0)| + \frac{C}{\sqrt{n}}. \tag{3.4}$$

Suppose $(X_0^{\check{\Gamma},i})_{i \in [n]}$ are exchangeable, though not necessarily independent, with $X_0^{\check{\Gamma},i} \sim \check{\mu}(0) = \eta$ for all $i \in [n]$. For $\pi = \check{\pi}$, the joint distribution induced by this initialization, we have from (3.4) that:

$$\sup_{t \in \mathbb{R}_+} \mathbb{E}_{\check{\pi}} |\mu_t^{\check{\Gamma}} - \check{\mu}(t)| \leq \frac{C}{\sqrt{n}}. \tag{3.5}$$

Remark 3.2. The assumption that $\check{\mu}(0)$ is not on the boundary of $\mathcal{P}([d])$ is so that we can assume there exists some constant $q > 0$ for which $\check{\mu}_x(t) > q$ for all $x \in [d]$ and for $t \in \mathbb{R}_+$. When $\check{\mu}(0)$ is on the boundary, we can replace the final display with:

$$\sup_{t \geq 1} \mathbb{E}_{\check{\pi}} |\mu_t^{\check{\Gamma}} - \check{\mu}(t)| \leq \frac{C}{\sqrt{n}},$$

though the choice of 1 is arbitrary—any positive constant would do. This is because we must allow a moment for the measure to enter the interior of the simplex. Since the transition rates between any states are bounded away from zero, the deterministic process $\check{\mu}$ will always move away from the boundary. Throughout the remainder of the paper, this technical point will remain implicit in any statement of a supremum over all $t \in \mathbb{R}_+$, except for Lemma 4.1, where we note a few details.

Remark 3.3. For the Markovian strategy profile $\check{\Gamma}$, each player plays according to her own state and a deterministic flow of measures. We emphasize that we allow the deviating player to use a strategy based on the current time as well as the knowledge of the current states of all the players. As we will show, this extra information does not help the deviating player that much, since they can only improve their cost by at most $C/\sqrt{n}$.

Remark 3.4. Recent work by Höfer and Soner [18] studies an example of non-unique solutions to (1.3) when $d = 2$ under a specified anti-monotone mean field running cost F, and quadratic running cost f. They show that, for certain choices of constants in the rates $\mathbb{A}$ and by scaling F, there are infinitely-many distinct MFE. Fortunately, Theorem 3.1 applies to these examples as well. Since we only require the reduced assumptions, we can accommodate these examples and prove that the infinitely-many equilibria studied in [18] constitute $(C/\sqrt{n})$-Nash equilibria as well.

With Theorem 3.1 in hand, which provides an asymptotic Markovian Nash equilibrium based on the time-dependent ergodic MFG system (1.3), we recall that the stationary ergodic MFG system (1.4) is a special case of (1.3) with the initial state $\eta = \bar{\mu}$. Consequently, strategies derived from the stationary system also yield an asymptotic Markovian Nash equilibrium. Namely, let $\mu_t^{\bar{\Gamma}}$ be the empirical distribution of n players using the time-stationary Markovian strategy profile:

$$\bar{\Gamma} := (\bar{\gamma})_{i \in [n]} \quad \text{where} \quad \bar{\gamma} := [\gamma_y^*(x, \Delta_x \bar{u})]_{x,y \in [d]}. \tag{3.6}$$

So, when $\check{\Gamma} = \bar{\Gamma}$, we can apply Theorem 3.1 to get the following corollary.

Corollary 3.1. Assume the reduced assumptions hold. Then, the stationary ergodic MFG system (1.4) has a solution $(\bar\varrho, \bar u, \bar\mu)$. Additionally, there exists $C > 0$, independent of n, such that the Markovian strategy profile $\bar\Gamma$ is a Markovian $(C/\sqrt{n})$-Nash equilibrium. Moreover, the cost for each player in the game, $J_0^n(\bar\Gamma)$, satisfies

$$|J_0^n(\bar\Gamma) - \bar\varrho| \le \frac{C}{\sqrt{n}}. \tag{3.7}$$

Also, for any $\pi \in \mathcal{P}([d]^n)$,

$$\mathbb{E}_\pi |\mu_t^{\bar\Gamma} - \bar\mu| \le Ce^{-ct}\mathbb{E}_\pi|\mu_0^{\bar\Gamma} - \bar\mu| + \frac{C}{\sqrt{n}}. \tag{3.8}$$

Suppose $(X_0^{\bar\Gamma,i})_{i\in[n]}$ are exchangeable, though not necessarily independent, with $X_0^{\bar\Gamma,i} \sim \bar\mu$ for all $i \in [n]$. For $\pi = \bar\pi$, the joint distribution induced by this initialization, we have from (3.8) that:

$$\sup_{t\in\mathbb{R}_+} \mathbb{E}_{\bar\pi}|\mu_t^{\bar\Gamma} - \bar\mu| \le \frac{C}{\sqrt{n}}.$$

The next result shows that the time-dependent Markovian strategy using the master equation's potential function U_0 is also a $(C/\sqrt{n})$-Nash equilibrium. For brevity's sake, we will write the Markovian strategy profile as:

$$\Gamma_0 := (\gamma_0)_{i\in[n]}, \quad \gamma_0(t) := [\gamma_y^*(x, \Delta_x U_0(\cdot, \mu_t^{\Gamma_0,-x}))]_{x,y\in[d]}. \tag{3.9}$$

where $(\mu_t^{\Gamma_0,-x})_{t\ge0}$ is the empirical measure $(\mu_t^{\Gamma_0})_{t\ge0}$, but with one fewer player in state x.

Note that in (3.9), we substitute the empirical distribution into the solution of the master equation, whereas in (3.6), the strategy is based solely on the stationary MFG solution. This difference implies that the following theorem cannot be treated as a corollary and requires a distinct proof.

Theorem 3.2. Assume the full assumptions are in force. There exists $C > 0$, independent of n, such that the strategy profile Γ_0 is a Markovian $(C/\sqrt{n})$-Nash equilibrium. Moreover, the cost for each player in the game, $J_0^n(\Gamma_0)$, satisfies

$$|J_0^n(\Gamma_0) - \varrho| \le \frac{C}{\sqrt{n}}.$$

With π_0, the stationary distribution of $\boldsymbol{X}^{\Gamma_0}$, and for all $\pi \in \mathcal{P}([d]^n)$:

$$\limsup_{T\to\infty} \mathbb{E}_\pi|\mu_T^{\Gamma_0} - \bar\mu| = \mathbb{E}_{\pi_0}|\mu_0^{\Gamma_0} - \bar\mu| \le \frac{C}{\sqrt{n}}. \tag{3.10}$$

As with $\check\Gamma$ and $\bar\Gamma$, we allow a player deviating from Γ_0 to use full information about their system, but that this information will not benefit them by more than $C/\sqrt{n}$.

3.2. **Large deviations results.** We start with a remark about $\bar\mu$ and how it is an equilibrium for Kolmogorov's equation. Specifically, the large deviation results we provide require a *globally asymptotically stable equilibrium* and we flesh out precisely what we mean by this. Then, we introduce the laws featured in the large deviations result and the good rate functions for each law. *For all large deviation results, we assume the full assumptions are in force.*

Proposition 3.1. Let $(\bar\varrho, \bar u, \bar\mu)$ be the unique stationary solution to (1.4). Then, $\bar\mu$ is the unique globally asymptotically stable equilibrium for (2.3).

Remark 3.5. As a reminder, $\bar\mu$ is an equilibrium to (2.3) means that $\bar\mu$ satisfies:

$$0 = \sum_{y\in[d]} \bar\mu_y \gamma_x^*(y, \Delta_y U_0(\cdot, \bar\mu)),$$

that is, the derivative term in (2.3) is 0. Recalling that $U_0(\cdot, \bar\mu) = \bar u$ from Proposition 2.2, we see that this amounts to $\bar\mu$ solving the measure equation from the stationary ergodic MFG system (1.4). Finally, to be globally asymptotically stable means that for all initial conditions $\eta \in \mathcal{P}([d])$, the solution μ_0 to (2.3) tends to $\mu_0(t) \to \bar\mu$ as $t \to \infty$. For more about this definition, see [34, Section 3].

The main theorem deals with large deviations of the empirical measures. We will state the theorem in terms of $\kappa^n \in \{\mu^{\Gamma_0}, \mu^{\bar{\Gamma}}\}$ and the corresponding rates $a^\kappa \in \{\gamma_0, \bar{\gamma}\}$ from (3.9) and (3.6). We provide three large deviation results, for:

$$\mathscr{L}_T^n := \mathbb{P} \circ (\kappa_T^n)^{-1}, \quad \text{for} \quad \mathscr{L}_{[0,T]}^n := \mathbb{P} \circ ((\kappa_t^n)_{t \in [0,T]})^{-1},$$

and for $\mathscr{L}_\infty^n$, the unique invariant measure for the process κ^n For each law, there are good rate functions that we take an aside to define. So, define:

$$\tau(a) := e^a - a - 1,$$

and its Legendre transform:

$$\tau^*(a) := \begin{cases} (a+1)\log(a+1) - a, & a > -1, \\ 1, & a = -1, \\ +\infty, & a < -1. \end{cases}$$

Denote the set of all absolutely continuous functions $\mathbb{R}_+ \to \mathcal{P}([d])$ as $\mathcal{C}_{\mathrm{abs}}$ and for a few forthcoming matrices, we will denote the transpose of a matrix B by B^*. Define the subset:

$$\mathcal{C}_{\mathrm{abs}}^\eta := \Big\{ \mu \in \mathcal{C}_{\mathrm{abs}} \mid \mu(0) = \eta, \quad \dot{\mu}(t) = L(t)^* \mu(t), \quad \text{for some } L : \mathbb{R}_+ \to \mathcal{Q} \text{ is measurable} \Big\}.$$

Define also a function $S_{[0,T]}(\cdot \mid \eta) : \mathcal{C}_{\mathrm{abs}}^\eta \to \mathbb{R}$ as:

$$S_{[0,T]}(\mu \mid \eta) := \int_0^T \xi(t,y) a_{x,y}^\kappa(\mu(t)) \tau^* \Big(\frac{L_{x,y}(t)}{a_{x,y}^\kappa(\xi(t))} - 1 \Big) dt.$$

For $\xi \in \mathcal{P}([d])$, set $A_\xi := [a_{x,y}^\kappa(\xi)]_{x,y \in [d]}$; so $a_{x,y}^\kappa(\xi) = \gamma_y^*(x, \Delta_x U_0(\cdot, \xi))$ for $\kappa^n = \mu^{\Gamma_0}$, and $a_{x,y}^\kappa(\xi) = \bar{\gamma}_{xy} = \gamma_y^*(x, \Delta_x \bar{u})$ for $\kappa^n = \mu^{\bar{\Gamma}}$. So for $\kappa^n = \mu^{\bar{\Gamma}}$, $a^\kappa(\xi)$ is independent of ξ. As proved in [4, Theorem 3.2], $S_{[0,T]}(\mu \mid \eta)$ can be written in an equivalent form as:

$$S_{[0,T]}(\mu \mid \eta) = \int_0^T |||\dot{\mu}(t) - A_{\mu(t)}^* \mu(t)|||_{\mu(t)} dt, \tag{3.11}$$

where we adopt the notation from [4] for the function $|||R|||_\xi$ meaning:

$$|||R|||_\xi := \sup_{\phi:[d]\to\mathbb{R}} \Big[\sum_{x\in[d]} R_x \phi_x - \sum_{x,y\in[d]} \tau(\phi_y - \phi_x) \xi_x a_{x,y}^\kappa(\xi) \Big].$$

We note that $||| \cdot |||_\xi$ is not a norm. Through $S_{[0,T]}(\cdot \mid \eta)$ we can define $S_T(\cdot \mid \eta) : \mathcal{P}([d]) \to \mathbb{R}$ as:

$$S_T(\xi \mid \eta) := \inf \big\{ S_{[0,T]}(\mu \mid \eta) \mid \mu \in \mathcal{C}_{\mathrm{abs}}^\eta, \quad \mu(T) = \xi \big\}. \tag{3.12}$$

Finally, we can define:

$$s(\eta) := \inf_\mu \int_0^\infty \sum_{x,y\in[d]} \mu(t,y) a_{x,y}^\kappa(\mu(t)) \tau^* \Big(\frac{\hat{L}_{x,y}(t)}{a_{x,y}^\kappa(\mu(t))} - 1 \Big) dt, \tag{3.13}$$

and the infimum is taken over all $\mu : \mathbb{R}_+ \to \mathcal{P}([d])$ such that $\mu(0) = \eta$, $\lim_{t\to\infty} \mu(t) = \bar{\mu}$, and μ satisfies the reversed Kolmogorov equation:

$$\dot{\mu}(t) = -\hat{L}(t)\mu(t), \quad t \geq 0,$$

for some measurable $\hat{L} : \mathbb{R}_+ \to \mathcal{Q}$. For a reference on large deviations for particle systems with jumps, see [29].

Theorem 3.3 (Large Deviations). For all $T > 0$, the laws $\mathscr{L}_{[0,T]}^n$, $\mathscr{L}_T^n$, and $\mathscr{L}_\infty^n$ satisfy large deviation principles with speed n and good rate functions given by $S_{[0,T]}$ (3.11), S_T (3.12), and s (3.13), respectively. Explicitly, let E be any Borel subset of $\mathcal{P}([d])$, denote its interior and closure by $\mathrm{int}(E)$ and $\mathrm{cls}(E)$, respectively, and we have:

$$-\inf_{\eta\in\mathrm{int}(E)} S_T(\eta \mid \bar{\mu}) \leq \liminf_{n\to\infty} n^{-1} \log(\mathscr{L}_T^n(E)) \leq \limsup_{n\to\infty} n^{-1} \log(\mathscr{L}_T^n(E)) \leq -\inf_{\eta\in\mathrm{cls}(E)} S_T(\eta \mid \bar{\mu}),$$

and
$$-\inf_{\eta\in\mathrm{int}(E)} s(\eta) \le \liminf_{n\to\infty} n^{-1}\log(\mathscr{L}^n_\infty(E)) \le \limsup_{n\to\infty} n^{-1}\log(\mathscr{L}^n_\infty(E)) \le -\inf_{\eta\in\mathrm{cls}(E)} s(\eta).$$

Now let G be a Borel set in the Skorohod topology of the space $\mathcal{C}^{\bar\mu}_{\mathrm{abs}}$. We also have:
$$-\inf_{\eta\in\mathrm{int}(G)} S_{[0,T]}(\eta \mid \bar\mu) \le \liminf_{n\to\infty} n^{-1}\log(\mathscr{L}^n_{[0,T]}(G))$$
$$\le \limsup_{n\to\infty} n^{-1}\log(\mathscr{L}^n_{[0,T]}(G)) \le -\inf_{\eta\in\mathrm{cls}(G)} S_{[0,T]}(\eta \mid \bar\mu).$$

4. Proofs for Theorem 3.1

In this section, we give the proof for Theorem 3.1 concerning the asymptotic Nash equilibria derived from the time-dependent ergodic MFG (1.3). The first lemma we require is Lemma 4.1, which gives a concentration inequality for players using a symmetric Markovian strategy profile when they are initialized the same way. The second, Lemma 4.2, compares an empirical measure under an arbitrary Markovian strategy profile to $\check\mu$. The bound we derive will depend on the time t, the initial conditions of each system, the number of players n, and the difference between the strategies. The two prior lemmas combine to obtain the propagation of chaos estimate (3.4) that allows us to compare mean field costs F at one point in the proof of the theorem.

