

Discrete-time average-cost mean-field games on Polish spaces

Naci SALDI* 

Department of Natural and Mathematical Sciences, Faculty of Engineering, Özyeğin University, İstanbul, Turkey

Received: 02.05.2019

Accepted/Published Online: 28.01.2020

Final Version: 17.03.2020

Abstract: In stochastic dynamic games, when the number of players is sufficiently large and the interactions between agents depend on empirical state distribution, one way to approximate the original game is to introduce infinite-population limit of the problem. In the infinite population limit, a generic agent is faced with a so-called mean-field game. In this paper, we study discrete-time mean-field games with average-cost criteria. Using average cost optimality equation and Kakutani's fixed point theorem, we establish the existence of Nash equilibria for mean-field games under drift and minorization conditions on the dynamics of each agent. Then, we show that the equilibrium policy in the mean-field game, when adopted by each agent, is an approximate Nash equilibrium for the corresponding finite-agent game with sufficiently many agents.

Key words: Mean-field games, average cost, approximate Nash equilibrium

1. Introduction

In this paper, we consider discrete-time mean-field games subject to average-cost criteria with Polish state and action spaces. These games arise as the infinite population limit of finite-agent dynamic games, where agents interact through the empirical distribution of their states. The main goal in mean-field games is to establish the existence of an optimal policy and a state distribution that are consistent with each other. Then, this optimal policy, when adopted by each agent, forms an approximate Nash equilibrium for finite-agent games with sufficiently many agents.

Mean-field games have been introduced by Huang et al. [18] and Lasry and Lions [20] to approximate continuous-time differential games with a large but finite number of identical agents interacting with each other via empirical distribution of their states (i.e. mean-field term). The key feature of this approach is to transform the game problem into a nonclassical stochastic control problem by passing to the infinite-population limit. In the infinite population limit, since empirical state distribution converges to a deterministic probability measure by the law of large numbers, agents are decoupled from each other and each agent is faced with a stochastic control problem that has a constraint on the distribution of its state. The latter problem is called mean-field game in the literature. The optimal solution of this stochastic control problem provides an approximate Nash equilibrium when the number of agents is sufficiently large. In continuous-time differential games, this optimal solution is characterized by a Fokker–Planck equation evolving forward in time and a Hamilton–Jacobi–Bellman equation evolving backward in time. We refer the reader to [3, 7, 8, 11, 16, 17, 23, 29] for studies of

*Correspondence: naci.saldi@ozyegin.edu.tr

2010 *AMS Mathematics Subject Classification:* 91A15, 91A10, 91A13, 93E20.

continuous-time mean-field games with different models and cost functions, such as games with major-minor players, risk-sensitive games, games with Markov jump parameters, and LQG games.

Discrete-time mean-field games have not been studied much in the literature. Existing works have mostly studied games with finite or countable state spaces subject to finite-horizon or infinite-horizon discounted cost criteria. Gomes et al. [10] considered a discrete-time mean-field game with a finite state space over a finite horizon. In [1], discrete-time mean-field game with countable state-space was studied subject to an infinite-horizon discounted cost criterion. The authors in [9, 21, 22, 24] consider discrete-time mean-field games with linear state dynamics. There are only three papers [5, 31, 32] studying discrete-time mean-field games subject to average cost criteria. In [32], the authors considered discrete set-up for average-cost mean-field games. In [5], the author considered average-cost mean-field games with σ -compact Polish state spaces. In that paper, it was assumed that, for the finite agent game problem, the dynamics of agents do not depend on the empirical distribution of the states. Under strong regularity conditions on system components, Biswas [5] established the existence of Nash equilibria for finite-agent games and then showed that these Nash equilibria converge to mean-field equilibria in the infinite-population limit. These imposed regularity conditions are in general prohibitive because they are stated in terms of a specific metric topology on the set of policies, and appear to be too strong to hold under reasonable assumptions. Wiecek [31] considered average-cost mean-field games with compact state spaces. This setup is the closest to the one studied in this paper. However, in addition to the state space being compact, the author in [31] also used a completely different technique to establish the existence of equilibrium in the infinite-population limit. Namely, the author in [31] proved the existence of mean-field equilibrium using ergodic properties of Markov chains induced by policies whereas we employ here the dynamic programming principle, stated through average cost equation, to establish the existence of mean-field equilibrium.

In [26] we studied infinite-horizon discounted-cost version of the same problem. Under mild assumptions on the system components, we first established the existence of mean-field equilibrium and then proved that the equilibrium in the mean-field game constitutes an approximate Nash equilibrium for finite-agent games when the number of agents is sufficiently large. However, in the average-cost set-up, it is infeasible to establish similar results under similar assumptions as analysis of average-cost criterion is much more difficult than discounted-cost criterion. Therefore, to establish similar existence and approximation results, we impose drift and minorization conditions on the system dynamics of each agent, which are a bit strong but are quite common to study average-cost stochastic control problems.

The paper is organized as follows. In Section 2, we introduce the infinite-population mean-field game and define mean-field equilibrium. In Section 3, we formulate the finite-agent game problem of the mean-field type. In Section 4, we prove the existence of a mean-field equilibrium. In Section 5 we establish that the mean-field equilibrium policies lead to an approximate Nash equilibrium for finite-agent games with sufficiently many agents. Section 6 concludes the paper.

Notation. For a metric space E , we let $C_b(E)$ denote the set of all bounded continuous real functions on E endowed with sup-norm $\|g\| = \sup_{e \in E} |g(e)|$, which turns $C_b(E)$ into a Banach space. Let $\mathcal{P}(E)$ denote the set of all Borel probability measures on E . A sequence $\{\mu_n\}$ of measures on E is said to converge weakly to a measure μ if $\int_E g(e) \mu_n(de) \rightarrow \int_E g(e) \mu(de)$ for all $g \in C_b(E)$. We endow $\mathcal{P}(E)$ with weak topology induced by weak convergence of probability measures. This topology is known to be metrizable, and if E is complete and separable, then $\mathcal{P}(E)$ is also complete and separable under weak topology. For metric spaces E_1 and E_2 , a

stochastic kernel $\gamma(\cdot|e_1)$ (or *regular conditional probability measure*) on E_2 given E_1 is a measurable function $\gamma : E_1 \rightarrow \mathcal{P}(E_2)$. A probability measure μ is called an invariant probability measure of a stochastic kernel γ on E given E if $\mu(\cdot) = \int_E \gamma(\cdot|e) \mu(de)$. For any subset B of E , we let ∂B and B^c denote the boundary and complement of B , respectively. The notation $v \sim \nu$ means that the random element v has distribution ν . Unless otherwise specified, the term “measurable” will refer to Borel measurability.

2. Mean-field games and mean-field equilibria

The discrete-time mean-field game model is specified by (X, A, p, c, μ_0) , where X and A are the Polish (complete and separable metric space) state and action spaces, respectively. The stochastic kernel $p : X \times A \times \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ denotes the transition probability law of the next state given the previous state-action pair and state-measure. The measurable function $c : X \times A \times \mathcal{P}(X) \rightarrow [0, \infty)$ is the one-stage cost function. The probability measure μ_0 denotes the initial distribution of the state. A policy π is a stochastic kernel on A given X ; that is, $\pi : X \rightarrow \mathcal{P}(A)$ is a measurable function, where $\mathcal{P}(A)$ is endowed with the Borel σ -algebra generated by the weak convergence of probability measures. Let Π denote the set of all policies. According to the Ionescu Tulcea Theorem [13], an initial distribution μ_0 , a policy π , and a transition probability p define a unique probability measure P^{π, μ_0} on $(X \times A)^\infty$. The expectation with respect to P^{π, μ_0} is denoted by E^{π, μ_0} .

