Dynamic Games and Applications
https://doi.org/10.1007/s13235-024-00570-x

Multiple-Population Discrete-Time Mean Field Games with Discounted and Total Payoffs: Approximation of Games with Finite Populations

Piotr Więcek[1]

Accepted: 6 June 2023

Abstract
In the paper we present a model of discrete-time mean-field game with multiple populations of players. Its main result shows that the equilibria obtained for the mean-field limit are approximate Markov–Nash equilibria for n-person counterparts of these mean-field games when the number of players in each population is large enough. We consider two payoff criteria: β-discounted payoff and total payoff. The existence of mean-field equilibria for games with both payoffs has been proven in our previous article, hence, the theorems presented here show in fact the existence of approximate equilibria in certain classes of stochastic games with large finite numbers of players. The results are provided under some rather general assumptions on one-step reward functions and individual transition kernels of the players. In addition, the results for total payoff case, when applied to a single population, extend the theory of mean-field games also by relaxing some strong assumptions used in the existing literature.

Keywords Mean-field game · Discrete time · Multiple-population game · Stationary mean-field equilibrium · Markov mean-field equilibrium · Discounted payoff · Total payoff

Mathematics Subject Classification 91A15 · 91A13 · 91A10

1 Introduction

The paper is the continuation of article [18], where we have presented a model of discrete-time mean-field game with multiple populations of players. In the paper we have provided the results about the conditions guaranteing existence of Markov or stationary equilibria in such games for two payoff criteria: β-discounted payoff and total payoff. These theorems were the first to deal with the problem of the existence of equilibrium in mean-field games with several populations in the discrete time setting, extending also some of the theory for single population discrete-time mean-field games. As mean-field games are only meant to

✉ Piotr Więcek
Piotr.Wiecek@pwr.edu.pl

1 Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Published online: 15 June 2024

serve as an approximation of real-life situations, where the populations of agents are large, but finite, it is crucial to complete the equilibrium-existence results for the class of games under investigation by the results showing that the equilibria obtained for the mean-field limit are approximate Markov–Nash equilibria for n-person counterparts of these mean-field games when the number of players in each population is large enough. This type of theorems have been provided for the single population case for different variants of the discrete-time mean-field game model in [10–14, 16, 17]. In this article, we build upon the theory presented in [11] to show that also in our case the Markov (or stationary) mean-field equilibria obtained in [18] are approximate equilibria in finite-player counterparts of the mean-field game when the number of players in each population goes to infinity.

The organization of the paper is as follows: In Sect. 2 we present the model of the discrete-time mean-field games with several populations of the players and its counterparts with finite number of players. In Sect. 3 we introduce some notation used in the remainder of the article. In Sect. 4 we give all the assumptions used in our theorems. Sections 5 contains all the main results of the article. We end with some concluding remarks in Sect. 6.

2 The Model

2.1 Multi-population Mean-Field Game Model

A *multi-population discrete-time mean-field game* is described by the following objects:

- We assume that the game is played in discrete time, that is $t \in \{1, 2, \ldots\}$.
- The game is played by an infinite number (continuum) of players divided into N *populations*. Each player has a *private state* s, changing over time. We assume that the set of individual states S^i is the same for each player in population i ($i = 1, \ldots, N$), and that it is a nonempty closed subset of a locally compact Polish space S.[1]
- A vector $\overline{\mu} = (\mu^1, \ldots, \mu^N) \in \Pi_{i=1}^N \Delta(S^i)$ of N probability distributions over Borel sets[2] of S^i, $i = 1, \ldots, N$, is called a *global state* of the game. Its i-th component describes the proportion of i-th population, which is in each of the individual states.
  We assume that at every stage of the game each player knows both his private state and the global state, and that his knowledge about individual states of his opponents is limited to the global state.
- The set of *actions* available to a player from population i in state $(s, \overline{\mu})$ is given by $A^i(s)$, with $A := \bigcup_{i\in\{1,\ldots,N\}, s\in S^i} A^i(s)$ – a compact metric space. For any i, $A^i(\cdot)$ is a non-empty compact valued correspondence such that

$$D^i := \{(s, a) \in S^i \times A : a \in A^i(s)\}$$

  is a measurable set. Note that we assume that the sets of actions available to a player only depend on his private state and not on the global state of the game.
- The global distribution of the state-action pairs is denoted by $\overline{\tau} = (\tau^1, \ldots, \tau^N) \in \Pi_{i=1}^N \Delta(D^i)$. Again, it gives the distributions of state-action pairs within the population divided into subpopulations $i = 1, \ldots, N$.

[1] As it can be clearly seen, the model encompasses in particular the situation when the state space for each population is the same and equal to S.

[2] Here and in the sequel, for any set X, $\Delta(X)$ denotes the set of probability distributions over the σ-algebra of Borel subsets of X, $\mathcal{B}(X)$.

- *Immediate reward* of an individual from population i is given by a measurable function $r^i : D^i \times \Pi_{i=1}^N \Delta(D^i) \to \mathbb{R}$. $r^i(s, a, \overline{\tau})$ gives the reward of a player at any stage of the game when his private state is s, his action is a and the distribution of state-action pairs among the entire player population is $\overline{\tau}$.
- *Transitions* are defined for each individual separately with stochastic kernels $Q^i : D^i \times \Pi_{i=1}^N \Delta(D^i) \to \Delta(S^i)$ denoting transition probability for players from i-th population. $Q^i(B \mid \cdot, \cdot, \overline{\tau})$ is product-measurable for any $B \in \mathcal{B}(S^i)$, any $\overline{\tau} \in \Pi_{i=1}^N \Delta(D^i)$ and $i \in \{1, \ldots, N\}$.
- The global state at time $t+1$, $\overline{\mu_t}$, is given by the aggregation of individual transitions of the players done by the formula

$$\mu_{t+1}^i(\cdot) = \Phi^i(\cdot \mid \overline{\tau_t}) := \int_{D^i} Q^i(\cdot \mid s, a, \overline{\tau_t})\tau_t^i(ds \times da).$$

As it can be clearly seen, the transition of the global state is deterministic.