Throughout this section, we require the reduced assumptions only. In the forthcoming proofs, C and c will be positive constants that may change value from one line to the next. Whenever C and c appear in the proofs, they depend only on the problem data and we emphasize that they are always independent of n, t, and any initialization of the players. The concentration inequality lemma is:

Lemma 4.1. Let $\boldsymbol{\beta}$ be a symmetric Markovian strategy profile with $\boldsymbol{\beta} := (\beta^i)_{i\in[n]}$, $\beta^i \in \mathcal{A}^{n,i}$, and let $\boldsymbol{X}^{\boldsymbol{\beta}}$ be the jump process of all players using $\boldsymbol{\beta}$. Suppose further that $(X_0^{\boldsymbol{\beta},i})_{i\in[n]}$ are exchangeable though not necessarily independent. So, all processes are initialized the same way; that is, for all $i \in [n]$, $\mathcal{L}(X_0^{\boldsymbol{\beta},i}) = \mathcal{L}(X_0^{\boldsymbol{\beta},1})$. Let $\pi^{\boldsymbol{\beta}} \in \mathcal{P}([d]^n)$ be the joint distribution induced by this initialization. Then, for all $i \in [n]$ and for all $t \in \mathbb{R}_+$, $\mathcal{L}(X_t^{\boldsymbol{\beta},i}) = \mathcal{L}(X_t^{\boldsymbol{\beta},1})$. Assume $\mathcal{L}(X_0^{\boldsymbol{\beta},1})$ is not on the boundary of $\mathcal{P}([d])$. Then there exists $q > 0$ depending only on $\mathbb{A}$ such that for all $t \in \mathbb{R}_+$, and all $i \in [n]$,
$$\mathcal{L}(X_t^{\boldsymbol{\beta},i})_x > q > 0 \quad\text{and}\quad 1 - \mathcal{L}(X_t^{\boldsymbol{\beta},i})_x > q > 0 \quad\text{for all}\quad x \in [d].$$
Moreover, there exists $C > 0$, independent of n, so that:
$$\sup_{t\in\mathbb{R}_+} \mathbb{P}_{\pi^{\boldsymbol{\beta}}}\left(\left|\frac{1}{n}\sum_{j\in[n]} \delta_{X^{\boldsymbol{\beta},j}}(x) - \mathcal{L}(X_t^{\boldsymbol{\beta},1})_x\right| \ge y\right) \le 2e^{-ny^2/4q}, \tag{4.1}$$
$$\sup_{t\in\mathbb{R}_+} \mathbb{E}_{\pi^{\boldsymbol{\beta}}}|\mu_t^{\boldsymbol{\beta}} - \mathcal{L}(X_t^{\boldsymbol{\beta},i})| \le \frac{C}{\sqrt{n}}. \tag{4.2}$$

When $\mathcal{L}(X_0^{\boldsymbol{\beta},1})$ is on the boundary of $\mathcal{P}([d])$, there exists such q for $t \ge 1$ and so we must modify the suprema in (4.1) and (4.2) to be over all $t \ge 1$.

Also let $(\check\varrho, \check u, \check\mu)$ be a solution to the time-dependent ergodic MFG system (1.3). As formulated in the statement of the theorem, recall that $\check\gamma_{xy}(t) := \gamma_y^*(x, \Delta_x\check u(t))$ and set $\check\gamma_x := (\check\gamma_{xy})_{y\in[d]}$, $\check\gamma := (\check\gamma_x)_{x\in[d]}$, and $\check\Gamma := (\check\gamma)_{i=1}^n$. Note that since (1.3) is the time-dependent system, the control $\check\gamma$ depends implicitly on time through the potential function $\check u$. The following lemma is helpful for the proof of Theorem 3.1 and for the proof of Theorem 3.2.

Lemma 4.2. Fix $\eta \in \mathcal{P}([d])$ and let $(\check\varrho, \check u, \check\mu)$ solve (1.3) with $\check\mu(0) = \eta$. There exist C, c positive constants such that for any $n \in \mathbb{N}_{>1}$ and any $\boldsymbol{\beta} \in \mathcal{A}^{(n)}$, any initialization $\pi \in \mathcal{P}([d]^n)$ with $\boldsymbol{X}_0^{\boldsymbol{\beta}} \sim \pi$, and any $t \in \mathbb{R}_+$,
$$\mathbb{E}_\pi|\mu_t^{\boldsymbol{\beta}} - \check\mu(t)| \le Ce^{-ct}\mathbb{E}_\pi|\mu_0^{\boldsymbol{\beta}} - \check\mu(0)| + \frac{C}{\sqrt{n}}$$
$$+ \mathbb{E}_\pi \int_0^t e^{-c(t-s)} \sum_{x\in[d]} \mu_{s,x}^{\boldsymbol{\beta}}\left|\frac{1}{n}\sum_{i\in[n]} \left[\beta_x^i(s, \boldsymbol{X}_s^{\boldsymbol{\beta}}) - \check\gamma_x(t)\right]\mathbb{1}_{\{X_s^{\boldsymbol{\beta},i}=x\}}\right| ds. \tag{4.3}$$

Suppose $(X_0^{\boldsymbol{\beta},i})_{i \in [n]}$ are exchangeable, though not necessarily independent, with $X_0^{\boldsymbol{\beta},i} \sim \check{\mu}(0)$ for all $i \in [n]$. For $\pi = \check{\pi}$, the joint distribution induced by this initialization, we have from (4.3) that

$$\mathbb{E}_{\check{\pi}} |\mu_t^{\boldsymbol{\beta}} - \check{\mu}(t)| \leq \frac{C}{\sqrt{n}} + \mathbb{E}_{\check{\pi}} \int_0^t e^{-c(t-s)} \sum_{x \in [d]} \mu_{s,x}^{\boldsymbol{\beta}} \Big| \frac{1}{n} \sum_{i \in [n]} \big[\beta_x^i(s, \boldsymbol{X}_s^{\boldsymbol{\beta}}) - \check{\gamma}_x(t) \big] \mathbb{1}_{\{X_s^{\boldsymbol{\beta},i}=x\}} \Big| ds. \tag{4.4}$$

The proofs of all lemmas are provided in the subsequent part of this section, following the conclusion of the proof of Theorem 3.1.

4.1. **Proof of Theorem 3.1.**
We split the proof into two parts. In the first part, we prove the propagation of chaos results (3.4) and (3.5). We will then use these results in the second part to prove that $\check{\Gamma}$ is a $(C/\sqrt{n})$-Nash equilibrium and (3.3).

4.1.1. *Propagation of chaos.*
To prove (3.4) and (3.5), we start by proving a related result to (3.4). We will show there exist C, c positive constants such that for any $n \in \mathbb{N}_{>1}$, any $\beta \in \mathcal{A}^{n,i}$, and any initialization $\pi \in \mathcal{P}([d]^n)$ with $\boldsymbol{X}_0^{[\check{\Gamma}^{-i};\beta]} \sim \pi$, and any $t \in \mathbb{R}_+$,

$$\mathbb{E}_\pi |\mu_t^{[\check{\Gamma}^{-i};\beta]} - \check{\mu}(t)| \leq Ce^{-ct} \mathbb{E}_\pi |\mu_0^{[\check{\Gamma}^{-i};\beta]} - \check{\mu}(0)| + \frac{C}{\sqrt{n}}.$$

First note that $\mu^{[\check{\Gamma}^{-i};\beta],i} = \mu^{\check{\Gamma},i}$ since the i-th player is removed from both processes and since the strategy $\check{\gamma}$ is independent of the players' process $\boldsymbol{X}^{[\check{\Gamma}^{-i};\beta]}$. Moreover, since

$$\mu^{\check{\Gamma}} = n^{-1}\big((n-1)\mu^{\check{\Gamma},i} + \delta_{X^{\check{\Gamma},i}}\big),$$

the measures $\mu^{\check{\Gamma}}$ and $\mu^{\check{\Gamma},i}$ differ by a term of order $\mathcal{O}(n^{-1})$. So, it suffices to show the main estimate for $\mu^{\check{\Gamma}}$. Applying Lemma 4.2 for $\boldsymbol{\beta} = \check{\Gamma}$, we note that $\beta^i = \check{\gamma}$ for every $i \in [n]$ and the integral term is zero, yielding:

$$\mathbb{E}_\pi |\mu_t^{\check{\Gamma}} - \check{\mu}(t)| \leq Ce^{-ct} \mathbb{E}_\pi |\mu_0^{\check{\Gamma}} - \check{\mu}(0)| + \frac{C}{\sqrt{n}}.$$

We recognize the prior display as (3.4). To get (3.5), when $\pi = \check{\pi}$, we note that the term in expectation on the right-hand side is no more than $C/\sqrt{n}$ by Lemma 4.1.

4.1.2. *Approximate Markovian Nash Equilibrium.*
We can now use the propagation of chaos results to obtain the remaining points of Theorem 3.1. We abuse notation slightly by writing $X^{\beta,i}$ for $X^{[\check{\Gamma}^{-i};\beta],i}$, keeping in mind that the dynamics of player i are implicitly dependent on the other players' dynamics, through player i's strategy β, and the rest of the players' strategy profile.

By Lemma 3.1 we can consider $J_0^{n,i}$ with any initial distribution for the agents. In the following display, we make use of this fact to study the value when a single player, player i, deviates from the strategy $\check{\gamma}$. Player i will use $\beta \in \mathcal{A}^{n,i}$ instead. Let $\check{\pi} \in \mathcal{P}([d]^n)$ be the joint distribution induced by initializing every player according to $\check{\mu}(0) \in \mathcal{P}([d])$. Then, adding and subtracting a term:

$$\begin{aligned}
J_0^{n,i}([\check{\Gamma}^{-i}; \beta]) &= \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\check{\pi}} \int_0^T [f(X_t^{\beta,i}, \beta(t, \boldsymbol{X}_t^{[\check{\Gamma}^{-i};\beta]})) + F(X_t^{\beta,i}, \mu_t^{[\check{\Gamma}^{-i};\beta],i})]dt \\
&= \limsup_{T \to \infty} \Big(\frac{1}{T} \mathbb{E}_{\check{\pi}} \int_0^T [f(X_t^{\beta,i}, \beta(t, \boldsymbol{X}_t^{[\check{\Gamma}^{-i};\beta]})) + F(X_t^{\beta,i}, \check{\mu}(t))]dt \\
&\qquad + \frac{1}{T} \mathbb{E}_{\check{\pi}} \int_0^T [F(X_t^{\beta,i}, \mu_t^{[\check{\Gamma}^{-i};\beta],i}) - F(X_t^{\beta,i}, \check{\mu}(t))]dt \Big).
\end{aligned} \tag{4.5}$$

We will consider each term on the right-hand side of (4.5) separately. For the first term, we note two things: (1) that we initialized $X_0^{\beta,i} \sim \check{\mu}(0)$ and (2) that $\check{\mu}$ is an MFE by Proposition 2.2. Then by Definition 2.3 and the definition of J_0, the strategy β that minimizes this function must be $\check{\gamma}$. Hence,

$$\inf_{\beta \in \mathcal{A}} \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\check{\pi}} \int_0^T [f(X_t^{\beta,i}, \beta(t, \boldsymbol{X}_t^{[\check{\Gamma}^{-i};\beta]})) + F(X_t^{\beta,i}, \check{\mu}(t))]dt = J_0(\check{\gamma}, \check{\mu}). \tag{4.6}$$

We will keep this fact aside for the moment and return to (4.5) in order to deal with the second term on the right-hand side. We note once more that each of the players has an initial state with law $\check{\mu}(0)$. Bringing the absolute value inside the integral, using that F is Lipschitz, using Fubini's theorem to exchange to the expectation and integral, and by $\mu^{[\check{\Gamma}^{-i};\beta],i} = \mu^{\check{\Gamma},i}$ with (3.5), we can bound the second term on the right-hand side of (4.5):

$$\Big| \frac{1}{T} \mathbb{E}_{\check{\pi}} \int_0^T [F(X_t^{\beta,i}, \mu_t^{[\check{\Gamma}^{-i};\beta],i}) - F(X_t^{\beta,i}, \check{\mu}(t))]dt \Big| \leq \frac{C}{\sqrt{n}}. \tag{4.7}$$

So using (4.5) and (4.7), followed by (4.6),

$$J_0^n(\check{\Gamma}) - J_0^{n,i}([\check{\Gamma}^{-i};\beta])$$

$$= J_0(\check{\gamma}, \check{\mu}) - \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\check{\pi}^n} \int_0^T [f(X_t^{\beta,i}, \beta(t, \boldsymbol{X}_t^{[\check{\Gamma}^{-i};\beta]})) + F(X_t^{\beta,i}, \check{\mu}(t))]dt + \mathcal{O}(n^{-1/2}) \tag{4.8}$$

$$\leq \mathcal{O}(n^{-1/2}).$$

Since β was arbitrary, $\check{\Gamma}$ is a $(C/\sqrt{n})$-Nash equilibrium. It only remains to show (3.3) and by (4.5), (4.6), (4.7), it suffices to show that $J_0(\check{\gamma}, \check{\mu}) = \check{\varrho}$. This follows immediately from follows from Proposition 2.2. □

4.2. **Proof of Lemma 4.1.** Denote $\kappa(t) := \mathcal{L}(X_t^{\beta,1})$. By assumption, all players are initialized independently according to $\kappa(0)$. Since β is symmetric, we claim the players are exchangeable with the same law $\kappa(t)$ for all $t \in \mathbb{R}_+$. That is, for all $j \in [n]$ and all $t \in \mathbb{R}_+$, we claim $\mathcal{L}(X_t^{\beta,j}) = \kappa(t)$.

Now, since the rates $\mathbb{A}$ are bounded away from zero and since $\mathcal{L}(X_t^{\beta,i}) = \kappa(t)$, we must have that there exists $q > 0$ such that for all $x \in [d]$ and all $t > 1$, $\mathcal{L}(X_t^{\beta,i})_x > q > 0$, even when $\kappa(0)$ is on the boundary. The choice of time 1 for this threshold is not crucial to the proof, any positive constant would do. By symmetry of the players, for any $t \in \mathbb{R}_+$, $(\delta_{X_t^{\beta,i}})_{i \in [n]}$ are exchangeable with the same marginal law $\kappa(t)$. Fix $x \in [d]$. Then $(\delta_{X_t^{\beta,i}}(x))_{i \in [n]}$ are exchangeable Bernoulli random variables, each with probability $\kappa_x(t)$ to be 1. By [13, Corollary 2.2] , we get (4.1) immediately. Using (4.1), and by definition of the expectation,

$$\mathbb{E}_{\pi^\beta}|\mu_{t,x}^{\beta} - \kappa_x(t)| = \int_0^\infty \mathbb{P}_{\pi^\beta}\Big(\Big| \frac{1}{n} \sum_{i \in [n]} \delta_{X_t^{\beta,i}}(x) - \kappa_x(t) \Big| \geq y\Big)dy \leq \int_0^\infty 2e^{-ny^2/4q}dy = 2\sqrt{\frac{\pi q}{n}},$$

which establishes (4.2). □

4.3. **Proof of Lemma 4.2.** Recall the definition of $\check{\Gamma}$ from Theorem 3.1. Define the transition matrix $Q(t)$ and its inverse by:

$$Q(t) := \exp\Big(\int_0^t \check{\Gamma}(s)ds \Big) \quad \text{and} \quad Q^{-1}(t) := \exp\Big(-\int_0^t \check{\Gamma}(s)ds \Big).$$

Also define:

$$\Gamma^{\beta}(s) := \Big(\frac{1}{n} \sum_{i \in [n]} \beta_{xy}^i(s, \boldsymbol{X}_s^{\beta}) \mathbb{1}_{\{X_s^{\beta,i}=x\}} \Big)_{x,y \in [d]}.$$

Note that for any time s and any realization of $\boldsymbol{X}_s^{\beta}$, we have $\Gamma_s^{\beta} \in \mathcal{Q}$. From the dynamics of μ^{β} and since $\check{\mu}$ solves Kolmogorov's equation from (1.3), we get that:

$$(\mu_s^{\beta} - \check{\mu}(s))^T Q^{-1}(s)\Big|_0^t$$

$$= \int_0^t (\mu^{\beta}(s)^T \Gamma^{\beta}(s) - \check{\mu}(s)^T \check{\Gamma}(s))Q^{-1}(s)ds + \int_0^t (\mu_s^{\beta} - \check{\mu}(s))^T[-\check{\Gamma}(s)Q^{-1}(s)]ds$$

$$+ \frac{1}{n} \sum_{x,y \in [d]} \int_0^t (e_y - e_x)Q^{-1}(s^-)\Big(dK_{xy}^{\beta}(s) - \sum_{i \in [n]} \beta_{xy}^i(s, \boldsymbol{X}_s^{\beta})\mathbb{1}_{\{X_s^{\beta,i}=x\}}ds\Big),$$

where $K^{\boldsymbol{\beta}}$ is given in (3.2). Multiplying both sides by $Q(t)$ and cancelling the terms with $\check{\mu}(s)^T \check{\Gamma}(s) Q^{-1}(s)$ from the first two integrals, we get,

$$
\begin{aligned}
(\mu_t^{\boldsymbol{\beta}} &- \check{\mu}(t)) - (\mu_0^{\boldsymbol{\beta}} - \check{\mu}(0))Q(t) \\
&= \int_0^t (\mu_s^{\boldsymbol{\beta}})^T (\Gamma^{\boldsymbol{\beta}}(s) - \check{\Gamma}(s))Q^{-1}(s)Q(t)ds + \frac{1}{n} \sum_{x,y \in [d]} \int_0^t (e_y - e_x)Q^{-1}(s^-)Q(t)dM_{xy}^{\boldsymbol{\beta}}(s),
\end{aligned}
\tag{4.9}
$$

where $M_{xy}^{\boldsymbol{\beta}}$ is a martingale given by,

$$
M_{xy}^{\boldsymbol{\beta}}(t) := \left(K_{xy}^{\boldsymbol{\beta}}(t) - \int_0^t \sum_{i \in [n]} \beta_{xy}^i(s, \boldsymbol{X}_s^{\boldsymbol{\beta}}) \mathbb{1}_{\{X_s^{\boldsymbol{\beta},i}=x\}} ds \right).
$$