Given any state-measure $\mu \in \mathcal{P}(X)$, a policy $\pi^* \in \Pi$ is optimal for μ if

$$J_\mu(\pi^*, \mu_0) = \inf_{\pi \in \Pi} J_\mu(\pi, \mu_0),$$

where

$$J_\mu(\pi, \mu_0) = \limsup_{T \rightarrow \infty} \frac{1}{T} E^{\pi, \mu_0} \left[\sum_{t=0}^{T-1} c(x(t), a(t), \mu) \right]$$

is the average cost of policy π with state-measure μ . In this case, the evolution of the states and actions is given by

$$\begin{aligned} x(0) &\sim \mu_0, \quad x(t) \sim p(\cdot | x(t-1), a(t-1), \mu), \quad t \geq 1, \\ a(t) &\sim \pi(\cdot | x(t)), \quad t \geq 0. \end{aligned}$$

Define the set-valued mapping $\Psi : \mathcal{P}(X) \rightarrow 2^\Pi$ as $\Psi(\mu) = \{\pi \in \Pi : \pi \text{ is optimal for } \mu \text{ and } \mu_0 = \mu\}$.

Conversely, we define another set-valued mapping $\Lambda : \Pi \rightarrow 2^{\mathcal{P}(X)}$ as follows: given $\pi \in \Pi$, the state-measure μ_π is in $\Lambda(\pi)$ if it is a fixed point of the following equation:

$$\mu_\pi(\cdot) = \int_{X \times A} p(\cdot | x, a, \mu_\pi) \pi(da|x) \mu_\pi(dx).$$

Under Assumption 1, which is given below, one can prove that $\Lambda(\pi)$ has a unique element for all π . Therefore, it is indeed a single-valued mapping.

The notion of an equilibrium for the average-cost mean-field game is defined via these mappings Ψ, Λ as follows.

Definition 2.1 A pair $(\pi, \mu) \in \Pi \times \mathcal{P}(X)$ is a mean-field equilibrium if $\pi \in \Psi(\mu)$ and $\mu \in \Lambda(\pi)$.

The main goal in average-cost mean-field games is to establish the existence of a mean-field equilibrium. To that end, we impose the assumptions below on the components of the mean-field game model. Note that a function $w : X \rightarrow [0, \infty)$ is called a moment function if there exists a nondecreasing sequence of compact sets $K_n \uparrow X$ such that

$$\lim_{n \rightarrow \infty} \inf_{x \notin K_n} w(x) = \infty.$$

Assumption 1 (a) *The cost function c is bounded and continuous.*

(b) *The stochastic kernel p is weakly continuous; that is, if $(x_n, a_n, \mu_n) \rightarrow (x, a, \mu)$, then $p(\cdot | x_n, a_n, \mu_n) \rightarrow p(\cdot | x, a, \mu)$ weakly.*

(c) *A is compact.*

(d) *There exists a nondegenerate subprobability measure λ on X such that*

$$p(\cdot | x, a, \mu) \geq \lambda(\cdot)$$

for all $x \in X$, $a \in A$, and $\mu \in \mathcal{P}(X)$.

(e) *There exist a constant $\alpha \in (0, 1)$ and a continuous moment function $w : X \rightarrow [0, \infty)$ such that*

$$\sup_{(a, \mu) \in A \times \mathcal{P}(X)} \int_X w(y) p(dy | x, a, \mu) \leq \alpha w(x) + \int_X w(y) \lambda(dy). \tag{2.1}$$

If we define the following substochastic kernel $\hat{p}(\cdot | x, a, \mu) = p(\cdot | x, a, \mu) - \lambda(\cdot)$, then (2.1) can be written as

$$\sup_{(a, \mu) \in A \times \mathcal{P}(X)} \int_X w(y) \hat{p}(dy | x, a, \mu) \leq \alpha w(x). \tag{2.2}$$

Note that condition (e) is so-called ‘drift inequality’ and condition (d) is so-called ‘minorization’ condition, both of which were used in the literature for studying ergodicity of Markov chains (see [14], and references therein). These assumptions are quite general for studying average cost stochastic control problems. Indeed, Assumption 1-(d) is true when the transition probability satisfies conditions R1(a) and R1(b) in [15] (see also [15, Remark 3.3] and references therein for further conditions). For Assumption 1-(e), we refer the reader to the examples in [14, Section 7.4] to see under which conditions on the system components Assumption 1-(e) holds.

The main result of this section is the existence of a mean-field equilibrium under Assumption 1.

Theorem 2.2 *Under Assumption 1, the mean-field game (X, A, p, c) admits a mean-field equilibrium (π^*, μ^*) .*

It is important to note that mean-field games are not games in the strict sense. They are stochastic control problems subject to a constraint on the distribution of the state at each time step. In other words, we have a single agent and represent the collective behavior of (a large population of) other agents by an exogenous state-measure $\mu \in \mathcal{P}(X)$. This measure μ should also be consistent with the state distributions of this single agent when the agent acts optimally. The proof of Theorem 2.2 is given Section 4. To establish the existence of a mean-field equilibrium, we use dynamic programming principle for average-cost criterion, which is stated via average cost optimality equation (ACOE), in addition to fixed point approach that is commonly used in classical game problems.

3. Finite player game

The motivation for studying mean-field games comes from the challenges to establish the existence of Nash equilibria for large population stochastic dynamic games with mean-field interactions. More precisely, suppose that we have a discrete-time N -agent stochastic game with state space \mathbf{X} and action space \mathbf{A} . For every $t \in \{0, 1, 2, \dots\}$ and every $i \in \{1, 2, \dots, N\}$, let $x_i^N(t) \in \mathbf{X}$ and $a_i^N(t) \in \mathbf{A}$ denote the state and the action of Agent i at time t , and let

$$e_t^{(N)}(\cdot) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i^N(t)}(\cdot) \in \mathcal{P}(\mathbf{X})$$

denote the empirical distribution of the states at time t (i.e., *mean-field term*), where $\delta_x \in \mathcal{P}(\mathbf{X})$ is the Dirac measure at x . The initial states $x_i^N(0)$ are independent and identically distributed according to μ_0 , and, for each $t \geq 0$, next states $(x_1^N(t+1), \dots, x_N^N(t+1))$ are generated according to the probability distribution

$$\prod_{i=1}^N p(dx_i^N(t+1) | x_i^N(t), a_i^N(t), e_t^{(N)}).$$

A policy for a generic agent is a stochastic kernel π on \mathbf{A} given \mathbf{X} . The set of all policies for Agent i is denoted by Π_i . Let ${}^{(N)} = \prod_{i=1}^N \Pi_i$. We let $\boldsymbol{\pi}^{(N)} = (\pi^1, \dots, \pi^N)$, $\pi^i \in \Pi_i$, denote the N -tuple of policies for all the agents in the game. Under such N -tuple of policies, actions $(a_1^N(t), \dots, a_N^N(t))$ at each time $t \geq 0$ are generated according to the probability distribution

$$\prod_{i=1}^N \pi^i(da_i^N(t) | x_i^N(t)).$$

Note that agents can only use their local states when constructing their control laws. For Agent i , the average cost under the initial distribution μ_0 and N -tuple of policies $\boldsymbol{\pi}^{(N)} \in {}^{(N)}$ is given by

$$J_i^{(N)}(\boldsymbol{\pi}^{(N)}, \mu_0) = \limsup_{T \rightarrow \infty} \frac{1}{T} E^{\boldsymbol{\pi}^{(N)}, \mu_0} \left[\sum_{t=0}^{T-1} c(x_i^N(t), a_i^N(t), e_t^{(N)}) \right].$$

We now define the notion of Nash equilibrium for this game problem as follows.

Definition 3.1 An N -tuple of policies $\boldsymbol{\pi}^{(N^*)} = (\pi^{1^*}, \dots, \pi^{N^*})$ constitutes a Nash equilibrium if

$$J_i^{(N)}(\boldsymbol{\pi}^{(N^*)}, \mu_0) = \inf_{\pi^i \in \Pi_i} J_i^{(N)}(\boldsymbol{\pi}_{-i}^{(N^*)}, \pi^i, \mu_0)$$

for each $i = 1, \dots, N$, where $\boldsymbol{\pi}_{-i}^{(N^*)} = (\pi^{j^*})_{j \neq i}$.

It is known that it is in general prohibitive to establish the existence of Nash equilibrium under decentralized information structure. Moreover, when the number of agents is large, obtaining Nash equilibrium is computationally intractable. Therefore, it is of interest to prove the existence of approximate Nash equilibrium. To that end, we introduce the following solution concept:

Definition 3.2 An N -tuple of policies $\pi^{(N^*)} \in {}^{(N)}$ is an ε -Nash equilibrium (for a given $\varepsilon > 0$) if

$$J_i^{(N)}(\pi^{(N^*)}, \mu_0) \leq \inf_{\pi^i \in \Pi_i} J_i^{(N)}(\pi_{-i}^{(N^*)}, \pi^i, \mu_0) + \varepsilon$$

for each $i = 1, \dots, N$.

In mean field games, the aim is now to show that the policy π^* in the mean-field equilibrium, when adopted by each agent, is ε -Nash equilibrium for games with sufficiently many agents. To that end, we need to impose additional assumptions on the components of the game model.

Let ρ denote the following metric on $\mathcal{P}(X)$ that metrizes the weak topology:

$$\rho(\mu, \nu) = \sum_{m=0}^{\infty} 2^{-m} \left| \int_X f_m(x) \mu(dx) - \int_X f_m(x) \nu(dx) \right|,$$

where $f_m \in C_b(X)$ and $\|f_m\| \leq 1$ for all m (see [25, Theorem 6.6, p. 47]). Define the following moduli of continuity:

$$\omega_c(r) = \sup_{(x,a) \in X \times A} \sup_{\substack{\mu, \nu: \\ \rho(\mu, \nu) \leq r}} |c(x, a, \mu) - c(x, a, \nu)|.$$

In addition to Assumption 1, we impose an additional assumption, which is stated below.

Assumption 2 (a) The transition probability $p(\cdot | x, a)$ does not depend on μ .

(b) $\omega_c(r) \rightarrow 0$ as $r \rightarrow 0$.

For average cost criterion, Assumption 2-(a) is quite common to establish the existence of approximate Nash equilibrium (see [5, 31]). The following theorem is the main result of this section, which states that the joint policy $\pi^{(N^*)} = (\pi^*, \dots, \pi^*)$, where π^* is repeated N times, which itself is obtained from the mean-field equilibrium, is an ε -Nash equilibrium for sufficiently large N .

Theorem 3.3 For any $\varepsilon > 0$, there exists $N(\varepsilon)$ such that for $N \geq N(\varepsilon)$, the N -tuple of policies $\pi^{(N^*)}$ is an ε -Nash equilibrium for the game with N agents.

The proof of this theorem is given in Section 5.

4. Proof of Theorem 2.2

Under Assumption 1, for any (π, μ) , there exists a unique probability measure $\mu_{\pi, \mu}$ such that

$$\mu_{\pi, \mu}(\cdot) = \int_{X \times A} p(\cdot | x, a, \mu) \pi(da|x) \mu_{\pi, \mu}(dx) \tag{4.1}$$

(see [30, Theorem 3.3], [12, Lemma 3.4]). Given $\mu \in \mathcal{P}(X)$, we define the operator $T_\mu : C_b(X) \rightarrow C_b(X)$ as

$$T_\mu u(x) = \min_{a \in A} \left[c(x, a, \mu) + \int_X u(y) p(dy|x, a, \mu) - \int_X u(y) \lambda(dy) \right],$$

$$= \min_{a \in A} \left[c(x, a, \mu) + \int_{\mathbf{X}} u(y) \hat{p}(dy|x, a, \mu) \right].$$

Under Assumption 1, T_μ is a well-defined operator. One can also prove that T_μ is a contraction operator with modulus $\beta = 1 - \lambda(\mathbf{X}) \in (0, 1)$ [27, Theorem 3.21]. Therefore, for each μ , there exists a fixed point $h_\mu \in C_b(\mathbf{X})$ of T_μ by Banach Fixed Point Theorem. The following theorem characterizes the optimal policies for each state-measure μ using the operator T_μ .

Theorem 4.1 *Given $\mu \in \mathcal{P}(\mathbf{X})$, a policy π is optimal for μ when $x(0) \sim \mu_{\pi, \mu}$, if*

$$\nu_{\pi, \mu} \left(\left\{ (x, a) : c(x, a, \mu) + \int_{\mathbf{X}} h_\mu(dy) \hat{p}(dy|x, a, \mu) = T_\mu h_\mu \right\} \right) = 1,$$

where $\nu_{\pi, \mu}(dx, da) = \mu_{\pi, \mu}(dx) \pi(da|x)$ and h_μ is the fixed point of T_μ .

Proof Note that if $T_\mu h_\mu = h_\mu$, then we have

$$h_\mu(x) + \rho_\mu = \min_{a \in A} \left[c(x, a, \mu) + \int_{\mathbf{X}} h_\mu(dy) p(dy|x, a, \mu) \right],$$

where $\rho_\mu = \int_{\mathbf{X}} h_\mu(dx) \lambda(dx)$. The last equation is called average cost optimality equation (ACOE) in the literature [13]. Since

$$\lim_{n \rightarrow \infty} \frac{E^\pi [h_\mu(x(n))]}{n} = 0,$$

for all π , as h is bounded, we have [13, Theorem 5.2.4]

$$\rho_\mu = \inf_{\pi \in \Pi} J_\mu(\pi, \mu_{\pi, \mu}) =: J_\mu^*.$$

Suppose that π satisfies the hypothesis in the theorem. For each $n \geq 1$ and $\gamma \in \mathcal{P}(\mathbf{X})$, we define

$$J_{\mu, n}(\pi, h_\mu, \gamma) = E^\pi \left[\sum_{t=0}^{n-1} c(x(t), a(t), \mu) + h_\mu(x(n)) \mid x(0) \sim \gamma \right].$$

We claim that, for each $n \geq 1$,

$$J_{\mu, n}(\pi, h_\mu, \mu_{\pi, \mu}) = n\rho_\mu + \int_{\mathbf{X}} h_\mu(dx) \mu_{\pi, \mu}(dx).$$

Claim clearly holds for $n = 0$. Suppose it holds for n and consider $n + 1$:

$$\begin{aligned} J_{\mu, n+1}(\pi, h_\mu, \mu_{\pi, \mu}) &= \int_{\mathbf{X} \times \mathbf{A}} \left[c(x, a, \mu) + \int_{\mathbf{X}} J_{\mu, n}(\pi, h_\mu, \delta_y) p(dy|x, a, \mu) \right] \pi(da|x) \mu_{\pi, \mu}(dx) \\ &\stackrel{(a)}{=} \int_{\mathbf{X} \times \mathbf{A}} c(x, a, \mu) \pi(da|x) \mu_{\pi, \mu}(dx) + \int_{\mathbf{X}} J_{\mu, n}(\pi, h_\mu, \delta_y) \mu_{\pi, \mu}(dy) \\ &\stackrel{(b)}{=} \int_{\mathbf{X} \times \mathbf{A}} \left[c(x, a, \mu) + \int_{\mathbf{X}} [n\rho_\mu + h_\mu(dy)] p(dy|x, a, \mu) \right] \pi(da|x) \mu_{\pi, \mu}(dx) \end{aligned}$$