Let $R(s,t) := (e_y - e_x)Q^{-1}(s)Q(t)$. For the following display, we use: the Burkholder–Davis–Gundy inequality; that the quadratic variation of a compensated Poisson process is the original process, so $\langle M_{xy}^{\boldsymbol{\beta}}(s) \rangle = K_{xy}^{\boldsymbol{\beta}}(s)$; Jensen's inequality for the concave function $\sqrt{\cdot}$; adding and subtracting the compensator for $K_{xy}^{\boldsymbol{\beta}}(s)$ and using that the martingale term has expectation zero; the boundedness of the rates β^i; and finally $R(s,t)R(s,t)^T = |R(s,t)|^2$ with [12, Lemma 3.3]:

$$
\begin{aligned}
\mathbb{E}\left| \int_0^t R(s^-,t)dM_{xy}^{\boldsymbol{\beta}}(s) \right| &\leq C\mathbb{E}\left[\left(\int_0^t R(s^-,t)R(s^-,t)^T \langle dM_{xy}^{\boldsymbol{\beta}}(s) \rangle \right)^{1/2} \right] \\
&= C\mathbb{E}\left[\left(\int_0^t R(s^-,t)R(s^-,t)^T dK_{xy}^{\boldsymbol{\beta}}(s) \right)^{1/2} \right] \\
&\leq C\left[\mathbb{E}\left(\int_0^t R(s^-,t)R(s^-,t)^T dK_{xy}^{\boldsymbol{\beta}}(s) \right) \right]^{1/2} \\
&= C\left[\mathbb{E}\left(\int_0^t R(s,t)R(s,t)^T \sum_{i \in [n]} \beta_{xy}^i(s, \boldsymbol{X}_s^{\boldsymbol{\beta}}) \mathbb{1}_{\{X_s^{\boldsymbol{\beta},i}=x\}} ds \right) \right]^{1/2} \\
&\leq Cn^{1/2}\left[\int_0^t R(s,t)R(s,t)^T ds \right]^{1/2} \\
&\leq Cn^{1/2}.
\end{aligned}
\tag{4.10}
$$

Also, let $\tilde{\mu}(t) := (\mu_0^{\boldsymbol{\beta}})^T Q(t)$ and note that while $Q(t)$ is deterministic, $\tilde{\mu}(t)$ is random since $\mu_0^{\boldsymbol{\beta}}$ is random. Note that $\check{\mu}(t) = (\check{\mu}(0))^T Q(t)$ too. So by [12, Lemma 3.3],

$$
|\tilde{\mu}(t) - \check{\mu}(t)| \leq Ce^{-ct}|\mu_0^{\boldsymbol{\beta}} - \check{\mu}(0)|.
$$

Taking the expectation,

$$
\mathbb{E}_\pi |(\mu_0^{\boldsymbol{\beta}} - \check{\mu}(0))Q(t)| \leq Ce^{-ct}\mathbb{E}_\pi |\mu_0^{\boldsymbol{\beta}} - \check{\mu}(0)|.
\tag{4.11}
$$

We now return to (4.9) and, first taking the expectation, use the martingale bound (4.10) along with (4.11) to get that:

$$
\mathbb{E}_\pi |\mu_t^{\boldsymbol{\beta}} - \check{\mu}(t)| \leq Ce^{-ct}\mathbb{E}_\pi |\mu_0^{\boldsymbol{\beta}} - \check{\mu}(0)| + \frac{C}{\sqrt{n}} + \mathbb{E}_\pi \left| \int_0^t (\mu_s^{\boldsymbol{\beta}})^T (\Gamma^{\boldsymbol{\beta}}(s) - \check{\Gamma}(s))Q^{-1}(s)Q(t)ds \right|.
$$

So applying triangle inequality, again using [12, Lemma 3.3], and expanding the matrix notation as a sum finishes the proof of (4.3). For (4.4), we note that when $\pi = \check{\pi}$, the first term on the right-hand side is no more than $C/\sqrt{n}$ by Lemma 4.1. $\qquad \square$

5. Proofs for Theorem 3.2

In addition to Lemma 4.2 from the prior section, we need two lemmas to approach the proof of Theorem 3.2. The first, Lemma 5.1, introduces a mock ergodic system with a solution (ϱ^i, W_0^i) meant to mimic the optimal value and associated potential for the deviating i-th player, optimizing against the other players, who are each using γ_0. Through this detour we prove that the deviating player's optimal

strategy is actually stationary. This is helpful to reduce the problem of proving Γ_0 is a Markovian $(C/\sqrt{n})$-Nash equilibrium from studying all strategies $\mathcal{A}^{n,i}$ for player i to just stationary strategies $\mathcal{A}_S^{n,i}$.

The second lemma, Lemma 5.2, is modeled after duality estimates originally used for the MFG system. Its estimate is crucial to avoiding a consequence of the deleterious feedback effect: to compare strategies, we must compare processes that we plug into them; and to compare processes, we must compare their strategies. The proof of the lemma relies on the full assumptions as well; namely, we will need that F is Lasry–Lions monotone and that H is concave.

In the forthcoming proofs, C and c will be positive constants that may change value from one line to the next. Whenever C and c appear in the proofs, they depend only on the problem data and we emphasize that they are always independent of n, t and any initialization of the players. Recall that for Theorem 3.2, the full assumptions hold, and will be in force throughout this section. Consequently, the entirety of Proposition 2.1 and Proposition 2.2 hold. We will let $(\varrho, \check{u}, \check{\mu})$ be the unique solution to (1.3) and let $(\varrho, \bar{u}, \bar{\mu})$ be the unique solution to (1.4).

For any $\boldsymbol{x} \in [d]^n$ we define the empirical distribution with and without the i-th player as:
$$\mu_{\boldsymbol{x}}^n := \frac{1}{n} \sum_{j \in [n]} e_{x^j}, \qquad \mu_{\boldsymbol{x}}^{n,i} := \frac{1}{n} \sum_{j \in [n], j \neq i} e_{x^j}.$$

Lemma 5.1. Based on γ_0 from (3.9) define $\gamma_0^j(\boldsymbol{x}) := \gamma^*(x^j, \Delta_{x^j} U_0(\cdot, \mu_{\boldsymbol{x}}^{n,j}))$. There exist $\varrho^i \in \mathbb{R}$, $W_0^i : [d] \times \mathcal{P}^{n-1}([d]) \to \mathbb{R}$ such that (ϱ^i, W_0^i) satisfies:
$$\varrho^i = H(x^i, \Delta_{x^i} W_0^i(\cdot, \mu_{\boldsymbol{x}}^{n,i})) + F(x^i, \mu_{\boldsymbol{x}}^{n,i}) + \sum_{j \neq i} \gamma_0^j(\boldsymbol{x}) \cdot \Delta_{x^j} U_0(\cdot, \mu_{\boldsymbol{x}}^{n,j}). \tag{5.1}$$

Moreover, for $\gamma_y^W(\boldsymbol{x}) := \gamma_y^*(x^i, \Delta_{x^i} W_0^i(\cdot, \mu_{\boldsymbol{x}}^{n,i}))$, $\gamma^W := (\gamma_y^W)_{y \in [d]}$, we have that:
$$\varrho^i = J_0^{n,i}([\Gamma_0^{-i}; \gamma^W]) = \inf_{\beta \in \mathcal{A}^{n,i}} J_0^{n,i}([\Gamma_0^{-i}; \beta]). \tag{5.2}$$

Interpreting (5.2) in words, a stationary Markovian strategy is optimal for a player deviating from Γ_0.

Lemma 5.2. There exists $C > 0$, independent of n, such that for all $\beta \in \mathcal{A}^{n,i}$ and all $\pi \in \mathcal{P}([d]^n)$:
$$\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} (\mu_{t,x}^{[\Gamma_0^{-i}; \beta]} + \check{\mu}_x(t)) |\Delta_x U_0(\cdot, \mu_t^{[\Gamma_0^{-i}; \beta]}) - \Delta_x \check{u}(t)|^2 dt \leq \frac{C}{n}.$$

5.1. **Proof of Theorem 3.2.** As with Theorem 3.1, we split the proof into two halves. The first part proves the propagation of chaos given in (3.10). For the propagation of chaos, we start by using Lemma 4.2 to pass off the problem of comparing processes to comparing their strategies. The strategies are Lipschitz, so it suffices to compare their different potential functions. Then we use the duality-style lemma, Lemma 5.2, to show that the integral term left over from the application of Lemma 4.2, and that is comprised of these potential functions, is small. Along the way, we will need to use that the stationary distribution π^β forces the processes inside the expectation to be independent of time; this helps us periodically move the processes outside of pesky, time-dependent integrals.

In order to use the propagation of chaos to prove the profile is an asymptotic Nash equilibrium, we must start by restricting the proof to stationary strategies. Otherwise, we would not necessarily have a stationary distribution and the prior part would not apply. After we apply propagation of chaos and obtain the result for stationary strategies, we can use Lemma 5.1 to upgrade the result so that it holds for any Markovian strategy.

5.1.1. *Propagation of chaos.* In this part we will prove:
$$\mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i}; \beta^S]} - \bar{\mu}| \leq \frac{C}{\sqrt{n}}, \tag{5.3}$$

for $\beta^S \in \mathcal{A}_S^{n,i}$, an arbitrary stationary Markovian strategy and $\pi^\beta \in \mathcal{P}([d]^n)$, the unique stationary distribution for the jump process $\boldsymbol{X}^{[\Gamma_0^{-i}; \beta^S]}$. Note that (5.3) is more general than (3.10). The full

strength of (5.3) is necessary for the latter part of the proof, where we prove Γ_0 is an asymptotic Nash equilibrium, and so deal with a deviating player.

Recall that Γ_0 is comprised of γ_0 defined in (3.9) and recall $\bar{\gamma}$ from (3.6). We now apply Lemma 4.2 with $\boldsymbol{\beta} = [\Gamma_0^{-i}; \beta^S]$ and $\pi = \pi^\beta$, use the fact that the rates are bounded to aggregate another term into $C/\sqrt{n}$, and then use that $(\frac{n-1}{n})(\mu_{s,x}^{[\Gamma_0^{-i};\beta^S],i}) \leq 1$ to get:

$$
\begin{aligned}
\mathbb{E}_{\pi^\beta} |\mu_t^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| &\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}} \\
&\quad + \mathbb{E}_{\pi^\beta} \int_0^t e^{cs-ct} \sum_{x\in[d]} \mu_{s,x}^{[\Gamma_0^{-i};\beta^S]} \Big| \big(\tfrac{n-1}{n}(\mu_{s,x}^{[\Gamma_0^{-i};\beta^S],i})[\gamma_{0,yx}(s) - \bar{\gamma}_{yx}] \\
&\qquad\qquad + \tfrac{1}{n}\mathbb{1}_{\left\{X_s^{[\Gamma_0^{-i};\beta^S],i}=x\right\}} [\beta_{yx}^S(\boldsymbol{X}_s^{[\Gamma_0^{-i};\beta^S]}) - \bar{\gamma}_{yx}]\big)_{y\in[d]} \Big| ds \\
&\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}} \\
&\quad + \mathbb{E}_{\pi^\beta} \int_0^t e^{cs-ct} \sum_{x\in[d]} \mu_{s,x}^{[\Gamma_0^{-i};\beta^S]} \big| (\tfrac{n-1}{n})(\mu_{s,x}^{[\Gamma_0^{-i};\beta^S],i})[\gamma_{0,yx}(s) - \bar{\gamma}_{yx}]\big| ds \\
&\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}} \\
&\quad + \mathbb{E}_{\pi^\beta} \int_0^t e^{cs-ct} \sum_{x\in[d]} \mu_{s,x}^{[\Gamma_0^{-i};\beta^S]} |\gamma_{0,yx}(s) - \bar{\gamma}_{yx}| ds.
\end{aligned}
$$

Now, leveraging the fact that π^β represents the stationary distribution, along with the definitions of γ_0 and $\bar{\gamma}$ in terms of γ^*, and considering the Lipschitz continuity of γ^*, we can derive:

$$
\begin{aligned}
\mathbb{E}_{\pi^\beta} |\mu_t^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| &\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}} \\
&\quad + C \mathbb{E}_{\pi^\beta} \Big[\sum_{x\in[d]} \mu_{0,x}^{[\Gamma_0^{-i};\beta^S]} \big| \Delta_x(U_0(\cdot, \mu_0^{[\Gamma_0^{-i};\beta^S]}) - \bar{u})\big| \Big].
\end{aligned}
\tag{5.4}
$$

We can also use that π^β is the stationary distribution, then Lemma 3.1 and then Lemma 5.2 to get:

$$
\begin{aligned}
\mathbb{E}_{\pi^\beta} \Big[\sum_{x\in[d]} (\mu_{0,x}^{[\Gamma_0^{-i};\beta^S]} &+ \bar{\mu}_x) \big| \Delta_x(U_0(\cdot, \mu_0^{[\Gamma_0^{-i};\beta^S]}) - \bar{u})\big|^2 \Big] \\
&= \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}_{\pi^\beta} \int_0^T \Big[\sum_{x\in[d]} (\mu_{t,x}^{[\Gamma_0^{-i};\beta^S]} + \bar{\mu}_x) \big| \Delta_x(U_0(\cdot, \mu_t^{[\Gamma_0^{-i};\beta^S]}) - \bar{u})\big|^2 dt \tag{5.5} \\
&\leq \frac{C}{n}.
\end{aligned}
$$

We can use Jensen's inequality for $\sqrt{\cdot}$ and that for any $\eta \in \mathcal{P}([d])$, $\eta_x^2 \leq \eta_x$ for all $x \in [d]$ in (5.4) and then (5.5) to get:

$$
\begin{aligned}
\mathbb{E}_{\pi^\beta} |\mu_t^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| &\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}} \\
&\quad + C \Big(\mathbb{E}_{\pi^\beta} \Big[\sum_{x\in[d]} \mu_{0,x}^{[\Gamma_0^{-i};\beta^S]} \big| \Delta_x(U_0(\cdot, \mu_0^{[\Gamma_0^{-i};\beta^S]}) - \bar{u})\big|^2 \Big] \Big)^{1/2} \\
&\leq C e^{-ct} \mathbb{E}_{\pi^\beta} |\mu_0^{[\Gamma_0^{-i};\beta^S]} - \bar{\mu}| + \frac{C}{\sqrt{n}}.
\end{aligned}
$$

We applied the stationary distribution on the left-hand side above as well, so it is independent of t. Passing the limit as $t \to \infty$ then yields (5.3). When $\beta = \gamma_0$ then, we get (3.10) where the limiting result follows by ergodicity of the Markov process $\mu_t^{\Gamma_0}$.

5.1.2. *Approximate Markovian Nash Equilibrium.* Now we prove the main part of the theorem. By a small abuse of notation, we will write $X^{[\Gamma_0^{-i};\beta^S],i}$ as $X^{\beta^S,i}$, keeping in mind that the position of the i-th player will depend on Γ_0 implicitly since β^S depends on the positions of the other players. Again we rely on *Lemma* 3.1 for the freedom to choose the initialization of the players. We will see that it is particularly convenient to use π^β, the stationary distribution. By the definition of $J_0^{n,i}$, using that π^β is the stationary distribution, adding and subtracting a term, and once more using that π^β is the stationary distribution:

$$
\begin{aligned}
J_0^{n,i}&([\Gamma_0^{-i};\beta^S]) \\
&= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{\pi^\beta} \int_0^T [f(X_t^{\beta^S,i}, \beta^S(\boldsymbol{X}_t^{[\Gamma_0^{-i};\beta^S]})) + F(X_t^{\beta^S,i}, \mu_t^{[\Gamma_0^{-i};\beta^S],i})]dt \\
&= \mathbb{E}_{\pi^\beta}\Big[f(X_0^{\beta^S,i}, \beta^S(\boldsymbol{X}_0^{[\Gamma_0^{-i};\beta^S]})) + F(X_0^{\beta^S,i}, \mu_0^{[\Gamma_0^{-i};\beta^S],i})\Big] \\
&= \mathbb{E}_{\pi^\beta}\Big[f(X_0^{\beta^S,i}, \beta^S(\boldsymbol{X}_0^{[\Gamma_0^{-i};\beta^S]})) + F(X_0^{\beta^S,i}, \bar{\mu}) + F(X_0^{\beta^S,i}, \mu_0^{[\Gamma_0^{-i};\beta^S],i}) - F(X_0^{\beta^S,i}, \bar{\mu})\Big] \\
&= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{\pi^\beta} \int_0^T [f(X_t^{\beta^S,i}, \beta^S(\boldsymbol{X}_t^{[\Gamma_0^{-i};\beta^S]})) + F(X_t^{\beta^S,i}, \bar{\mu})]dt \\
&\quad + \mathbb{E}_{\pi^\beta}\Big[F(X_0^{\beta^S,i}, \mu_0^{[\Gamma_0^{-i};\beta^S],i}) - F(X_0^{\beta^S,i}, \bar{\mu})\Big].
\end{aligned}
\tag{5.6}
$$

For the second term in (5.6), we can use the Lipschitz continuity of F, followed by (5.3) to get:

$$
\begin{aligned}
\left|\mathbb{E}_{\pi^\beta}\Big[F(X_0^{\beta^S,i}, \mu_0^{[\Gamma_0^{-i};\beta^S],i}) - F(X_0^{\beta^S,i}, \bar{\mu})\Big]\right| &\le C_{L,F}\mathbb{E}_{\pi^\beta}|\mu_0^{[\Gamma_0^{-i};\beta^S],i} - \bar{\mu}| \\
&\le \frac{C}{\sqrt{n}}.
\end{aligned}
$$

By the same arguments as in the proof of Theorem 3.1 then, specifically for (4.6) and (4.8), we have that:

$$
J_0^{n,i}(\Gamma_0) - J_0^{n,i}([\Gamma_0^{-i};\beta^S]) \le \mathcal{O}(n^{-1/2}).
$$

So far in the proof, we had to assume β^S was stationary in order to invoke the stationary distribution. Using Lemma 5.1 we can now upgrade this result to hold for all $\beta \in \mathcal{A}^{n,i}$ non-stationary. Recalling γ^W from Lemma 5.1, and using the prior display with $\beta^S = \gamma^W$:

$$
J_0^{n,i}(\Gamma_0) - \inf_{\beta\in\mathcal{A}^{n,i}} J_0^{n,i}([\Gamma_0^{-i};\beta]) = J_0^{n,i}(\Gamma_0) - J_0^{n,i}([\Gamma_0^{-i};\gamma^W]) \le \mathcal{O}(n^{-1/2}).
$$

$\square$

5.2. Proof of Lemma 5.1.

The proof proceeds in two main steps. For the first, we establish a unique solution to an auxiliary discounted system. Then, by a vanishing discount argument, we obtain a solution to the auxiliary ergodic system (5.1). Finally, by a verification argument—that is, plugging in a carefully chosen process to the mock potential function from (5.1) and using Itô's lemma—we get (5.2).

5.2.1. *Existence for a discounted system.* We start by obtaining a solution to an auxiliary discounted system in order to ultimately construct a solution to the ergodic system (5.1). We proceed by defining a map S for which any fixed point will solve the discounted system. We prove that for any discount r, S is a contraction and therefore there exists a unique solution W_r^i to the auxiliary discounted system.

Let $r > 0$ be the discount factor. We will first prove that there exists $W_r^i : [d] \times \mathcal{P}^{n-1}([d]) \to [0, C_{f+F}/r]$ solving:

$$rW_r^i(x^i, \mu_{\boldsymbol{x}}^{n,i}) = H(x^i, \Delta_{x^i} W_r^i(\cdot, \mu_{\boldsymbol{x}}^{n,i})) + F(x^i, \mu_{\boldsymbol{x}}^{n,i}) + \sum_{j \neq i} \gamma_0^j(\boldsymbol{x}) \cdot \Delta_{x^j} U_0(\cdot, \mu_{\boldsymbol{x}}^{n,j}). \tag{5.7}$$

Define:

$$\mathcal{V} := \{w : [d]^n \to [0, C_{f+F}/r] \mid \exists \phi^w : [d] \times \mathcal{P}^{n-1}([d]) \to [0, C_{f+F}/r] \text{ s.t. } w(\boldsymbol{x}) = \phi^w(x^i, \mu_{\boldsymbol{x}}^{n,i})\}.$$

Define a map $S : \mathcal{V} \to \mathcal{V}$ by:

$$S(w)(\boldsymbol{x}) := \inf_{a \in \mathbb{A}_{x^i}^d} R(w, a)(\boldsymbol{x})$$

where:

$$R(w, a)(\boldsymbol{x}) := \frac{1}{r + \sum_{y \neq x^i} a_y} \Big[\sum_{y, y \neq x^i} a_y \phi^w(y, \mu_{\boldsymbol{x}}^{n,i}) + F(x^i, \mu_{\boldsymbol{x}}^{n,i}) + f(x^i, a)$$
$$+ \sum_{j \neq i} \gamma_0^j(\boldsymbol{x}) \cdot \Delta_{x^j} U_0(\cdot, \mu_{\boldsymbol{x}}^{n,j}) \Big].$$

Now,

$$|(S(w) - S(\tilde{w}))(\boldsymbol{x})| \leq \sup_{a \in \mathbb{A}_{x^i}^d} |(R(w, a) - R(\tilde{w}, a))(\boldsymbol{x})|$$
$$\leq \sup_{a \in \mathbb{A}_{x^i}^d} \left\{ \frac{1}{r + \sum_{y \neq x^i} a_y} \Big[\sum_{y, y \neq x^i} a_y |\phi^w(y, \mu_{\boldsymbol{x}}^{n,i}) - \phi^{\tilde{w}}(y, \mu_{\boldsymbol{x}}^{n,i})| \Big] \right\} \tag{5.8}$$
$$\leq \sup_{a \in \mathbb{A}_{x^i}^d} \left\{ \frac{\sum_{y \neq x^i} a_y}{r + \sum_{y \neq x^i} a_y} [\|w - \tilde{w}\|_{\sup}] \right\}$$
$$\leq \tilde{r} \|w - \tilde{w}\|_{\sup},$$

where

$$\|w\|_{\sup} := \max_{\boldsymbol{x} \in [d]^n} |w(\boldsymbol{x})|$$

and

$$\tilde{r} := \sup_{x^i \in [d]} \sup_{a \in \mathbb{A}_{x^i}^d} \left\{ \frac{\sum_{y \neq x^i} a_y}{r + \sum_{y \neq x^i} a_y} \right\} \in (0, 1).$$

Taking the supremum of the left-hand side of (5.8) over all $\boldsymbol{x}$, we obtain that S is a contraction. By Banach fixed point theorem there exists a unique fixed point of S, which we call w_r^i, with $w_r^i(\boldsymbol{x}) = W_r^i(x^i, \mu_{\boldsymbol{x}}^{n,i})$ for some $W_r^i : [d] \times \mathcal{P}^{n-1}([d]) \to [0, C_{f+F}/r]$ since $w_r^i \in \mathcal{V}$. Unfurling the definition of S and using that w_r^i is its fixed point, we find that W_r^i satisfies (5.7).

5.2.2. *Verification and vanishing discount.* Now we can make use of the discounted system (5.7) from the prior step in order to approach (5.1). We first confirm that $w_r^i(\boldsymbol{x})$ equals a discounted cost under the strategy profile Υ_r. This gives an alternate form for w_r^i that we will use to show $w_r^i(\cdot) - w_r^i(\hat{\boldsymbol{x}})$ is bounded in maximum norm, for some $\hat{\boldsymbol{x}} \in [d]^n$, by a stopping time argument. With a uniform bound, we pass to a subsequence of discounts decreasing to zero and show that the discounted system (5.7) converges to (5.1). We then use another verification argument to get (5.2).

For brevity, define $\Upsilon_r := [\Gamma_0^{-i}; \gamma^{W_r}]$. To get ϱ^i and W_0^i, we proceed by a vanishing discount argument. To do this, we first verify that W_r^i is the discounted cost:

$$J_r^{n,i}(\boldsymbol{x}, \Upsilon_r) := \mathbb{E}_{\boldsymbol{x}} \int_0^\infty e^{-rt} \left(f(X_t^{\Upsilon_r,i}, \gamma^{W_r}(\boldsymbol{X}_t^{\Upsilon_r})) + F(X_t^{\Upsilon_r,i}, \mu_t^{\Upsilon_r,i}) \right) dt.$$

By applying Itô's lemma to $e^{-rt} w_r^i(\boldsymbol{X}_t^{\Upsilon_r})$ and taking the expectations, we get that

$$\mathbb{E}_{\boldsymbol{x}} \left[e^{-rt} w_r^i(\boldsymbol{X}_t^{\Upsilon_r}) - w_r^i(\boldsymbol{X}_0^{\Upsilon_r}) \right] \tag{5.9}$$

$$= \mathbb{E}_{\boldsymbol{x}}\Big[\int_0^t e^{-rs}\Big(-rw_r^i(\boldsymbol{X}_s^{\Upsilon_r}) + \sum_{j\neq i}\gamma_0^j(\boldsymbol{X}_s^{\Upsilon_r})\cdot\Delta_{X_s^{\Upsilon_r,j}}U_0(\cdot,\mu_s^{\Upsilon_r,j})$$
$$+\gamma^{W_r}(\boldsymbol{X}_s^{\Upsilon_r})\cdot\Delta_{X_s^{\Upsilon_r,i}}W_r^i(\cdot,\mu_s^{\Upsilon_r,i})\Big)ds\Big].$$

We use (5.7) in (5.9) and rearranging the terms, it follows that

$$w_r^i(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}\Big[e^{-rt}w_r^i(\boldsymbol{X}_t^{\Upsilon_r})\Big] + \mathbb{E}_{\boldsymbol{x}}\Big[\int_0^t e^{-rs}\left(f(X_s^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_s^{\Upsilon_r})) + F(X_s^{\Upsilon_r,i},\mu_s^{\Upsilon_r,i})\right)ds\Big].$$

Note that w_r^i is bounded. Hence, passing to the limit implies:

$$w_r^i(\boldsymbol{x}) = \lim_{t\to\infty}\Big\{\mathbb{E}_{\boldsymbol{x}}\Big[e^{-rt}w_r^i(\boldsymbol{X}_t^{\Upsilon_r})\Big] + \mathbb{E}_{\boldsymbol{x}}\Big[\int_0^t e^{-rs}\left(f(X_s^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_s^{\Upsilon_r})) + F(X_s^{\Upsilon_r,i},\mu_s^{\Upsilon_r,i})\right)ds\Big]\Big\} \quad (5.10)$$
$$= \mathbb{E}_{\boldsymbol{x}}\Big[\int_0^\infty e^{-rs}\left(f(X_s^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_s^{\Upsilon_r})) + F(X_s^{\Upsilon_r,i},\mu_s^{\Upsilon_r,i})\right)ds\Big]$$
$$= J_r^{n,i}(\boldsymbol{x},\Upsilon_r).$$

Recall that players choose their rates from the set $[\mathfrak{a}_l,\mathfrak{a}_u]$ where $0 < \mathfrak{a}_l < \mathfrak{a}_u < \infty$. Since there are finitely many configurations, every configuration is positive recurrent so there exists $N \in (0,\infty)$, independent of r, such that for any arbitrary $\boldsymbol{x},\hat{\boldsymbol{x}} \in [d]^n$, one has $\mathbb{E}_{\boldsymbol{x}}[\hat\tau] \leq N$ where $\hat\tau := \inf\{t \geq 0 \mid \boldsymbol{X}_t = \hat{\boldsymbol{x}}\}$.

Fix an arbitrary $\hat{\boldsymbol{x}} \in [d]^n$. By definition and some arithmetic:

$$J_r^{n,i}(\boldsymbol{x},\Upsilon_r) = \mathbb{E}_{\boldsymbol{x}}\Big[\int_0^{\hat\tau} e^{-rt}\left(f(X_t^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_t^{\Upsilon_r})) + F(X_t^{\Upsilon_r,i},\mu_t^{\Upsilon_r,i})\right)dt + J_r^{n,i}(\boldsymbol{X}_{\hat\tau},\Upsilon_r)$$
$$- (1-e^{-r\hat\tau})J_r^{n,i}(\boldsymbol{X}_{\hat\tau},\Upsilon_r)\Big].$$

By definition of $\hat\tau$, we know that $\mathbb{E}[J_r^{n,i}(\boldsymbol{X}_{\hat\tau},\Upsilon_r)] = J_r^{n,i}(\hat{\boldsymbol{x}},\Upsilon_r)$ and so by the previous computation and the introduction of absolute value, we see that:

$$|J_r^{n,i}(\boldsymbol{x},\Upsilon_r) - J_r^{n,i}(\hat{\boldsymbol{x}},\Upsilon_r)| \leq \mathbb{E}_{\boldsymbol{x}}\left[\int_0^{\hat\tau} e^{-rt}\left|f(X_t^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_t^{\Upsilon_r})) + F(X_t^{\Upsilon_r,i},\mu_t^{\Upsilon_r,i})\right|dt\right]$$
$$+ \mathbb{E}_{\boldsymbol{x}}[(1-e^{-r\hat\tau})|J_r^{n,i}(\boldsymbol{X}_{\hat\tau},\Upsilon_r)|].$$

To deal with the first term after the inequality we note that since f and F are bounded:

$$\mathbb{E}_{\boldsymbol{x}}\left[\int_0^{\hat\tau} e^{-rt}\left|f(X_t^{\Upsilon_r,i},\gamma^{W_r}(\boldsymbol{X}_t^{\Upsilon_r})) + F(X_t^{\Upsilon_r,i},\mu_t^{\Upsilon_r,i})\right|dt\right] \leq \mathbb{E}_{\boldsymbol{x}}[\hat\tau]C_{f+F} \leq NC_{f+F}.$$

For the second term, we use a max bound followed by the inequality $(1-e^{-rz})/r \leq z$ for all $r \geq 0$ to see that:

$$\mathbb{E}_{\boldsymbol{x}}[(1-e^{-r\hat\tau})|J_r^{n,i}(\boldsymbol{X}_{\hat\tau},\Upsilon_r)|] \leq \mathbb{E}_{\boldsymbol{x}}[r^{-1}(1-e^{-r\hat\tau})]\max_{\boldsymbol{y}\in[d]^n} r|J_r^{n,i}(\boldsymbol{y},\Upsilon_r)|$$
$$\leq \mathbb{E}_{\boldsymbol{x}}[\hat\tau]\max_{\boldsymbol{y}\in[d]^n} r|J_r^{n,i}(\boldsymbol{y},\Upsilon_r)|$$
$$\leq NC_{f+F},$$

where recall that the solution w_r^i to (5.7) was bounded above by C_{f+F}/r for all configurations $\boldsymbol{y} \in [d]^n$ and so $J_r^{n,i}(\boldsymbol{y},\Upsilon_r) = w_r^i(\boldsymbol{y}) \leq C_{f+F}r^{-1}$ implies that $rJ_r^{n,i}(\boldsymbol{y},\Upsilon_r) \leq C_{f+F}$. This gives the last inequality above. Combining these results:

$$|w_r^i(\boldsymbol{x}) - w_r^i(\hat{\boldsymbol{x}})| \leq 2NC_{f+F}.$$

Since the bound is uniform with respect to r, we see that for any sequence $(r_m)_{m\geq0} \to 0$:

$$\max_{\boldsymbol{x}\in[d]^n}|w_{r_m}^i(\boldsymbol{x}) - w_{r_m}^i(\hat{\boldsymbol{x}})| \leq 2NC_{f+F}.$$

We note that $\max_{\boldsymbol{x}\in[d]^n}|\cdot|$ as above is a norm on the space of functions $[d]^n \to [0, C_{f+F}/r]$ and that the difference $w^i_{r_m}(\boldsymbol{x}) - w^i_{r_m}(\hat{\boldsymbol{x}})$ is bounded uniformly in this norm. Therefore there exists a subsequence of $(r_m)_m$, which by abuse of notation we also refer to as $(r_m)_m$, such that $w^i_{r_m}(\cdot) - w^i_{r_m}(\hat{\boldsymbol{x}})$ converges as $r_m \to 0$. Denote this limit by $w^i_0(\cdot)$ and since $w^i_{r_m} \in \mathcal{V}$ there exists W^i_0 such that $w^i_0(\boldsymbol{x}) = W^i_0(x^i, \mu^{n,i}_{\boldsymbol{x}})$. So, for all $\boldsymbol{x} \in [d]^n$, $W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}}) - W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}}) \to W^i_0(x^i, \mu^{n,i}_{\boldsymbol{x}})$ as $m \to \infty$ too.