$$\begin{aligned}
 &= n\rho_\mu + \int_{\mathbf{X} \times \mathbf{A}} \left[c(x, a, \mu) + \int_{\mathbf{X}} h_\mu(dy) p(dy|x, a, \mu) \right] \nu_{\pi, \mu}(dx, da) \\
 &\stackrel{(c)}{=} n\rho_\mu + \int_{\mathbf{X} \times \mathbf{A}} \left[T_\mu h_\mu(x) + \rho_\mu \right] \nu_{\pi, \mu}(dx, da) \\
 &\stackrel{(d)}{=} (n+1)\rho_\mu + \int_{\mathbf{X} \times \mathbf{A}} h_\mu(x) \mu_{\pi, \mu}(dx),
 \end{aligned}$$

where (a) follows from (4.1), (b) follows from

$$J_{\mu, n}(\pi, h_\mu, \mu_{\pi, \mu}) = \int_{\mathbf{X}} J_{\mu, n}(\pi, h_\mu, \delta_x) \mu_{\pi, \mu}(dx)$$

and induction hypothesis, (c) follows from the hypothesis in the theorem, and (d) follows from $T_\mu h_\mu = h_\mu$. This completes the proof of the claim. Note that

$$\begin{aligned}
 J_\mu(\pi, \mu_{\pi, \mu}) &= \limsup_{n \rightarrow \infty} \frac{1}{n} J_{\mu, n}(\pi, 0, \mu_{\pi, \mu}) \\
 &= \limsup_{n \rightarrow \infty} \frac{1}{n} \left[J_{\mu, n}(\pi, h_\mu, \mu_{\pi, \mu}) - E^\pi[h_\mu(x(n))] \right] \\
 &\stackrel{(a)}{=} \limsup_{n \rightarrow \infty} \frac{1}{n} J_{\mu, n}(\pi, h_\mu, \mu_{\pi, \mu}) \\
 &= \limsup_{n \rightarrow \infty} \frac{1}{n} \left[n\rho_\mu + \int_{\mathbf{X}} h_\mu(dx) \mu_{\pi, \mu}(dx) \right] \\
 &= \rho_\mu = J_\mu^*,
 \end{aligned}$$

where (a) follows from the fact that h_μ is bounded. Therefore, π is optimal. □

Define the set $\mathcal{P}_c(\mathbf{X})$ as

$$\mathcal{P}_c(\mathbf{X}) = \left\{ \mu \in \mathcal{P}(\mathbf{X}) : \int_{\mathbf{X}} w(x) \mu(dx) \leq \frac{\int_{\mathbf{X}} w(x) \lambda(dx)}{1 - \alpha} \right\}.$$

Since w is a continuous moment function, $\mathcal{P}_c(\mathbf{X})$ is compact [13, Proposition E.8]. We also define

$$\Xi = \{ \nu \in \mathcal{P}(\mathbf{X} \times \mathbf{A}) : \nu(dx, da) = \mu(dx) \pi(da|x) \text{ for some } (\pi, \mu) \in \Pi \times \mathcal{P}_c(\mathbf{X}) \}.$$

Note that compactness of \mathbf{A} and compactness of $\mathcal{P}_c(\mathbf{X})$ imply that Ξ is tight [13, Definition E.5], and so, Ξ is relatively compact [13, Theorem E.6]. Since w is continuous, Ξ is closed. Therefore, Ξ is compact. For any $\nu \in \mathcal{P}(\mathbf{X} \times \mathbf{A})$, we let ν_1 denote the marginal distribution on \mathbf{X} .

Note that (π, μ) is a mean-field equilibrium if $\pi \in \Psi(\Lambda(\pi))$. We transform this fixed point equation $\pi \in \Psi(\Lambda(\pi))$ into a fixed point equation of a set-valued mapping from Ξ to $2^{\mathcal{P}(\mathbf{X} \times \mathbf{A})}$. To that end, we define the set-valued mapping $\Gamma : \Xi \rightarrow 2^{\mathcal{P}(\mathbf{X} \times \mathbf{A})}$ as follows:

$$\Gamma(\nu) = C(\nu) \cap B(\nu),$$

where

$$C(\nu) = \left\{ \nu' : \nu'_1(\cdot) = \int_{\mathbf{X} \times \mathbf{A}} p(\cdot | x, a, \nu_1) \nu(dx, da) \right\}$$

and

$$B(\nu) = \left\{ \nu' : \nu' \left(\left\{ (x, a) : c(x, a, \nu_1) + \int_{\mathbf{X}} h_{\nu_1}(y) \hat{p}(dy|x, a, \nu_1) = T_{\nu_1} h_{\nu_1}(x) \right\} \right) = 1 \right\}.$$

Note that the set $C(\nu)$ characterizes the equation $\mu = \Lambda(\pi)$ and the set $B(\nu)$ characterizes the equation $\pi \in \Psi(\mu)$. Under Assumption 1-(e), one can easily prove that the image of Ξ under Γ is contained in 2^Ξ ; that is, $\Gamma(\nu) \subset \Xi$. Indeed, fix any $\nu \in \Xi$. It is sufficient to prove that $C(\nu) \subset \Xi$. Let $\nu' \in C(\nu)$. We have

$$\begin{aligned} \int_{\mathbf{X}} w(y) \nu'_1(dy) &= \int_{\mathbf{X} \times \mathbf{A}} \int_{\mathbf{X}} w(y) p(dy|x, a, \nu_1) \nu(dx, da) \\ &\leq \int_{\mathbf{X}} \alpha w(x) \nu_1(dx) + \int_{\mathbf{X}} w(y) \lambda(dy) \quad (\text{by Assumption 1-(e)}) \\ &\leq \frac{\int_{\mathbf{X}} w(x) \lambda(dx)}{1 - \alpha} \quad (\text{as } \nu_1 \in \mathcal{P}_c(\mathbf{X})). \end{aligned}$$

Hence, $\nu'_1 \in \mathcal{P}_c(\mathbf{X})$. This completes the proof of the claim. Therefore,

$$\Gamma : \Xi \rightarrow 2^\Xi.$$

A point $\nu \in \Xi$ is a fixed point of Γ if $\nu \in \Gamma(\nu)$. The following proposition makes the connection between mean-field equilibria and the fixed points of Γ .

Proposition 4.2 *Suppose that Γ has a fixed point ν . Disintegrate ν as $\nu(dx, da) = \nu_1(dx) \pi(da|x)$. Then the pair (π, ν_1) is a mean-field equilibrium.*

Proof Let $\nu \in \Gamma(\nu)$. Since $\nu \in C(\nu)$, we have $\Lambda(\pi) = \nu_1 = \mu_{\pi, \nu_1}$. Moreover, since $\nu \in B(\nu)$, π is optimal for ν_1 when $x(0) \sim \nu_1$ by Theorem 4.1; that is, $\pi \in \Psi(\nu_1)$. This completes the proof. \square

By Proposition 4.2, it suffices to prove that Γ has a fixed point in order to establish the existence of a mean-field equilibrium. To prove this, we use Kakutani's fixed point theorem ([2, Corollary 17.55]). Note that Ξ is convex. Furthermore, it can be proved that $C(\nu) \cap B(\nu) \neq \emptyset$ for any $\nu \in \Xi$. Indeed, we define

$$\mu(\cdot) = \int_{\mathbf{X} \times \mathbf{A}} p(\cdot | x, a, \nu_1) \nu(dx, da).$$

Moreover, let $f : \mathbf{X} \rightarrow \mathbf{A}$ be the minimizer of the following optimality equation:

$$\begin{aligned} c(x, f(x), \nu_1) + \int_{\mathbf{X}} h_{\nu_1}(y) \hat{p}(dy|x, f(x), \nu_1) \\ = \inf_{a \in \mathbf{A}} \left[c(x, a, \nu_1) + \int_{\mathbf{X}} h_{\nu_1}(y) \hat{p}(dy|x, a, \nu_1) \right]. \end{aligned}$$

Existence of such an f follows from the Measurable Selection Theorem [13, Section D]. If we define $\nu'(dx, da) = \mu(dx) \delta_{f(x)}(da)$, then it is straightforward to prove that $\nu' \in C(\nu) \cap B(\nu)$, and so, $C(\nu) \cap B(\nu) \neq \emptyset$. We can also show that both $C(\nu)$ and $B(\nu)$ are convex; thus, their intersection is also convex. Moreover, Ξ is a convex compact subset of a locally convex topological space $\mathcal{M}(\mathbf{X} \times \mathbf{A})$, where $\mathcal{M}(\mathbf{X} \times \mathbf{A})$ denotes the set of finite signed measures on $\mathbf{X} \times \mathbf{A}$. Therefore, to establish the existence of a fixed point of Γ via Kakutani's fixed point, we only need to prove the following:

Proposition 4.3 *The graph of Γ , i.e., the set*

$$\text{Gr}(\Gamma) := \{(\nu, \xi) \in \Xi \times \Xi : \xi \in \Gamma(\nu)\},$$

is closed.