We can now apply this vanishing discount to (5.7) to obtain a solution for (5.1). We note that:

$$
\begin{aligned}
\lim_{m\to\infty} \Delta_{x^i} W^i_{r_m}(\cdot, \mu^{n,i}_{\boldsymbol{x}}) &= \lim_{m\to\infty} \left(W^i_{r_m}(y, \mu^{n,i}_{\boldsymbol{x}}) - W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}})\right)_{y\in[d]} \\
&= \lim_{m\to\infty} \left(W^i_{r_m}(y, \mu^{n,i}_{\boldsymbol{x}}) - W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}}) + W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}}) - W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}})\right)_{y\in[d]} \\
&= \left(W^i_0(y, \mu^{n,i}_{\boldsymbol{x}}) - W^i_0(x^i, \mu^{n,i}_{\boldsymbol{x}})\right)_{y\in[d]} \\
&= \Delta_{x^i} W^i_0(\cdot, \mu^{n,i}_{\boldsymbol{x}}).
\end{aligned}
$$

Again by passing to a subsequence, we can denote ϱ^i to be the limit of the real-valued and bounded sequence $\{r_m W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}})\}_{m\in\mathbb{N}}$. Once more, recall that $|W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}}) - W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}})|$ is bounded for any configuration $\boldsymbol{x} \in [d]^n$. Now we can apply the continuity of H to obtain that:

$$
\begin{aligned}
\varrho^i &= \lim_{m\to\infty} [r_m W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}}) + r_m(W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}}) - W^i_{r_m}(\hat{x}^i, \mu^{n,i}_{\hat{\boldsymbol{x}}}))] \\
&= \lim_{m\to\infty} r_m W^i_{r_m}(x^i, \mu^{n,i}_{\boldsymbol{x}}) \\
&= \lim_{m\to\infty} \left[H(x^i, \Delta_{x^i} W^i_{r_m}(\cdot, \mu^{n,i}_{\boldsymbol{x}})) + F(x^i, \mu^{n,i}_{\boldsymbol{x}}) + \sum_{j\neq i} \gamma^j_0(\boldsymbol{x}) \cdot \Delta_{x^j} U_0(\cdot, \mu^{n,j}_{\boldsymbol{x}})\right] \\
&= H(x^i, \Delta_{x^i} W^i_0(\cdot, \mu^{n,i}_{\boldsymbol{x}})) + F(x^i, \mu^{n,i}_{\boldsymbol{x}}) + \sum_{j\neq i} \gamma^j_0(\boldsymbol{x}) \cdot \Delta_{x^j} U_0(\cdot, \mu^{n,j}_{\boldsymbol{x}}).
\end{aligned}
$$

Hence, we established (5.1).

As the final step of this proof, we now turn to prove (5.2). Let $\Upsilon := [\Gamma^{-i}_0; \beta]$, $\beta \in \mathcal{A}^{n,i}$. By another verification-style argument like for (5.10), we can show that:

$$
\mathbb{E}_{\boldsymbol{x}}(w^i_0(\boldsymbol{X}^\Upsilon_0)) \leq \mathbb{E}_{\boldsymbol{x}}\left(w^i_0(\boldsymbol{X}^\Upsilon_T)\right) + \mathbb{E}_{\boldsymbol{x}}\left(\int_0^T \left[f(X^{\Upsilon,i}_s, \beta(s, \boldsymbol{X}^{\Upsilon,i}_s)) + F(X^{\Upsilon,i}_s, \mu^{\Upsilon,i}_s)\right] ds\right) - T\varrho^i,
$$

with equality above when $\beta = \gamma^W$ since γ^W is the minimizer in (5.1). Dividing by T and passing the $\limsup$ as $T \to \infty$, we get:

$$
\varrho^i \leq \mathbb{E}_{\boldsymbol{x}}\left(\int_0^T \left[f(X^{\Upsilon,i}_s, \beta(s, \boldsymbol{X}^{\Upsilon,i}_s)) + F(X^{\Upsilon,i}_s, \mu^{\Upsilon,i}_s)\right] ds\right) = J^{n,i}_0(\Upsilon).
$$

Since we have equality when $\beta = \gamma^W$ and inequality otherwise, we conclude (5.2). $\qquad\square$

5.3. **Proof of Lemma 5.2.** The proof is modeled after duality estimates for the MFG, see [12, Lemma 3.6], for example. In that case, the functions in play are solutions to ODEs and there is no randomness. Here, however, we must consider random processes and consequently, the proof is much more involved. Specifically, we will have to deal with three main issues: (1) martingale terms arising from the jump dynamics of the empirical measure, and (2) a deviating player with a potentially different strategy, causing error terms to aggregate, and (3) that because the process we consider is a jump process on $\mathcal{P}^n([d])$, it may be modified slightly each time it appears, thus accumulating more error. The ultimate goal is to drag the accumulating errors through the proof and into the estimate at the end.

Fix arbitrary $\beta \in \mathcal{A}^{n,i}$ and $\pi \in \mathcal{P}([d]^n)$. For brevity, denote the strategy profile $\Upsilon := [\Gamma^{-i}_0; \beta]$. Define $dL^\Upsilon_{yz}(t)$ as the martingale increment:

$$
dL^\Upsilon_{yz}(t) = dK^\Upsilon_{yz}(t) - \left[(n-1)\mu^{\Upsilon,i}_{t,y}\gamma^*_z(y, \Delta_y U_0(\cdot, \mu^{\Upsilon,-y}_t)) + \mathbb{1}_{\{X^{\Upsilon,i}_t=y\}}\beta_{yz}(t, \boldsymbol{X}^\Upsilon_t)\right] dt,
$$

where recall that K^β is defined generally in (3.1). *We now go through several displays to compute* $dU_0(x, \mu^\Upsilon_t)$. Recall that $\mu^{\Upsilon,-x}$ is the empirical distribution of the players, minus one player in position x.

By Itô's chain rule,

$$dU_0(x, \mu_t^\Upsilon) = \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dK_{yz}^\Upsilon(t)$$

[adding and subtracting a rate term to get a martingale]

$$= \sum_{y,z \in [d]} \left\{ \left[U_0(x, \mu_t^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_t^\Upsilon) \right] \right.$$
$$\left. \times \left[(n-1)\mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) + \mathbb{1}_{\{X_t^{\Upsilon,i} = y\}} \beta_{yz}(t, \boldsymbol{X}_t^\Upsilon) \right] dt \right\}$$
$$+ \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dL_{yz}^\Upsilon(t)$$

[adding and subtracting a $D^\eta U_0$ term]

$$= \sum_{y,z \in [d]} D_{yz}^\eta U_0(x, \mu_t^\Upsilon) \left[\mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) + \tfrac{1}{n} \mathbb{1}_{\{X_s^{\Upsilon,i} = y\}} \beta_{yz}(t, \boldsymbol{X}_t^\Upsilon) \right] dt$$
$$+ \sum_{y,z \in [d]} \left[(n-1)((U_0(x, \mu_t^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_t^\Upsilon)) - D_{yz}^\eta U_0(x, \mu_t^\Upsilon) \right]$$
$$\times \left[\mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) + \tfrac{1}{n} \mathbb{1}_{\{X_s^{\Upsilon,i} = y\}} \beta_{yz}(t, \boldsymbol{X}_t^\Upsilon) \right] dt$$
$$+ \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dL_{yz}^\Upsilon(t)$$

[using Lipschitz continuity of $D^\eta U_0$ and the boundedness of the rates,

we can aggregate terms of order $\mathcal{O}(n^{-1})$]

$$= \sum_{y,z \in [d]} D_{yz}^\eta U_0(x, \mu_t^\Upsilon) \mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) dt$$
$$+ \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dL_{yz}^\Upsilon(t)$$
$$+ \mathcal{O}(n^{-1})$$

[From the Lipschitz continuity of γ^* and U_0 as well as the facts that

$$|\mu_{t,y}^{\Upsilon,i} - \mu_{t,y}^\Upsilon| \le C/n, \qquad |\mu_t^{\Upsilon,-y} - \mu_t^\Upsilon| \le C/n,$$

we can aggregate more terms into $\mathcal{O}(n^{-1})$]

$$= \sum_{y,z \in [d]} D_{yz}^\eta U_0(x, \mu_t^\Upsilon) \mu_{t,y}^\Upsilon \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^\Upsilon)) dt$$
$$+ \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dL_{yz}^\Upsilon(t)$$
$$+ \mathcal{O}(n^{-1}).$$

And finally by (1.5):

$$dU_0(x, \mu_t^\Upsilon) = (\varrho - H(x, \Delta_x U_0(\cdot, \mu_t^\Upsilon)) - F(x, \mu_t^\Upsilon)) dt$$
$$+ \sum_{y,z \in [d]} \left[U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon) \right] dL_{yz}^\Upsilon(t) \tag{5.11}$$
$$+ \mathcal{O}(n^{-1}),$$

Set the processes:

$$v_x(t) := U_0(x, \mu_t^\Upsilon) - \check{u}_x(t) \qquad \text{and} \qquad m_x(t) := \mu_{t,x}^\Upsilon - \check{\mu}_x(t), \qquad (t, x) \in \mathbb{R}_+ \times [d].$$

Now, an application of Itô's product rule implies that:

$$m(T) \cdot v(T) - m(0) \cdot v(0) = \int_0^T \sum_{x \in [d]} d[m_x(t) v_x(t)]$$

$$= \int_0^T \sum_{x \in [d]} (m_x(t) dv_x(t) + v_x(t) dm_x(t) + dm_x(t) dv_x(t)) .$$

We will treat the three terms in the integrand of the previous display separately. Taking expectations on both sides, using (1.3), (3.1), (3.2), and that L_{yz}^Υ is a martingale, we get (with more detailed explanations to follow),

$$\mathbb{E}_\pi[m(T) \cdot v(T) - m(0) \cdot v(0)] = \sum_{i=1}^3 \Psi_i(T) + \mathcal{O}(n^{-1}), \tag{5.12}$$

where the $\mathcal{O}(n^{-1})$ term comes from (5.11), where:

$$\Psi_1(T) := \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} m_x(t) \big(\varrho - H(x, \Delta_x U_0(\cdot, \mu_t^\Upsilon)) - F(x, \mu_t^\Upsilon) - \check{\varrho} + H(x, \Delta_x \check{u}(t)) + F(x, \check{\mu}(t))\big) dt,$$

$$\Psi_2(T) := \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} v_x(t) \Big[\sum_{y,z \in [d]} \frac{(e_z - e_y)_x}{n} \big[(n-1)\mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) + \mathbb{1}_{\{X_t^{\Upsilon,i} = y\}} \beta_{yz}(t, \boldsymbol{X}_t^\Upsilon)\big]$$

$$- \sum_{y \in [d]} \check{\mu}_y(t) \check{\gamma}_{yx}(t) \Big] dt,$$

$$\Psi_3(T) := \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} \sum_{y,z \in [d]} (U_0(x, \mu_{t-}^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_{t-}^\Upsilon)) \frac{(e_z - e_y)_x}{n} dK_{yz}^\Upsilon(t)$$

$$= \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} \sum_{y,z \in [d]} \Big\{ (U_0(x, \mu_t^\Upsilon + \tfrac{e_z - e_y}{n}) - U_0(x, \mu_t^\Upsilon)) \frac{(e_z - e_y)_x}{n}$$

$$\times \big[(n-1)\mu_{t,y}^{\Upsilon,i} \gamma_z^*(y, \Delta_y U_0(\cdot, \mu_t^{\Upsilon,-y})) + \mathbb{1}_{\{X_t^{\Upsilon,i} = y\}} \beta_{yz}(t, \boldsymbol{X}_t^\Upsilon)\big] \Big\} dt.$$

We note that we also used (1.3) to substitute in for $\frac{d}{dt}\check{u}(t)$ and $\frac{d}{dt}\check{\mu}(t)$ in Ψ_1 and Ψ_2, respectively. Moreover, to derive the first expression for Ψ_3 we used the facts that the quadratic variation of a compensated random Poisson measure is the non-compensated measure and that $dK_{xy}^\Upsilon(t) dK_{wz}^\Upsilon(t) = 1$ if and only if $(x, y) = (w, z)$ and otherwise $dK_{xy}^\Upsilon(t) dK_{wz}^\Upsilon(t) = 0$. We can use the Lipschitz continuity of U_0 and the boundedness of γ^* and β to get that:

$$|\Psi_3(T)| \leq CTn^{-1}. \tag{5.13}$$

We can use $\sum_x m_x(t) = 0$ in the integrand of $\Psi_1(T)$ and then the monotonicity of F to get:

$$\Psi_1(T) = \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} m_x(t) \big(H(x, \Delta_x \check{u}(t)) - H(x, \Delta_x U_0(\cdot, \mu_t^\Upsilon)) + F(x, \check{\mu}(t)) - F(x, \mu_t^\Upsilon)\big) dt$$

$$\leq \mathbb{E}_\pi \int_0^T \sum_{x \in [d]} m_x(t) \big(H(x, \Delta_x \check{u}(t)) - H(x, \Delta_x U_0(\cdot, \mu_t^\Upsilon))\big) dt. \tag{5.14}$$

For $\Psi_2(T)$, we will deal with the term on the first line for the moment, and return to the full expression after. We can use the definition of $(e_z - e_y)_x$ first to separate into two sums, the fact that $\sum_{y \in [d]} \gamma_y^*(x, \cdot) = 0$, recombine the sums into one, the Lipschitz continuity of γ^* and U_0 as well as the fact that $((n-1)\mu_{t,y}^{\Upsilon,i} +$

$\delta_{X_t^{\Upsilon,i}}(y))/n = \mu_{t,y}^{\Upsilon}$ to aggregate more terms of order $\mathcal{O}(n^{-1})$, and get that it equals:

$$\mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t) \Big[\sum_{y,y\neq x} \frac{1}{n}\big[(n-1)\mu_{t,y}^{\Upsilon,i}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon,-y})) + \mathbb{1}_{\{X_t^{\Upsilon,i}=y\}}\beta_{yx}(t,\boldsymbol{X}_t^{\Upsilon})\big]\Big]dt$$

$$+ \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\sum_{z,z\neq x} \frac{-1}{n}\big[(n-1)\mu_{t,x}^{\Upsilon,i}\gamma_z^*(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon,-x})) + \mathbb{1}_{\{X_t^{\Upsilon,i}=x\}}\beta_{xz}(t,\boldsymbol{X}_t^{\Upsilon})\big]\Big]dt$$

$$= \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\sum_{y,y\neq x} \frac{1}{n}\big[(n-1)\mu_{t,y}^{\Upsilon,i}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon,-y})) + \mathbb{1}_{\{X_t^{\Upsilon,i}=y\}}\beta_{yx}(t,\boldsymbol{X}_t^{\Upsilon})\big]\Big]dt$$

$$+ \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\frac{(n-1)\mu_{t,x}^{\Upsilon,i}}{n}\big[\gamma_x^*(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon,-x})) + \frac{1}{n-1}\mathbb{1}_{\{X_t^{\Upsilon,i}=x\}}\beta_{xx}(t,\boldsymbol{X}_t^{\Upsilon})\big]\Big]dt$$

$$= \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\sum_{y\in[d]} \frac{n-1}{n}\big[\mu_{t,y}^{\Upsilon,i}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon,-y})) + \mathcal{O}(n^{-1})\big]\Big]dt$$

$$= \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\sum_{y\in[d]} \big[\mu_{t,y}^{\Upsilon}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon})) + \mathcal{O}(n^{-1})\big]\Big]dt$$

Therefore, returning to the definition of Ψ_2, we have:

$$\Psi_2(T) = \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} v_x(t)\Big[\sum_{y\in[d]} \big[\mu_{t,y}^{\Upsilon}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon})) - \check{\mu}_y(t)\check{\gamma}_{yx}(t) + \mathcal{O}(n^{-1})\big]\Big]dt. \qquad (5.15)$$

Using (5.14), (5.15), (5.13) in (5.12):

$$\mathbb{E}_\pi[m(T)\cdot v(T) - m(0)\cdot v(0)] \leq \hat{\mathbb{E}}_\pi \int_0^T \Big\{ \sum_{x\in[d]}[H(x,\Delta_x\check{u}(t)) - H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon}))]m_x(t)$$

$$+ v_x(t)\Big[\sum_{y\in[d]}\mu_{t,y}^{\Upsilon}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon})) - \check{\mu}_y(t)\check{\gamma}_{yx}(t)\Big]\Big\}dt \qquad (5.16)$$

$$+ T\mathcal{O}(n^{-1}).$$

Next, utilizing the identity:

$$\sum_{x\in[d]}\check{\gamma}_{yx}(t) = \sum_{x\in[d]}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon})) = 0,$$

we get that:

$$\sum_{x\in[d]}\Big[\big\{H(x,\Delta_x\check{u}(t)) - H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon}))\big\}m_x(t)$$

$$+ v_x(t)\sum_{y\in[d]}\{\mu_{t,y}^{\Upsilon}\gamma_x^*(y,\Delta_y U_0(\cdot,\mu_t^{\Upsilon})) - \check{\mu}_y(t)\check{\gamma}_{yx}(t)\}\Big]$$

$$= \sum_{x\in[d]}\Big[\big\{H(x,\Delta_x\check{u}(t)) - H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon})) + \Delta_x v(t)\cdot\gamma^*(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon}))\big\}\mu_{t,x}^{\Upsilon}\Big] \qquad (5.17)$$

$$+ \sum_{x\in[d]}\Big[\big\{H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon})) - H(x,\Delta_x\check{u}(t)) - \Delta_x v(t)\cdot\gamma^*(x,\Delta_x\check{u}(t))\big\}\check{\mu}_x(t)\Big].$$

Using (2.2) and assumption (B_2), we obtain the inequalities:

$$H(x,\Delta_x\check{u}(t)) - H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon})) + \Delta_x v(t)\cdot\gamma^*(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon}))) \leq -C_{2,H}|\Delta_x v(t)|^2,$$

$$H(x,\Delta_x U_0(\cdot,\mu_t^{\Upsilon})) - H(x,\Delta_x\check{u}(t)) - \Delta_x v(t)\cdot\gamma^*(x,\Delta_x\check{u}(t))) \leq -C_{2,H}|\Delta_x v(t)|^2.$$

The last two bounds applied to the expansion in (5.17) imply:

$$
\begin{aligned}
\sum_{x\in[d]} \Big[& \big\{ H(x, \Delta_x \check{u}(t)) - H(x, \Delta_x U_0(\cdot, \mu_t^\Upsilon)) \big\} m_x(t) \\
& + v_x(t) \sum_{y\in[d]} \{ \mu_{t,y}^\Upsilon \gamma_x^*(y, \Delta_y U_0(\cdot, \mu_t^\Upsilon)) - \check{\mu}_y(t) \check{\gamma}_{yx}(t) \} \Big] \\
& \qquad\qquad\qquad\qquad \leq -C_{2,H} \sum_{x\in[d]} (\mu_{t,x}^\Upsilon + \check{\mu}_x(t)) |\Delta_x v(t)|^2.
\end{aligned}
\tag{5.18}
$$

Plugging in (5.18) into (5.16), moving the negated terms to the left-hand side,

$$
C_{2,H} \mathbb{E}_\pi \int_0^T \sum_{x\in[d]} (\mu_{t,x}^\Upsilon + \check{\mu}_x(t)) |\Delta_x v(t)|^2 dt \leq T\mathcal{O}(n^{-1}) + \mathbb{E}_\pi[m(0) \cdot v(0) - m(T) \cdot v(T)].
$$

Dividing the prior display by T, taking $\limsup_{T\to\infty}$, and using the boundedness of m and v finishes the proof. $\qquad\square$

6. Proofs of Proposition 3.1 and Theorem 3.3

We finish the main theoretical portion of the paper by proving the results pertaining to the large deviations of the empirical measures.

6.1. **Proof of Proposition 3.1.** Consider a globally asymptotically stable equilibrium for (2.3), call it $\mathring{\mu}$. Define $\mathring{u} := (U_0(x, \mathring{\mu}))_{x\in[d]}$ and note that since $\mathring{\mu}$ is globally asymptotically stable, we have by taking a limit of (1.3) as $t \to \infty$ that:

$$
\bar{\varrho} = H(x, \Delta_x \mathring{u}) + F(x, \mathring{\mu}).
$$

Hence, $(\bar{\varrho}, \mathring{u}, \mathring{\mu})$ solves (1.4). Yet by Proposition 2.1, (1.4) has a unique solution, where recall that $\bar{u}$ is unique up to an additive constant vector $k\vec{1}$, and therefore for all $x \in [d]$, $\Delta_x \mathring{u} = \Delta_x \bar{u}$ with $\mathring{\mu} = \bar{\mu}$ in the first place.

By definition, $\Delta_x u_0(t, \cdot) = \Delta_x U_0(\cdot, \mu_0(t))$. And since U_0 is Lipschitz by Proposition 2.2, there exists $C > 0$, independent of t, such that:

$$
|\Delta_x U_0(\cdot, \mu_0(t))| \leq C(|\Delta_x \bar{u}| + |\mu_0(t) - \bar{\mu}|),
$$

which means $|\Delta_x u_0(t, \cdot)|$ is bounded. Therefore, we may re-use the proof from [16, Theorem 3] to obtain that $\mu_0(t) \to \bar{\mu}$ as $t \to \infty$.[1] Recalling that the initial condition η was arbitrary, we have that $\bar{\mu}$ is a globally asymptotically stable equilibrium. $\qquad\square$

6.2. **Proof of Theorem 3.3.** The proof is an application of theorems from [4], so we verify the assumptions of the theorems. Namely, the possible transitions on $[d]$ form a complete, finite graph and a fortiori such a graph is irreducible. In either case for κ^n, the rate mapping is Lipschitz because γ^* is Lipschitz by assumption, and U_0 is Lipschitz, by Proposition 2.2. Moreover, the rates are assumed to be bounded below by $\mathfrak{a}_l > 0$. Proposition 3.1 verifies the final requirement for each of the following theorems. So, [4, Theorem 2.1] proves the large deviation result for $\mathscr{L}_\infty^n$; [4, Theorem 3.2] for $\mathscr{L}_{[0,T]}^n$; and [4, Theorem 3.3] for $\mathscr{L}_T^n$. As a small comment, the set of paths over which the infimum in the good rate function $s(\cdot)$ is taken is nonempty due to [4, Lemma 4.2]. $\qquad\square$

---

[1]It is worth mentioning that this theorem relies on the additional contraction assumptions of that work. However, their proof only uses the contractive assumptions to bound their potential functions $u_0(t)$ in a sub-additive (non-norm) function $\|\cdot\|_\sharp$. Here, $\|u_0(t)\|_\sharp$ measures the maximum difference between potentials of any two states from $[d]$ at time $t \in \mathbb{R}_+$ and so is equivalent to $|\Delta u_0(t)|$ in our case. Namely, for any measurable function $b : \mathbb{R}_+ \to \mathbb{R}^d$, the quantity $\|b(t)\|_\sharp$ is bounded for all $t \in \mathbb{R}_+$ if and only if $|\Delta b(t)|$ is bounded for all $t \in \mathbb{R}_+$.

7. Numerics Comparing $(C/\sqrt{n})$-Nash Equilibria

By Corollary 3.1 and Theorem 3.2, we saw that there are, under the full assumptions, at least two distinct types of $(C/\sqrt{n})$-Nash equilibria: one driven by the master equation's solution, Γ_0, and one driven by the solution to the stationary ergodic MFG, $\bar{\Gamma}$. One might ask whether, qualitatively speaking, players adhering to the approximate Nash equilibria profiles $\bar{\Gamma}$ or Γ_0 experience different costs, or player behaviors. In Section 7.1, we solve a class of quadratic examples for the stationary ergodic MFG system (1.4). This allows us to explicitly state the strategies for the stationary Markovian strategy profile $\bar{\Gamma}$ in our numerics. To approximate the strategies making up Γ_0 however, no such explicit formula is available due to the dependence on U_0. To tackle this, the following Section 7.2 adapts the deep Galerkin method to approximate the solution to (1.5). We then begin our comparison of the strategies forming $\bar{\Gamma}$ and Γ_0, compare their respective stationary distributions for each empirical measure, and contrast the realized costs under each regime.

7.1. **Explicit solutions to the stationary ergodic MFG.** The example of quadratic costs is ubiquitous in MFG literature and will serve as our benchmark problem for the numerics that follow this section. Here, we will explicitly solve the example problem in the finite-state ergodic case. This particular example was studied for the finite-horizon, finite-state case in [9, Example 1], [2, Example 3.1], and [11, Example 7.1].

Fix a parameter $\delta \geq 0$ that will serve as a bias to avoid state $1 \in [d]$ and fix a scaling for the rate costs $b > 0$. We will consider an example with the parameters:

$$\mathbb{A} = [1,3], \qquad d = 2, \qquad f(x,a) = \delta \mathbb{1}_{\{x=1\}} + b\sum_{y \neq x}(a_{xy} - 2)^2.$$

Pausing to consider the form of f, we note that we have introduced two parameters δ and b whose interplay will determine the equilibrium state of the game. The larger δ is, the more players will be incentivized to move quickly away from state 1 to avoid the additional cost of remaining in that state. Meanwhile, increasing b means that players who want to move faster or slower than the "free" rate of 2, to avoid state 1 for instance, will pay a higher cost. With this formulation, we can explicitly compute the Hamiltonian H and its minimizer γ^* as:

$$H(x,p) = \sum_{y \neq x}\left(2p_y - \frac{p_y^2}{4b}\right), \qquad \gamma_y^*(x,p) = 2 - \frac{p_y}{2b}.$$

And moreover,

$$H(x,\Delta_x\bar{u}) = \sum_{y \neq x}\left(2(\bar{u}_y - \bar{u}_x) - \frac{(\bar{u}_y - \bar{u}_x)^2}{4b}\right), \qquad \gamma_y^*(x,\Delta_x\bar{u}) = 2 + \frac{\bar{u}_x - \bar{u}_y}{2b}.$$

To accommodate the full assumptions, we will consider a typical Lasry–Lions monotone mean field cost:

$$F(x,\eta) = \eta_x.$$

With these assumptions, (1.4) reads:

$$\begin{aligned}
\bar{\varrho} &= 2(\bar{u}_2 - \bar{u}_1) - \frac{1}{4b}(\bar{u}_2 - \bar{u}_1)^2 + \bar{\mu}_1 + \delta, \\
\bar{\varrho} &= 2(\bar{u}_1 - \bar{u}_2) - \frac{1}{4b}(\bar{u}_2 - \bar{u}_1)^2 + \bar{\mu}_2, \\
\bar{\mu}_1(-\phi_1) + \bar{\mu}_2\phi_2 &= 0, \\
\bar{\mu}_1\phi_1 + \bar{\mu}_2(-\phi_2) &= 0,
\end{aligned} \tag{7.1}$$

where:

$$\phi_1 := 2 + \frac{\bar{u}_1 - \bar{u}_2}{2b}, \quad \phi_2 := 2 + \frac{\bar{u}_2 - \bar{u}_1}{2b}. \tag{7.2}$$

We solve this system of equations explicitly. Using the fact that $\bar{\mu}_1 = 1 - \bar{\mu}_2$ in the measure equations of (7.1), we can show that:

$$-\phi_1 + \bar{\mu}_2(\phi_1 + \phi_2) = 0 \implies \bar{\mu}_2 = \phi_1/4,$$

where the implication follows from (7.2). For notational simplicity, define:

$$A := \bar{u}_2 - \bar{u}_1.$$

We can write the value equations from (7.1) as:

$$\varrho = 2A - \frac{1}{4b}A^2 + \bar{\mu}_1 + \delta$$

$$\varrho = -2A - \frac{1}{4b}A^2 + \bar{\mu}_2.$$

Subtracting the equations in the prior display,

$$0 = 4A + \bar{\mu}_1 - \bar{\mu}_2 + \delta.$$

Since $\bar{\mu}_2 = \phi_1/4$ and $\bar{\mu}_1 = 1 - \bar{\mu}_2$,

$$4A + \frac{A}{4b} + \delta = 0,$$

and so rearranging gives:

$$A = \delta\frac{4b}{16b + 1}.$$

And since we can write $\bar{\mu}_1$ and $\bar{\mu}_2$ in terms of A, we plug in the prior display to get:

$$\bar{\mu}_1 = \frac{1}{2}\left(1 - \frac{\delta}{16b + 1}\right), \qquad \bar{\mu}_2 = \frac{1}{2}\left(1 + \frac{\delta}{16b + 1}\right).$$

With this, we can return to the second value equation of (7.1) to get an expression for ϱ in terms of δ and b,

$$\varrho = \frac{1}{2} + \delta\frac{128b^2 + 24b - 4b\delta + 1}{(16b + 1)^2}.$$

Suppose $b = 4$, and $\delta = 0$ for instance. Then the solution reads:

$$\bar{\varrho} = 1/2, \qquad \bar{u}_2 - \bar{u}_1 = 0, \qquad \bar{\mu} = (1/2, 1/2).$$

When we take $\delta > 0$ we will no longer have symmetry in the stationary measure and in the potential function. For example, take $\delta = 1$ and $b = 4$ to get:

$$\bar{\varrho} = 8483/8450 \approx 1.003, \qquad \bar{u}_2 - \bar{u}_1 = 16/65, \qquad \bar{\mu} = (32/65, 33/65).$$

For the remainder of the paper, we will study the case when $b = 4$ and $\delta = 0$.

7.2. Approximate solutions to the ergodic master equation.

In this section, we use the DGM to solve the ergodic master equation and in order to numerically compare the strategies. The main difference we find is that the master equation-derived strategy is sensitive to the players' empirical distribution in the simplex and that this effect is most pronounced at the boundary. In a sense, this means that the master equation-derived strategy has a better short-term response than the stationary strategy. For the master equation's corresponding empirical n-player measure then, we would expect a faster transition away from the boundary of the simplex compared to the measure of players using the stationary strategy.

We solve the master equation corresponding to this data for U_0 using the DGM. While the solution of the master equation is a pair (ϱ, U_0), we note that since $\varrho = \bar{\varrho}$ by Proposition 2.2, we have already found that $\varrho = 1/2$ in Section 7.1. In Algorithm 1, we present an algorithm for approximating the potential function U_0 of the ergodic master equation, given the game's value ϱ.

Algorithm 1 is, in a sense, an extension of the DGM method for the stationary ergodic MFG, presented in [8, Section 4] for the MFG on the torus. Setting aside the different state-spaces, we note that [8] approximates the stationary ergodic MFG system while Algorithm 1 approximates the ergodic master equation. Moreover, our Algorithm 1 proceeds with ϱ as a known value which we solved for a priori. We showed in Section 7.1 an explicit expression for ϱ in the case of quadratic (and potentially asymmetric)

costs. In general, to compute ϱ, one would first solve (1.4) by using [8] and then using the value for ϱ determined, solve for U_0 using Algorithm 1.

As in the DGM, Algorithm 1 uses a family of parameterized functions $\mathcal{U}(\cdot, \cdot; \theta) : [d] \times \mathcal{P}([d]) \to \mathbb{R}$ to approximate U_0, where θ is a vector of the parameters in the neural network $\mathcal{U}$. For any run of the algorithm, we fix the architecture of $\mathcal{U}$ beforehand; here we use a fully-connected, feed-forward architecture like in [11], while [33] used an LSTM architecture. Regardless of the architecture choice, we need to consider a class of neural networks that is rich enough to be able to approximate U_0. In our case, we used a smoothed ReLU activation at each layer, ELU. We train the parameter θ using stochastic gradient descent so that $\mathcal{U}(\cdot, \cdot; \theta)$ fits (1.5) as well as possible.

Algorithm 1 Ergodic DGM

1: **Input:** An initial parameter vector θ.
2: **Output:** A trained vector $\hat{\theta}$ such that $(\varrho, \mathcal{U}(\cdot, \cdot; \hat{\theta}))$ approximately solves (1.5).
3: Compute:
$$\hat{\theta} \in \underset{\theta \in \mathbb{R}^{\bar{\delta}}}{\arg\min} \, L(\theta)$$
where for a finite set of samples $\mathcal{S} := \{S \mid S = (x, \eta)\}$, we define L to be the mean squared error:
$$L(\theta) := \frac{1}{|\mathcal{S}|} \sum_{(x,\eta) \in \mathcal{S}} \left\{ \left| H(x, \Delta_x \mathcal{U}(\cdot, \eta; \theta)) + \eta_x + \sum_{y,z \in [d]} \eta_y D_{yz}^\eta \mathcal{U}(x, \eta; \theta) \gamma_z^*(y, \Delta_y \mathcal{U}(\cdot, \eta; \theta)) - \varrho \right|^2 \right.$$
$$\left. + \left| \sum_{y \in [d]} \mathcal{U}(y, \eta; \theta) \right|^2 \right\}.$$

Remark 7.1. The first term in the loss function of Algorithm 1 is the master equation evaluated at uniformly sampled points of $[d] \times \mathcal{P}([d])$. The second term, however, centralizes the potential function around 0. As noted in Proposition 2.2, U_0 is only unique up to an additive constant. Therefore, adding the centralization term encourages the algorithm to reach the same solution on different runs.