Proof Here, we adapt the proof of [26, Proposition 3.9] to the average cost criterion. Let $\{(\nu^{(n)}, \xi^{(n)})\}_{n \geq 1} \subset \Xi \times \Xi$ be such that $\xi^{(n)} \in \Gamma(\nu^{(n)})$ for all n and $(\nu^{(n)}, \xi^{(n)}) \rightarrow (\nu, \xi)$ as $n \rightarrow \infty$ for some $(\nu, \xi) \in \Xi \times \Xi$. We prove that $\xi \in \Gamma(\nu)$, which completes the proof.

We first prove that $\xi \in C(\nu)$. Note that, for all n , we have

$$\xi_1^{(n)}(\cdot) = \int_{\mathbf{X} \times \mathbf{A}} p(\cdot | x, a, \nu_1^{(n)}) \nu^{(n)}(dx, da). \tag{4.2}$$

Let $g \in C_b(\mathbf{X})$. Then, by [19, Theorem 3.5], we have

$$\lim_{n \rightarrow \infty} \int_{\mathbf{X} \times \mathbf{A}} \int_{\mathbf{X}} g(y) p(dy | x, a, \nu_1^{(n)}) \nu^{(n)}(dx, da) = \int_{\mathbf{X} \times \mathbf{A}} \int_{\mathbf{X}} g(y) p(dy | x, a, \nu_1) \nu(dx, da)$$

since $\nu^{(n)} \rightarrow \nu$ weakly and $\int_{\mathbf{X}} g(y) p(dy | x, a, \nu_1^{(n)})$ converges to $\int_{\mathbf{X}} g(y) p(dy | x, a, \nu_1)$ continuously* (see [28, p. 388]). This implies that the sequence of measures on the right-hand side of (4.2) converges weakly to $\int_{\mathbf{X} \times \mathbf{A}} p(\cdot | x, a, \nu_1) \nu(dx, da)$. Therefore, we have

$$\xi_1(\cdot) = \int_{\mathbf{X} \times \mathbf{A}} p(\cdot | x, a, \nu_1) \nu(dx, da),$$

which means that $\xi \in C(\nu)$.

We now prove that $\xi \in B(\nu)$, which implies that $\xi \in \Gamma(\nu)$. To prove this, we use the continuity properties of the average cost optimality equation. For each n , let us define

$$F^{(n)}(x, a) = c(x, a, \nu_1^{(n)}) + \int_{\mathbf{X}} h_{\nu_1^{(n)}}(y) \hat{p}(dy | x, a, \nu_1^{(n)})$$

and

$$F(x, a) = c(x, a, \nu_1) + \int_{\mathbf{X}} h_{\nu_1}(y) \hat{p}(dy | x, a, \nu_1).$$

By definition,

$$h_{\nu_1^{(n)}}(x) = \inf_{a \in \mathbf{A}} F^{(n)}(x, a) \quad \text{and} \quad h_{\nu_1}(x) = \inf_{a \in \mathbf{A}} F(x, a).$$

By assumption, we have

$$1 = \xi^{(n)} \left(\left\{ (x, a) : F^{(n)}(x, a) = h_{\nu_1^{(n)}}(x) \right\} \right), \quad \text{for all } n.$$

*Suppose g, g_n ($n \geq 1$) are measurable functions on metric space \mathbf{E} . The sequence g_n is said to converge to g continuously if $\lim_{n \rightarrow \infty} g_n(e_n) = g(e)$ for any $e_n \rightarrow e$ where $e \in \mathbf{E}$.

Let $A^{(n)} = \{(x, a) : F^{(n)}(x, a) = h_{\nu_1^{(n)}}(x)\}$. Since both $F^{(n)}$ and $h_{\nu_1^{(n)}}$ are continuous, $A^{(n)}$ is closed. Define $A = \{(x, a) : F(x, a) = h_{\nu_1}(x)\}$ which is also closed as both F and h_{ν_1} are continuous.

We first prove $\xi \in B(\nu)$ by supposing that $F^{(n)}$ converges to F continuously and $h_{\nu_1^{(n)}}$ converges to h_{ν_1} continuously, as $n \rightarrow \infty$. Then, we prove these statements.

For each $M \geq 1$, define closed set $B^M = \{(x, a) : F(x, a) \geq h_{\nu_1}(x) + \epsilon(M)\}$, where $\epsilon(M) \rightarrow 0$ as $M \rightarrow \infty$. Since both F and h_{ν_1} are continuous, we can choose $\{\epsilon(M)\}_{M \geq 1}$ so that $\xi(\partial B^M) = 0$ for each M . As $A^c = \bigcup_{M=1}^{\infty} B^M$ and $B^M \subset B^{M+1}$, we have by the monotone convergence theorem

$$\xi^{(n)}(A^c \cap A^{(n)}) = \liminf_{M \rightarrow \infty} \xi^{(n)}(B^M \cap A^{(n)}).$$

Hence, we have

$$\begin{aligned} 1 &= \limsup_{n \rightarrow \infty} \liminf_{M \rightarrow \infty} \left\{ \xi^{(n)}(A \cap A^{(n)}) + \xi^{(n)}(B^M \cap A^{(n)}) \right\} \\ &\leq \liminf_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \left\{ \xi^{(n)}(A \cap A^{(n)}) + \xi^{(n)}(B^M \cap A^{(n)}) \right\}. \end{aligned}$$

Fix any M . Note that $\xi^{(n)}$ converges weakly to ξ as $n \rightarrow \infty$ when both measures are restricted to B^M , as B^M is closed and $\xi(\partial B^M) = 0$ (see [6, Theorem 8.2.3]). Furthermore, $1_{A^{(n)} \cap B^M}$ converges continuously to 0. Indeed, if $(x^{(n)}, a^{(n)}) \rightarrow (x, a)$ in B^M , then

$$\begin{aligned} \lim_{n \rightarrow \infty} F^{(n)}(x^{(n)}, a^{(n)}) &= F(x, a) \\ &\geq h_{\nu_1}(x) + \epsilon(M) \\ &= \lim_{n \rightarrow \infty} h_{\nu_1^{(n)}}(x^{(n)}) + \epsilon(M). \end{aligned}$$

Hence, for large enough n , we have $F^{(n)}(x^{(n)}, a^{(n)}) > h_{\nu_1^{(n)}}(x^{(n)})$, which implies that $(x^{(n)}, a^{(n)}) \notin A^{(n)}$, so $1_{A^{(n)} \cap B^M}(x^{(n)}, a^{(n)}) = 0$. Therefore, by [19, Theorem 3.5], for each M , we have

$$\limsup_{n \rightarrow \infty} \xi^{(n)}(B^M \cap A^{(n)}) = 0.$$

Therefore, we obtain

$$\begin{aligned} 1 &\leq \limsup_{n \rightarrow \infty} \xi^{(n)}(A \cap A^{(n)}) \\ &\leq \limsup_{n \rightarrow \infty} \xi^{(n)}(A) \\ &\leq \xi(A), \end{aligned}$$

where the last inequality follows from the Portmanteau Theorem (see [4, Theorem 2.1]). Hence, $\xi(A) = 1$. This means that $\xi \in B(\nu)$. Therefore, $\xi \in \Gamma(\nu)$ which completes the proof under the condition that $F^{(n)}$ converges to F continuously and $h_{\nu_1^{(n)}}$ converges to h_{ν_1} continuously, as $n \rightarrow \infty$, which we prove next.