Remark 7.2. Unfortunately, the master equation is a different kind of PDE than those studied in [33], and so the theoretical guarantees of the algorithm's correctness do not apply. While we do not offer a proof of correctness, applying the DGM to the finite-horizon master equation was pioneered in [28, Algorithm 7] and proofs of algorithmic correctness for that case were offered in [11].

In Figure 1, we present the difference in strategies, where the strategy $\gamma_y^*(x, \Delta_x U_0(\cdot, \eta))$ is approximated by plugging in the result of the ergodic DGM Algorithm 1; namely, $\gamma_y^*(x, \Delta_x \mathcal{U}(\cdot, \eta; \hat{\theta}))$. In the case we study numerically in Section 7.1, for any $D \in \mathbb{R}^d$, $\gamma_y^*(x, \Delta_x D)$ is a linear transformation of the vector $\Delta_x D$. We note that Γ_0 has a greater effect near the boundary of the simplex; this has the effect of moving a representative player faster toward the center.

We perform the same analysis for the case $d = 3$ and observe the differences in the strategies in Figure 2. The vertical axis represents the proportion of players in state 1 and the horizontal the proportion in state 2. The proportion of players in state 3 is 1 in the lower, left corner of the plots. Like in Figure 1, we note that Γ_0 results in a greater rate of transition when more players are in one's own state.

In Figure 3, we note that the difference in rates leads to a greater concentration of the stationary distribution $\mathscr{L}_\infty^n$ around the center of the simplex $\mathcal{P}([2])$ for Γ_0 than for $\bar{\Gamma}$. We computed the stationary distribution according to the formula for a birth-death process. Specifically, [21, Part III, Chapter 14] lists the formula for the stationary distribution of a birth-death process with specified birth and death rates for every state; we then consider "births" as when the empirical distribution increases its proportion in state 1 and "deaths" when it increases its proportion in state 2.

Given the stationary distributions for μ^{Γ_0} and $\mu^{\bar{\Gamma}}$, we can compute the exact cost to a single player in each ε-Nash equilibria. We write:
$$\varrho^{n, U_0} := J^n(\Gamma_0), \qquad \varrho^{n, \bar{u}} := J^n(\bar{\Gamma}).$$
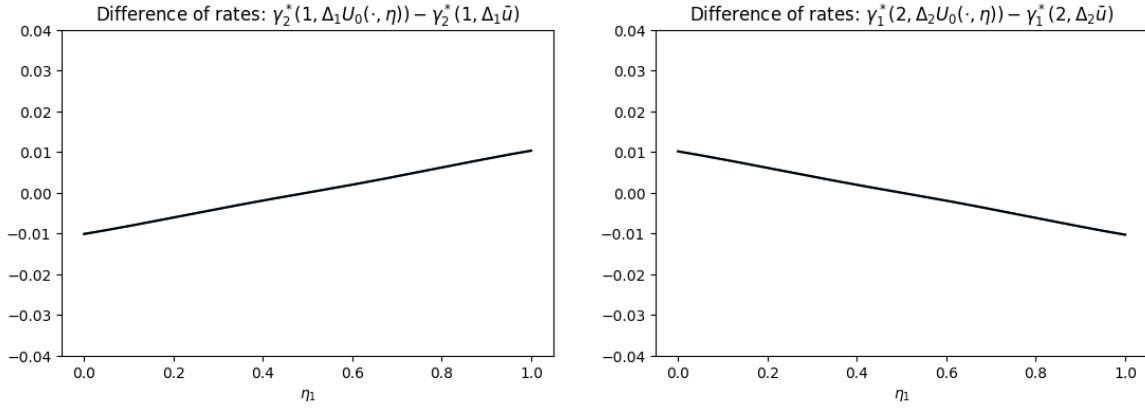
FIGURE 1. We can compare Γ_0 and $\bar{\Gamma}$ by observing the difference in rates $\gamma_y^*(x, \Delta_x U_0(\cdot, \eta)) - \gamma_y^*(x, \Delta_x \bar{u})$ in the case $d = 2$.
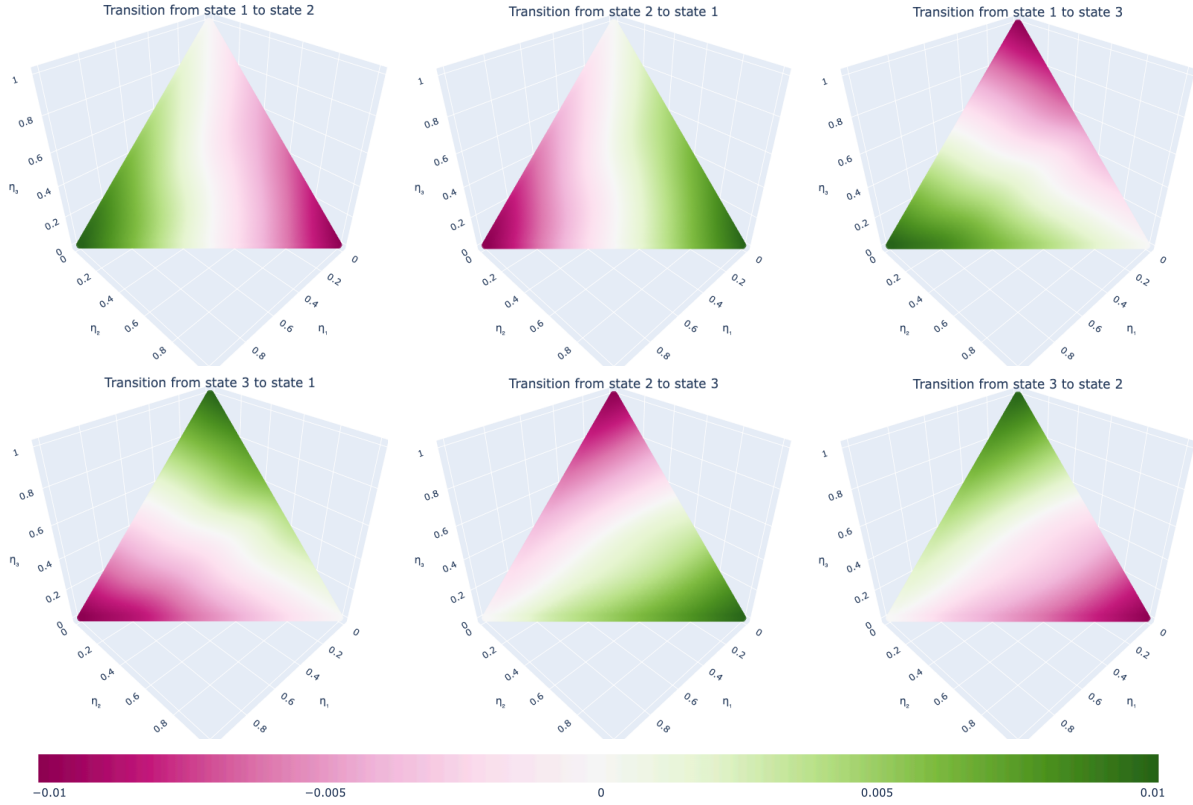

FIGURE 2. We can compare Γ_0 and $\bar{\Gamma}$ by observing the difference in rates $\gamma_y^*(x, \Delta_x U_0(\cdot, \eta)) - \gamma_y^*(x, \Delta_x \bar{u})$ in the case $d = 3$.

In Figure 4, we represent ϱ^{n,U_0} and $\varrho^{n,\bar{u}}$ as n increases. We see that ϱ^{n,U_0} is slightly smaller and so there is a marginal advantage to playing the profile Γ_0 associated with the master equation.

8. Large Deviations Numerics

In what follows, we provide an explicit formula for the good rate function of $\mathscr{L}_\infty^n$ when $d = 2$. We do this by introducing natural choices for functions $r, \bar{r}$ that satisfy an asymptotic property sufficient for proving Theorem 8.1.
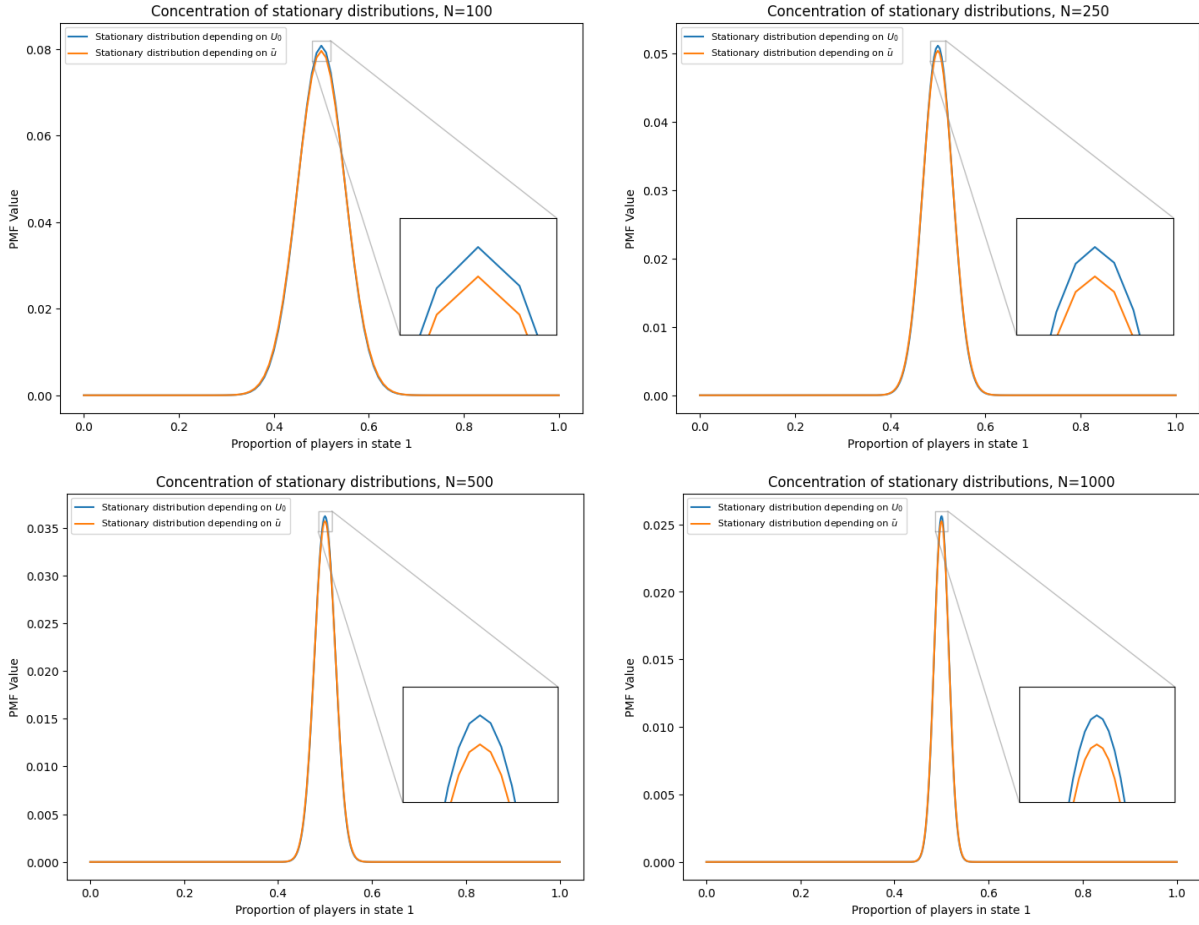
FIGURE 3. Due to the elevated rate for players of Γ_0 near the boundary of $\mathcal{P}([2])$, we see a greater concentration of mass toward the center of the simplex, relative to $\bar{\Gamma}$.
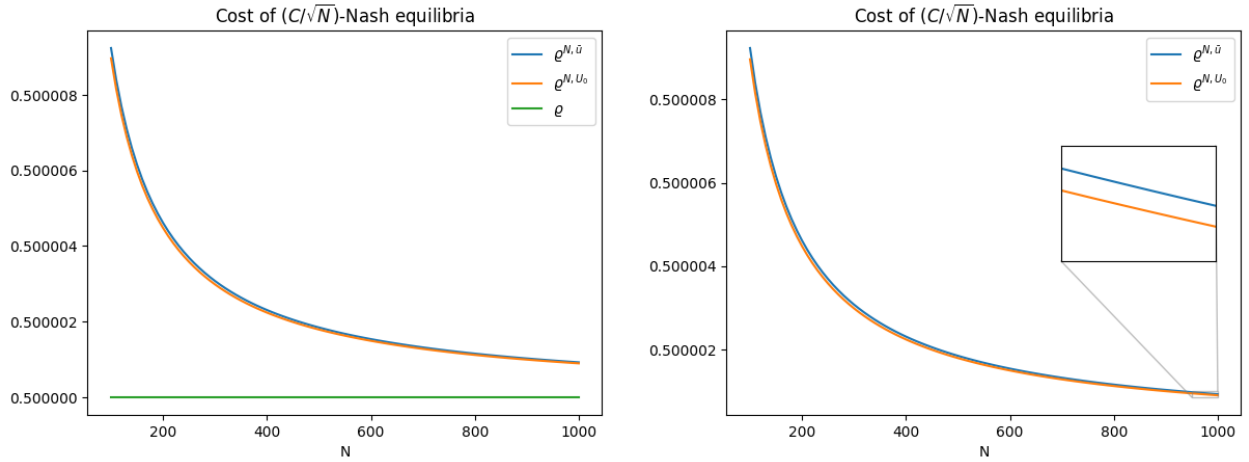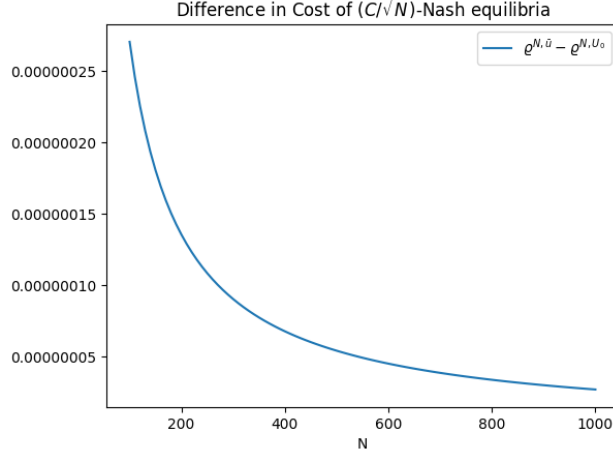

FIGURE 4. The left graph shows the convergence of each realized cost to ϱ the optimal cost in the mean field limit. The right graph highlights the fact that $\varrho^{n,U_0} < \varrho^{n,\bar{u}}$.

FIGURE 5. The difference in realized value $\varrho^{n,\bar{u}} - \varrho^{n,U_0}$.

As noted briefly in Section 7, when $d = 2$ we can interpret the empirical measures under $\bar{\Gamma}$ and Γ_0 as birth-death processes where movement into state 1 is interpreted as a birth and movement into state 2 is interpreted as a death. For $\bar{\Gamma}$ then, we can write the birth and death rates as the functions:

$$
\begin{aligned}
\bar{b}^{(n)}(k/n) &= \gamma_1^*(2, \Delta_2 \bar{u})n(1 - k/n), \\
\bar{d}^{(n)}(k/n) &= \gamma_2^*(1, \Delta_1 \bar{u})n(k/n).
\end{aligned}
\tag{8.1}
$$

Similarly, for Γ_0:

$$
\begin{aligned}
b^{(n)}(k/n) &= \gamma_1^*(2, \Delta_2 U_0(\cdot, (k/(n-1), 1 - k/(n-1))))n(1 - k/n), \\
d^{(n)}(k/n) &= \gamma_2^*(1, \Delta_1 U_0(\cdot, (k/(n-1), 1 - k/(n-1))))n(k/n).
\end{aligned}
\tag{8.2}
$$

We now propose two functions $\bar{r}, r_0 : [0,1] \to \mathbb{R}$ that, for the case $d = 2$ only, will allow us to define a simpler form for the rate function than in Theorem 3.3. Set:

$$
\begin{aligned}
\bar{r}(\lambda) &:= \left(\frac{\gamma_2^*(1, \Delta_1 \bar{u})}{\gamma_1^*(2, \Delta_2 \bar{u})} \right) \frac{\lambda}{1 - \lambda}, \\
r_0(\lambda) &:= \left(\frac{\gamma_2^*(1, \Delta_1 U_0(\cdot, (\lambda, 1-\lambda)))}{\gamma_1^*(2, \Delta_2 U_0(\cdot, (\lambda, 1-\lambda)))} \right) \frac{\lambda}{1 - \lambda}.
\end{aligned}
\tag{8.3}
$$

Lemma 8.1. *The functions* $\log(\bar{r}), \log(r_0) \in L^1([0,1])$; *that is, the log of each function is absolutely integrable. Moreover,*

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n-1} \left(\log \left(\frac{\bar{d}^{(n)}((k+1)/n)}{\bar{b}^{(n)}(k/n)} \right) - \log(\bar{r}(k/n)) \right) = 0,
\tag{8.4}
$$

and r_0 satisfies the same limit with $d^{(n)}$ and $b^{(n)}$.