For continuous functions, continuous convergence is the same as uniform convergence on compact sets (see [19, Lemma 2.1]). Therefore, we show that $F^{(n)}$ uniformly converges to F over compact sets and $h_{\nu_1^{(n)}}$

uniformly converges to h_{ν_1} over compact sets, as these functions are continuous. Furthermore, if $h_{\nu_1^{(n)}}$ converges to h_{ν_1} continuously, then $F^{(n)}$ also converges to F continuously. Indeed, let $(x^{(n)}, a^{(n)}) \rightarrow (x, a)$. By [19, Theorem 3.5] we have

$$\begin{aligned} \lim_{n \rightarrow \infty} F^{(n)}(x^{(n)}, a^{(n)}) &= \lim_{n \rightarrow \infty} \left[c(x, a, \nu_1^{(n)}) + \int_{\mathbb{X}} h_{\nu_1^{(n)}}(y) \hat{p}(dy|x, a, \nu_1^{(n)}) \right] \\ &= \left[c(x, a, \nu_1) + \int_{\mathbb{X}} h_{\nu_1}(y) \hat{p}(dy|x, a, \nu_1) \right] \\ &= F(x, a). \end{aligned}$$

Therefore, it is enough to prove that $h_{\nu_1^{(n)}}$ uniformly converges to h_{ν_1} over compact sets. The following proposition establishes this result.

Proposition 4.4 *For any compact $K \subset \mathbb{X}$, we have*

$$\lim_{n \rightarrow \infty} \sup_{x \in K} |h_{\nu_1^{(n)}}(x) - h_{\nu_1}(x)| = 0.$$

Proof To ease the notation, let $T^{(n)} = T_{\nu_1^{(n)}}$, $h^{(n)} = h_{\nu_1^{(n)}}$, $T = T_{\nu_1}$, and $h = h_{\nu_1}$. Let $u_0^{(n)} = u_0 = 0$ and

$$u_{k+1}^{(n)} = T^{(n)}u_k^{(n)} \text{ and } u_{k+1} = Tu_k.$$

Since $T^{(n)}$ and T are contractive operators on $C_b(\mathbb{X})$ with modulus β , it can be proved that

$$\|u_k^{(n)} - h^{(n)}\|, \|u_k - h\| \leq L_0 \beta^k, \tag{4.3}$$

for some constant L_0 .

For any compact $K \subset \mathbb{X}$, one can prove that

$$\lim_{n \rightarrow \infty} \sup_{x \in K} |u_k^{(n)}(x) - u_k(x)| = 0, \tag{4.4}$$

for all $k \geq 0$. Indeed, fix any compact $K \subset \mathbb{X}$. Since $u_0^{(n)} = u_0 = 0$, the claim trivially holds for $k = 0$. Suppose that the claim holds for k . Then, consider $k + 1$:

$$\begin{aligned} &\sup_{x \in K} |u_{k+1}^{(n)}(x) - u_{k+1}(x)| \\ &= \sup_{x \in K} \left| \min_{a \in \mathbb{A}} \left[c(x, a, \nu_1^{(n)}) + \int_{\mathbb{X}} u_k^{(n)}(y) \hat{p}(dy|x, a, \nu_1^{(n)}) \right] \right. \\ &\quad \left. - \min_{a \in \mathbb{A}} \left[c(x, a, \nu_1) + \beta \int_{\mathbb{X}} u_k(y) \hat{p}(dy|x, a, \nu_1) \right] \right| \\ &\leq \sup_{(x,a) \in K \times \mathbb{A}} |c(x, a, \nu_1^{(n)}) - c(x, a, \nu_1)| \\ &\quad + \sup_{(x,a) \in K \times \mathbb{A}} \left| \int_{\mathbb{X}} u_k^{(n)}(y) \hat{p}(dy|x, a, \nu_1^{(n)}) \right. \end{aligned}$$

$$- \int_{\mathbf{X}} u_k(y) \hat{p}(dy|x, a, \nu_1) \Big|.$$

Note that $c(\cdot, \cdot, \nu^{(n)})$ converges to $c(\cdot, \cdot, \nu)$ continuously as $n \rightarrow \infty$. Furthermore, since $u_k^{(n)}$ converges to u_k continuously, by [19, Theorem 3.5] $\int_{\mathbf{X}} u_k^{(n)}(y) \hat{p}(dy|\cdot, \cdot, \nu_1^{(n)})$ converges to $\int_{\mathbf{X}} u_k(y) \hat{p}(dy|\cdot, \cdot, \nu_1)$ continuously as $n \rightarrow \infty$. Since continuous convergence is equivalent to uniform convergence over compact sets for continuous functions, the last expression goes to zero as $n \rightarrow \infty$. This proves (4.4).

Using (4.4), we now complete the proof of Proposition 4.4. Fix any compact $K \subset \mathbf{X}$. For all $k \geq 0$, we have

$$\begin{aligned} \sup_{x \in K} |h^{(n)}(x) - h(x)| &\leq \|h^{(n)} - u_k^{(n)}\| + \sup_{x \in K} |u_k^{(n)}(x) - u_k(x)| + \|u_k - h\| \\ &\leq 2L_0\beta^k + \sup_{x \in K} |u_k^{(n)}(x) - u_k(x)| \quad (\text{by (4.3)}). \end{aligned}$$

This last expression can be made arbitrarily small by first choosing large enough k and then large enough n . \square

This completes the proof of Proposition 4.3 as $F^{(n)}$ converges to F continuously and $h_{\nu_1^{(n)}}$ converges to h_{ν_1} continuously by Proposition 4.4. \square

Recall that Ξ is a compact convex subset of the locally convex topological space $\mathcal{M}(\mathbf{X} \times \mathbf{A})$. Furthermore, Γ has closed graph by Proposition 4.3, and it takes nonempty convex values. Therefore, by Kakutani’s fixed point theorem (see [2, Corollary 17.55]), Γ has a fixed point. Then, Theorem 2.2 follows from Proposition 4.2.

5. Proof of Theorem 3.3

To prove Theorem 3.3, we mimic the idea used in [31] of formulating N -agent game as a stochastic control problem over an extended state space and action space. Namely, we consider N -agent game as a stochastic control problem with state space $\mathbf{X}^N = \prod_{i=1}^N \mathbf{X}$, action space $\mathbf{A}^N = \prod_{i=1}^N \mathbf{A}$, transition probability

$$P^N(dy_1, \dots, dy_N|x_1, \dots, x_N, a_1, \dots, a_N) = \prod_{i=1}^N p(dy_i|x_i, a_i),$$

and one-stage cost function

$$C^N(x_1, \dots, x_N, a_1, \dots, a_N) = c \left(x_1, a_1, \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right).$$

Here, the initial state $(x_1(0), \dots, x_N(0))$ is generated by the product measure $\mu_0^N = \prod_{i=1}^N \mu_0$. For any N -tuple of policies $\pi^{(N)}$, the average-cost of this stochastic control problem is the same as the average-cost of Agent 1 in the game; that is,

$$J_1^N(\pi^{(N)}, \mu_0) = \limsup_{T \rightarrow \infty} \frac{1}{T} E^{\pi^{(N)}, \mu_0} \left[\sum_{t=0}^{T-1} C^N(x_1^N(t), \dots, x_N^N(t), a_1^N(t), \dots, a_N^N(t)) \right].$$

Suppose that Assumptions 1 and 2 hold. Then, for all N , the following are satisfied:

- (i) The one-stage cost function C^N is bounded and continuous.
- (ii) The stochastic kernel P^N is weakly continuous and satisfies

$$P^N(\cdot | x_1, \dots, x_N, a_1, \dots, a_N) \geq \lambda^N(\cdot) = \prod_{i=1}^N \lambda(\cdot), \quad \forall (x_1, \dots, x_N) \in \mathbb{X}^N, (a_1, \dots, a_N) \in \mathbb{A}^N \tag{5.1}$$

$$\sup_{(a_1, \dots, a_N)} \int_{\mathbb{X}^N} w^N(y_1, \dots, y_N) \hat{p}^N(dy_1, \dots, y_N | x_1, \dots, x_N, a_1, \dots, a_N) \leq \alpha^N w^N(x_1, \dots, x_N), \quad \forall (x_1, \dots, x_N) \in \mathbb{X}^N \tag{5.2}$$

where $w^N(x_1, \dots, x_N) = \prod_{i=1}^N w(y_i)$ and $\hat{p}^N(\cdot | x_1, \dots, x_N, a_1, \dots, a_N) = p(\cdot | x_1, \dots, x_N, a_1, \dots, a_N) - \lambda^N(\cdot)$. Under (5.1) and (5.2), for any $\pi^{(N)}$, the following stochastic kernel on \mathbb{X}^N given \mathbb{X}^N

$$P^{\pi^{(N)}}(\cdot | x_1, \dots, x_N) = \int_{\mathbb{A}^N} P^N(\cdot | x_1, \dots, x_N, a_1, \dots, a_N) \prod_{i=1}^N \pi(da_i | x_i)$$

has an unique invariant distribution $\mu_{\pi^{(N)}} \in \mathcal{P}(\mathbb{X}^N)$ and we have

$$J_1^N(\pi^{(N)}, \mu_0) = \int_{\mathbb{X}^N \times \mathbb{A}^N} c\left(x_1, a_1, \frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right) \prod_{i=1}^N \pi_i(da_i | x_i) \mu_{\pi^{(N)}}(dx_1, \dots, dx_N).$$

(see [30, Theorem 3.3], [12, Lemma 3.4]). Moreover,

$$\mu_{\pi^{(N)}}(dx_1, \dots, dx_N) = \prod_{i=1}^N \mu_{\pi_i}(dx_i),$$

where, for each $i = 1, \dots, N$, μ_{π_i} is the unique invariant probability measure of the stochastic kernel

$$\int_{\mathbb{A}} p(\cdot | x, a) \pi_i(da | x).$$

The following result is the key to prove Theorem 3.3.

Proposition 5.1 *For any $\pi \in \Pi_1$, we have*

$$|J_{\mu^*}(\pi, \mu_0) - J_1^N(\pi, \pi_{-1}^{(N*)}, \mu_0)| \leq \epsilon(N),$$

where

$$\epsilon(N) = \int_{\mathbb{X}^{N-1} \times \mathbb{A}^{N-1}} \omega_c\left(\rho\left(\mu^*, \frac{1}{N-1} \sum_{i=1}^N \delta_{x_i}\right)\right) \mu^*(dx_2) \dots \mu^*(dx_N) + \omega_c\left(2 \cdot \left(1 - \frac{N-1}{N} + \frac{1}{N}\right)\right).$$

Note that, by law of large numbers and the fact that $\omega_c(r) \rightarrow 0$ as $r \rightarrow 0$, $\epsilon(N) \rightarrow 0$ as $N \rightarrow \infty$.

Proof Fix any policy π for Agent 1. Since μ_π is the unique invariant distribution of the stochastic kernel $\int_{\mathbf{A}} p(\cdot | x, a) \pi(da|x)$, we can write

$$J_{\mu^*}(\pi, \mu_0) = \int_{\mathbf{X} \times \mathbf{A}} c(x_1, a_1, \mu^*) \pi(da_1|x_1) \mu_\pi(dx_1)$$

(see [30, Theorem 3.3], [12, Lemma 3.4]). Then, we have

$$\begin{aligned} & |J_{\mu^*}(\pi, \mu_0) - J_1^N(\pi, \boldsymbol{\pi}_{-1}^{(N*)}, \mu_0)| \\ &= \left| \int_{\mathbf{X} \times \mathbf{A}} c(x_1, a_1, \mu^*) \pi(da_1|x_1) \mu_\pi(dx_1) \right. \\ &\quad \left. - \int_{\mathbf{X}^N \times \mathbf{A}^N} c\left(x_1, a_1, \frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right) \pi(da_1|x_1) \prod_{i=2}^N \pi^*(da_i|x_i) \mu_{\pi, \boldsymbol{\pi}_{-1}^{(N*)}}(dx_1, \dots, dx_N) \right| \\ &\stackrel{(a)}{=} \left| \int_{\mathbf{X} \times \mathbf{A}} c(x_1, a_1, \mu^*) \pi(da_1|x_1) \mu_\pi(dx_1) \right. \\ &\quad \left. - \int_{\mathbf{X} \times \mathbf{A}} \left[\int_{\mathbf{X}^{N-1}} c\left(x_1, a_1, \frac{\delta_{x_1}}{N} + \frac{N-1}{N} \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}\right) \mu^*(dx_2) \dots \mu^*(dx_N) \right] \pi(da_1|x_1) \mu_\pi(dx_1) \right| \\ &\leq \left| \int_{\mathbf{X} \times \mathbf{A}} c(x_1, a_1, \mu^*) \pi(da_1|x_1) \mu_\pi(dx_1) \right. \\ &\quad \left. - \int_{\mathbf{X} \times \mathbf{A}} \left[\int_{\mathbf{X}^{N-1}} c\left(x_1, a_1, \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}\right) \mu^*(dx_2) \dots \mu^*(dx_N) \right] \pi(da_1|x_1) \mu_\pi(dx_1) \right| \\ &\quad + \left| \int_{\mathbf{X} \times \mathbf{A}} \left[\int_{\mathbf{X}^{N-1}} c\left(x_1, a_1, \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}\right) \mu^*(dx_2) \dots \mu^*(dx_N) \right] \pi(da_1|x_1) \mu_\pi(dx_1) \right. \\ &\quad \left. - \int_{\mathbf{X} \times \mathbf{A}} \left[\int_{\mathbf{X}^{N-1}} c\left(x_1, a_1, \frac{\delta_{x_1}}{N} + \frac{N-1}{N} \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}\right) \mu^*(dx_2) \dots \mu^*(dx_N) \right] \pi(da_1|x_1) \mu_\pi(dx_1) \right| \\ &\stackrel{(b)}{\leq} \int_{\mathbf{X}^{N-1}} \omega_c\left(\rho(\mu^*, \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i})\right) \mu^*(dx_2) \dots \mu^*(dx_N) + \omega_c\left(2 \cdot \left(1 - \frac{N-1}{N} + \frac{1}{N}\right)\right), \end{aligned}$$

where (a) follows from the fact that μ^* is the unique invariant probability measure of the stochastic kernel $\int_{\mathbf{A}} p(\cdot | x, a) \pi^*(da|x)$ and μ_π is the unique invariant probability measure of the stochastic kernel $\int_{\mathbf{A}} p(\cdot | x, a) \pi(da|x)$, and (b) follows from

$$\rho\left(\frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}, \frac{\delta_{x_1}}{N} + \frac{N-1}{N} \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}\right) = 2 \cdot \left(1 - \frac{N-1}{N} + \frac{1}{N}\right).$$