Proof. Recall that the rates are bounded in a compact set away from zero. Therefore, using that $\log(ab) = \log(a) + \log(b)$, for any $r \in \{\bar{r}, r_0\}$,

$$
\int_0^1 |\log r(\lambda)| d\lambda \le C + \int_0^1 \left| \log \left(\tfrac{\lambda}{1-\lambda} \right) \right| d\lambda = C + \log(4) < \infty.
$$

We check (8.4) for $r = \bar{r}$. We now combine the log's in (8.4), use (8.1), and use that the product is telescoping to get:

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n-1} \log \left(\frac{k+1}{n-k} \cdot \frac{n-k}{k} \right) = \lim_{n \to \infty} \frac{1}{n} \log \left(\prod_{k=1}^{n} (k+1)/k \right) = \lim_{n \to \infty} \frac{\log(n+1)}{n} = 0.
\tag{8.5}
$$

The proof for r_0 is similar to that of $\bar{r}$, though we use at one point that U_0 is Lipschitz. $\qquad\square$

Theorem 8.1 (Large Deviations, Special Case $d = 2$). Recall from Theorem 3.3 that $\mathscr{L}_\infty^n \in \mathcal{P}(\mathcal{P}^n([d]))$ is the stationary distribution of the empirical distribution $\kappa^n \in \{\mu^{\bar{\Gamma}}, \mu^{\Gamma_0}\}$ and let the $r \in \{\bar{r}, r_0\}$ be the corresponding function from Lemma 8.1. When $d = 2$, $\mathscr{L}_\infty^n$ satisfies a large deviation principle with a simplified good rate function s given by:

$$s(\eta) = \tilde{s}(\eta) - \tilde{s}_{min},$$

where:

$$\tilde{s}(\eta) := \int_0^{\eta_1} \log(m(\lambda))d\lambda,$$
$$\tilde{s}_{min} := \min_{\eta \in \mathcal{P}([2])} \tilde{s}(\eta). \tag{8.6}$$

And in the case of $\kappa^n = \mu^{\bar{\Gamma}}$, we have explicit formulas for $\tilde{s}$ and $\tilde{s}_{min}$:

$$\tilde{s}(\eta) = \eta_1 \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) + \log(\eta_2) + \eta_1 \log(\eta_1/\eta_2),$$
$$\tilde{s}_{min} = \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) - \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})} + 1\right). \tag{8.7}$$

Proof. The purpose of Lemma 8.1 is to satisfy the assumptions from [10, (3.3a), (3.3b)]. As a consequence, [10, Theorem 3.1] asserts that s is the good rate function as claimed. It remains to show that, in the case of $\bar{r}$, we have an explicit form of s.

That is, we can integrate, using that $\log(ab) = \log(a) + \log(b)$, integration by parts, and the fact that $\eta_1 = 1 - \eta_2$,

$$\tilde{s}(\eta) = \int_0^{\eta_1} \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\frac{\lambda}{1-\lambda}\right)d\lambda$$
$$= \eta_1 \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) + \int_0^{\eta_1} \log\left(\frac{\lambda}{1-\lambda}\right)d\lambda$$
$$= \eta_1 \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) + \log(1-\lambda) + \lambda \log\left(\frac{\lambda}{1-\lambda}\right)\Big|_0^{\eta_1}$$
$$= \eta_1 \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) + \log(\eta_2) + \eta_1 \log(\eta_1/\eta_2).$$

Taking the derivative,

$$\frac{\partial}{\partial \eta_1}\tilde{s}(\eta) = \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\frac{\eta_1}{1-\eta_1}\right).$$

Therefore, we can find that $\tilde{s}$ has a critical point at:

$$\eta = (\eta_1, \eta_2), \quad \text{where} \quad \eta_1 = \left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})} + 1\right)^{-1}, \quad \eta_2 = 1 - \eta_1.$$

By the first derivative test, this point is the minimum of $\tilde{s}$. Plugging this point into $\tilde{s}$ and by some arithmetic,

$$\tilde{s}_{min} = \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})}\right) - \log\left(\frac{\gamma_2^*(1,\Delta_1\bar{u})}{\gamma_1^*(2,\Delta_2\bar{u})} + 1\right).$$

$\square$

In the numerical case we visited in the prior section, this means that the good rate function for the $\bar{\Gamma}$-system is given by:

$$s(\eta) = \log(\eta_2) + \eta_1 \log(\eta_1/\eta_2) + \log(2).$$

Unfortunately, the good rate function associated with the Γ_0-system is not explicitly solvable since the integral contains U_0. Nonetheless, since we approximately solved for U_0 using Algorithm 1, we can use Monte Carlo integration to approximate $\tilde{s}$ in that case and thus obtain the good rate function. In Figure 6, we compare the good rate functions for $\bar{\Gamma}$ and Γ_0.
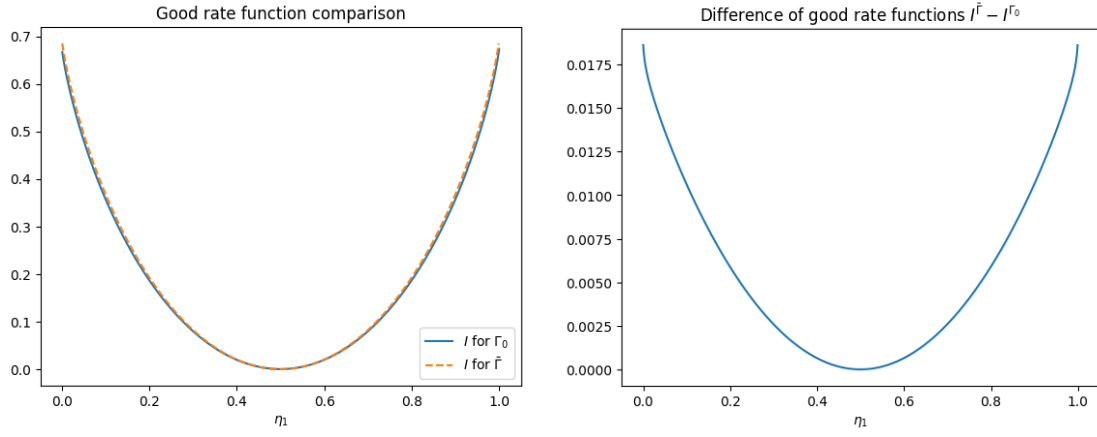
FIGURE 6. On the left-hand side, we compare the good rate functions associated with the invariant measures for μ^{Γ_0} and $\mu^{\bar{\Gamma}}$. Theorem 8.1 gives the rate function associated with $\mu^{\bar{\Gamma}}$ while Algorithm 1 and Monte Carlo integration obtain the rate function for μ^{Γ_0}. On the right, we highlight the difference between the rate functions. We note that the graph on the right-hand side implies that playing the profile $\bar{\Gamma}$ is associated with larger deviations as compared with Γ_0, and that the distinction is most pronounced toward the boundary of $\mathcal{P}([2])$.

APPENDIX A. PROOF OF LEMMA 3.1

Note that $(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{t\geq 0}$ is a Markov process. Denote its distributional law at time t by $(\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{\boldsymbol{x}} = \mathbb{P}(\boldsymbol{X}_t^{\boldsymbol{\beta}} = \boldsymbol{x}))_{\boldsymbol{x}\in[d]^n}$. Then, it satisfies the Kolmogorov equation:

$$\frac{d}{dt}\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{\boldsymbol{x}} = \sum_{\boldsymbol{y}\in[d]^n} \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}})_{\boldsymbol{y}} Q_{\boldsymbol{y}\boldsymbol{x}}^{\boldsymbol{\beta}}(t, \boldsymbol{X}_t^{\boldsymbol{\beta}}), \quad (t, \boldsymbol{x}) \in \mathbb{R}_+ \times [d]^n,$$

where $Q_{\boldsymbol{y}\boldsymbol{x}}^{\boldsymbol{\beta}}(t, \boldsymbol{X}_t^{\boldsymbol{\beta}})$ is the rate matrix:

$$Q_{\boldsymbol{y}\boldsymbol{x}}^{\boldsymbol{\beta}}(t, \boldsymbol{X}_t^{\boldsymbol{\beta}}) := \begin{cases} 0, & \boldsymbol{y} \neq \boldsymbol{x} + e_y - e_x, \text{ for any } y \neq x, \\ \sum_{i\in[n]} \beta_{xy}^i(t, \boldsymbol{X}_t^{\boldsymbol{\beta}})\mathbb{1}_{\{X_t^{\boldsymbol{\beta},i}=x\}}, & \boldsymbol{y} = \boldsymbol{x} + e_y - e_x, \text{ for some } y \neq x, \\ \sum_{i\in[n]} \beta_{xx}^i(t, \boldsymbol{X}_t^{\boldsymbol{\beta}})\mathbb{1}_{\{X_t^{\boldsymbol{\beta},i}=x\}}, & \boldsymbol{y} = \boldsymbol{x}. \end{cases}$$

Depending on the initial conditions π_0 or π_1, we know by [12, Lemma 3.3] that:

$$|\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}} \mid \boldsymbol{X}_0^{\boldsymbol{\beta}} \sim \pi_0) - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}} \mid \boldsymbol{X}_0^{\boldsymbol{\beta}} \sim \pi_1)| \leq K_n e^{-k_n t}|\pi_0 - \pi_1|, \tag{A.1}$$

where K_n, k_n are positive constants depending only on $\mathbb{A}$ and n. Hence,

$$\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}_{\pi_0}\int_0^T g(\boldsymbol{X}_t^{\boldsymbol{\beta}})dt = \limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}}dt$$

$$= \limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\big[\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}} - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}\big]dt$$

$$+ \limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}dt \qquad\text{(A.2)}$$

$$= \limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\big[\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}} - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}\big]dt$$

$$+ \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}_{\pi_1}\int_0^T g(\boldsymbol{X}_t^{\boldsymbol{\beta}})dt.$$

But then by triangle inequality, the boundedness of g, and finally (A.1),

$$\bigg|\limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\big[\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}} - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}\big]dt\bigg|$$

$$\le \limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}g(\boldsymbol{x})\big|\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}} - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}\big|dt$$

$$\le C_g\limsup_{T\to\infty}\frac{1}{T}\int_0^T\sum_{\boldsymbol{x}\in[d]^n}\big|\mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_0)_{\boldsymbol{x}} - \mathcal{L}(\boldsymbol{X}_t^{\boldsymbol{\beta}}\mid\boldsymbol{X}_0^{\boldsymbol{\beta}}\sim\pi_1)_{\boldsymbol{x}}\big|dt \qquad\text{(A.3)}$$

$$\le C_g\limsup_{T\to\infty}\frac{1}{T}\int_0^T K_n e^{-k_n t}|\pi_0 - \pi_1|dt$$

$$= 0.$$

Using this fact in (A.2) concludes the proof. $\qquad\qquad\square$

References

[1] A. Arapostathis, A. Biswas, and J. Carroll. On solutions of mean field games with ergodic cost. Journal de Mathématiques Pures et Appliquées, 107(2):205–251, Feb. 2017.

[2] E. Bayraktar and A. Cohen. Analysis of a finite state many player game using its master equation. SIAM Journal on Control and Optimization, 56(5):3538–3568, 2018.

[3] A. Bensoussan, J. Frehse, and S. C. P. Yam. The master equation in mean field theory. J. Math. Pures Appl. (9), 103(6):1441–1474, 2015.

[4] V. S. Borkar and R. Sundaresan. Asymptotics of the invariant measure in mean field models with jumps. Stoch. Syst., 2(2):322–380, 2012.

[5] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. The master equation and the convergence problem in mean field games, volume 201 of Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2019.

[6] P. Cardaliaguet and A. Porretta. Long time behavior of the master equation in mean-field game theory. Analysis & PDE, 2017.

[7] P. Cardaliaguet and C. Rainer. An example of multiple mean field limits in ergodic differential games. NoDEA Nonlinear Differential Equations Appl., 27(3):Paper No. 25, 19, 2020.

[8] R. Carmona and M. Laurière. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games I: The ergodic case. SIAM J. Numer. Anal., 59(3):1455–1485, 2021.

[9] A. Cecchin and G. Pelino. Convergence, fluctuations and large deviations for finite state mean field games via the master equation. Stochastic Processes and their Applications, 129(11):4510 – 4555, 2019.

[10] T. Chan. Large deviations and quasi-stationarity for density-dependent birth-death processes. J. Austral. Math. Soc. Ser. B, 40(2):238–256, 1998.

[11] A. Cohen, M. Laurière, and E. Zell. Deep Backward and Galerkin Methods for the Finite State Master Equation. arXiv e-prints, Mar. 2024.

[12] A. Cohen and E. Zell. Analysis of the Finite-State Ergodic Master Equation. Applied Mathematics and Optimization, 87, 2023.

[13] T. Daras. Hoeffding Type Inequalities and their Applications in Statistics and Operations Research. AIP Conference Proceedings, 936, Sep 2007.

[14] M. Fischer. On the connection between symmetric n-player games and mean field games. Ann. Appl. Probab., 127(2):757–810, 2017.

[15] W. Gangbo, A. R. Mészáros, C. Mou, and J. Zhang. Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity. Ann. Probab., 50(6):2178–2217, 2022.

[16] D. A. Gomes, J. Mohr, and R. R. Souza. Discrete time, finite state space mean field games. J. Math. Pures Appl. (9), 93(3):308–328, 2010.

[17] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. Appl. Math. Optim., 68(1):99–143, 2013.

[18] F. Höfer and M. Soner. Synchronization games. arxiv preprint, 2024.

[19] M. Huang, P. E. Caines, and R. P. Malhamé. The Nash certainty equivalence principle and Mckean-Vlasov systems: An invariance principle and entry adaptation. In Decision and Control, 2007 46th IEEE Conference on, pages 121–126. IEEE, 2007.

[20] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. Commun. Inf. Syst., 6(3):221–251, 2006.

[21] H. Kobayashi, B. L. Mark, and W. Turin. Probability, random processes, and statistical analysis. Cambridge University Press, Cambridge, 2012.

[22] D. Lacker. A general characterization of the mean field limit for stochastic differential games. Probab. Theory Related Fields, pages 1–68, 2015.

[23] D. Lacker. Limit theory for controlled mckean-vlasov dynamics. SIAM J. Control Optim., 55:1641–1672, 2017.

[24] D. Lacker. On the convergence of closed-loop nash equilibria to the mean field game limit. Ann. Appl. Probab., 30, 2020.

[25] D. Lacker and L. Le Flem. Closed-loop convergence for mean field games with common noise. Ann. Appl. Probab., 33(4):2681–2733, 2023.

[26] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. C. R. Math. Acad. Sci. Paris, 343(9):619–625, 2006.

[27] J.-M. Lasry and P.-L. Lions. Mean field games. Jpn. J. Math., 2(1):229–260, 2007.

[28] M. Laurière. Numerical methods for mean field games and mean field type control. In Mean field games, volume 78 of Proc. Sympos. Appl. Math., pages 221–282. Amer. Math. Soc., Providence, RI, 2021.

[29] C. Léonard. Large deviations for long range interacting particle systems with jumps. Ann. Inst. H. Poincaré Probab. Statist., 31(2):289–323, 1995.

[30] C. Mou and J. Zhang. Wellposedness of second order master equations for mean field games with nonsmooth data. arXiv e-prints, page arXiv:1903.09907, Mar. 2019.

[31] C. Mou and J. Zhang. Mean Field Game Master Equations with Anti-monotonicity Conditions. arXiv e-prints, page arXiv:2201.10762, Jan. 2022.

[32] B. A. Neumann. Stationary equilibria of mean field games with finite state and action space. Dynamic Games and Applications, 10(4):845–871, 2020.

[33] J. Sirignano and K. Spiliopoulos. DGM: a deep learning algorithm for solving partial differential equations. J. Comput. Phys., 375:1339–1364, 2018.

[34] L. Ying. On the approximation error of mean-field models. Stochastic Systems, 8(2):126–142, 2018.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, United States
Email address: shloshim@gmail.com

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, United States
Email address: ezell@umich.edu