This completes the proof. □

Now, we are ready to prove Theorem 3.3. We prove that, for sufficiently large N , we have

$$J_i^N(\boldsymbol{\pi}^{(N)}, \mu_0) \leq \inf_{\pi \in \Pi_i} J_i^N(\boldsymbol{\pi}_{-i}^{(N)}, \pi, \mu_0) + \varepsilon \tag{5.3}$$

for each $i = 1, \dots, N$. Since the transition probabilities and the one-stage cost functions are the same for all agents in the game, it is sufficient to prove (5.3) for Agent 1 only. Given $\epsilon > 0$, for each $N \geq 1$, let $\tilde{\pi}^{(N)} \in \Pi_1$ be such that

$$J_1^N(\tilde{\pi}^{(N)}, \pi^*, \dots, \pi^*, \mu_0) < \inf_{\pi \in \Pi_1} J_1^N(\pi, \pi^*, \dots, \pi^*, \mu_0) + \frac{\epsilon}{3}.$$

Let $N(\epsilon)$ be such that $\epsilon(N) \leq \frac{\epsilon}{3}$ for all $N \geq N(\epsilon)$. Then, by Proposition 5.1, we have

$$\begin{aligned} \inf_{\pi \in \Pi_1} J_1^N(\pi, \pi^*, \dots, \pi^*, \mu_0) + \epsilon &> J_1^N(\tilde{\pi}^{(N)}, \pi^*, \dots, \pi^*, \mu_0) + \frac{2\epsilon}{3} \\ &\geq J_{\mu^*}(\tilde{\pi}^{(N)}, \mu_0) + \frac{\epsilon}{3} \\ &\geq \inf_{\pi \in \Pi} J_{\mu^*}(\pi, \mu_0) + \frac{\epsilon}{3} \\ &= J_{\mu^*}(\pi^*, \mu_0) + \frac{\epsilon}{3} \\ &\geq J_1^N(\pi^*, \pi^*, \dots, \pi^*, \mu_0). \end{aligned}$$

This completes the proof of Theorem 3.3.

6. Conclusion

This paper has considered discrete-time mean-field games subject to average cost criteria, for Polish state and action spaces. Under drift and minorization conditions on the state transition kernel, the existence of a mean-field equilibrium has been established for infinite-population game models using the average cost optimality equation. We have then applied the policy in mean-field equilibrium to the finite population game and have proved that it constitutes an approximate Nash equilibrium for games with a sufficiently large number of agents.

One interesting future direction of research to pursue is to study partially observed version of the same problem. By converting partially-observed system into a fully-observed one over a belief-state space, it might be possible to establish similar results for the partially-observed case.

Acknowledgment

This research was supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) BİDEB 2232 Research Grant. The author is grateful to Professor Tamer Başar and Professor Maxim Raginsky for their constructive comments.

References

- [1] Adlakha S, Johari R, Weintraub GY. Equilibria of dynamic games with many players: Existence, approximation, and market structure. *Journal of Economic Theory* 2015; 156: 269-316.
- [2] Aliprantis CD, Border KC. *Infinite Dimensional Analysis*. 3rd ed. Berlin, Germany: Springer, 2006.
- [3] Bensoussan A, Frehse J, Yam P. *Mean Field Games and Mean Field Type Control Theory*. Springer, New York, NY, USA, 2013.
- [4] Billingsley P. *Convergence of Probability Measures*. 2nd ed. New York, NY, USA: Wiley, 1999.
- [5] Biswas A. Mean field games with ergodic cost for discrete time Markov processes; 2015. arXiv:1510.08968.

- [6] Bogachev VI. Measure Theory: Volume II. Springer; 2007.
- [7] Cardaliaguet P. Notes on mean-field games, 2011.
- [8] Carmona R, Delarue F. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization* 2013; 51 (4): 2705-2734.
- [9] Elliot R, Li X, Ni Y. Discrete time mean-field stochastic linear-quadratic optimal control problems. *Automatica* 2013; 49: 3222-3233.
- [10] Gomes DA, Mohr J, Souza RR. Discrete time, finite state space mean field games. *Journal de Mathematiques Pures et Appliquees* 2010; 93: 308-328.
- [11] Gomes DA, Saúde J. Mean field games models - a brief survey. *Dynamic Games and Applications* 2014; 4 (2): 110-154.
- [12] Gordienko E, Hernandez-Lerma O. Average Cost Markov Control Processes with Weighted Norms: Existence of Canonical Policies. *Applicationes Mathematicae* 1995; 23 (2): 199-218.
- [13] Hernández-Lerma O, Lasserre JB. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. New York, NY, USA: Springer, 1996.
- [14] Hernández-Lerma O, Lasserre JB. *Further Topics on Discrete-Time Markov Control Processes*. Springer; 1999.
- [15] Hernández-Lerma O, Montes-De-Oca R, Cavazos-Cadena R. Recurrence conditions for Markov decision processes with Borel state space: a survey. *Annals of Operations Research* 1991; 28 (1): 29-46.
- [16] Huang M. Large-population LQG games involving major player: The Nash certainty equivalence principle. *SIAM Journal on Control Optimization* 2010; 48 (5): 3318-3353.
- [17] Huang M, Caines PE, Malhamé RP. Large-population cost coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Transactions on Automatic Control* 2007; 52 (9): 1560-1571.
- [18] Huang M, Malhamé RP, Caines PE. Large population stochastic dynamic games: Closed loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information Systems* 2006; 6: 221-252.
- [19] Langen HJ. Convergence of dynamic programming models. *Mathematics of Operations Research* 1981; 6 (4): 493-512.
- [20] Lasry J, Lions P. Mean field games. *Japanese Journal of Mathematics* 2007; 2: 229-260.
- [21] Moon J, Başar T. Discrete-time decentralized control using the risk-sensitive performance criterion in the large population regime: a mean field approach. In: *ACC 2015*. Chicago; 2015.
- [22] Moon J, Başar T. Discrete-time mean field Stackelberg games with a large number of followers. In: *CDC 2016*. Las Vegas, USA; 2016.
- [23] Moon J, Başar T. Robust mean field games for coupled Markov jump linear systems. *International Journal of Control* 2016; 89 (7): 1367-1381.
- [24] Nourian M, Nair GN. Linear-quadratic-Gaussian mean field games under high rate quantization. In: *CDC 2013*. Florence, Italy; 2013.
- [25] Parthasarathy KR. *Probability Measures on Metric Spaces*. AMS Bookstore; 1967.
- [26] Saldi N, Başar T, Raginsky M. Markov-Nash Equilibria in Mean-Field Games with Discounted Cost. *SIAM Journal on Control and Optimization* 2018 ;56 (6): 4256-4287.
- [27] Saldi N, Linder T, Yüksel S. *Finite approximations in discrete-time stochastic control: Quantized models and asymptotic optimality*. Cham, Switzerland: Springer, 2018.
- [28] Serfozo R. Convergence of Lebesgue integrals with varying measures. *Sankhya SerA* 1982; p. 380-402.
- [29] Tembine H, Zhu Q, Başar T. Risk-sensitive mean field games. *IEEE Transactions on Automatic Control* 2014; 59 (4): 835-850.

- [30] Vega-Amaya O. The average cost optimality equation: a fixed point approach. *Boletin de la Sociedad Matematica Mexicana* 2003; 9 (3): 185-195.
- [31] Wiecek P. Discrete-time ergodic mean-field games with average reward on compact spaces. *Dynamic Games and Applications* 2019; p. 1-35.
- [32] Wiecek P, Altman E. Stationary anonymous sequential games with undiscounted rewards. *Journal of Optimization Theory and Applications* 2015; 166 (2): 686-710.