A sequence $\pi^i = \{\pi_t^i\}_{t=0}^\infty$ of functions $\pi_t^i : S^i \to \Delta(A)$, such that $\pi_t^i(B \mid \cdot)$ is measurable for any $B \in \mathcal{B}(A)$ and any t, satisfying $\pi_t^i(A^i(s) \mid s) = 1$ for every $s \in S^i$ and every t, is called a *Markov strategy* for a player of population i. A function $f^i : S^i \to \Delta(A)$, such that $f^i(B \mid \cdot)$ is measurable for any $B \in \mathcal{B}(A)$, satisfying $f^i(A^i(s) \mid s) = 1$ for every $s \in S^i$ is called a *stationary strategy*. The set of all Markov strategies for players from i-th population is denoted by $\mathcal{M}^i$ while that of stationary strategies by $\mathcal{F}^i$. As in MDPs, stationary strategies can be seen as a specific case of Markov strategies that do not depend on t. In the paper we never consider general (history-dependent) strategies.

Next, let $\Pi_t^i(\pi^i, \mu^i)$ denote the state-action distribution of the i-th population players at time t in the mean-field game corresponding to the distribution of individual states in population i, μ^i and a Markov strategy for players of population i, $\pi^i \in \mathcal{M}^i$, that is

$$\Pi_t^i(\pi^i, \mu^i)(B) := \int_B \pi_t^i(da \mid s)\mu^i(ds) \quad \text{for } B \in \mathcal{B}(D^i).$$

The vector $(\Pi_t^1(\pi^1, \mu^1), \ldots, \Pi_t^N(\pi^N, \mu^N))$ will be denoted by $\overline{\Pi}_t(\overline{\pi}, \overline{\mu})$. When we use this notation for stationary strategies, we skip the subscript t.

Given the evolution of the global state, which depends on the strategies of the players in a deterministic manner, we can define the individual history of a player α (from any given population i) as the sequence of his consecutive individual states and actions $h = (s_0^\alpha, a_0^\alpha, s_1^\alpha, a_1^\alpha, \ldots)$. By the Ionescu-Tulcea theorem (see Chap. 7 in [2]), for any Markov strategies π^α of player α and $\sigma^1, \ldots, \sigma^N$ of other players (including all other players of the same population), any initial global state $\overline{\mu_0}$ and any initial private state of player α, s, there exists a unique probability measure $\mathbb{P}^{s,\overline{\mu_0},\overline{Q},\pi^\alpha,\overline{\sigma}}$ on the set of all infinite individual histories of the game $H = (D^i)^\infty$ endowed with Borel σ-algebra, such that for any $B \in \mathcal{B}(S^i)$, $E \in \mathcal{B}(A)$ and any partial history $h_t^\alpha = (s_0^\alpha, a_0^\alpha, \ldots, s_{t-1}^\alpha, a_{t-1}^\alpha, s_t^\alpha) \in (D^i)^t \times S^i =: H_t$, $t \in \mathbb{N}$,

$$\mathbb{P}^{s,\overline{\mu_0},\overline{Q},\pi^\alpha,\overline{\sigma}}(h \in H : s_0^\alpha = s) = 1, \tag{1}$$

$$\mathbb{P}^{s,\overline{\mu_0},\overline{Q},\pi^\alpha,\overline{\sigma}}(h \in H : a_t^\alpha \in E \mid h_t^\alpha) = \pi_t^\alpha(E \mid s_t^\alpha),$$

$$\mathbb{P}^{s,\overline{\mu_0},\overline{Q},\pi^\alpha,\overline{\sigma}}(h \in H : s_{t+1}^\alpha \in B \mid (h_t^\alpha, a_t^\alpha)) = Q^i(B \mid s_t^\alpha, a_t^\alpha, \overline{\tau_t}), \tag{2}$$

with state-action distributions defined by $\tau_0^j = \Pi_0^j(\sigma^j, \mu_0^j)$, $\tau_{t+1}^j = \Pi_t^i(\sigma^j, \Phi^j(\cdot \mid \overline{\tau_t}))$ for $t = 1, 2, \ldots$ and $j = 1, \ldots, N$.

For $\beta \in (0, 1)$, the *β-discounted reward*[3] for a player α from population i using policy $\pi^i \in \mathcal{M}^i$ when other players use policies $\sigma^j \in \mathcal{M}^j$ (depending on the population j they belong to) and the initial global state is $\overline{\mu_0}$, with the initial individual state of player α being s_0^i is defined as follows:

$$J_\beta^i(s_0^i, \overline{\mu_0}, \pi^i, \overline{\sigma}) = \mathbb{E}^{s_0^i, \overline{\mu_0}, \overline{Q}, \pi^i, \overline{\sigma}} \sum_{t=0}^{\infty} \beta^t r^i(s_t^i, a_t^i, \overline{\tau_t}),$$

where $\tau_0^j = \Pi_0^j(\sigma^j, \mu_0^j)$, $\tau_{t+1}^j = \Pi_t^j(\sigma^j, \Phi^j(\cdot \mid \overline{\tau_t}))$ for $t = 1, 2, \ldots$ and $j = 1, \ldots, N$.

To define the total reward in our game let us distinguish one state in S, say s^*, isolated from $S \setminus \{s^*\}$ and assume that $A^i(s^*) = \{a^*\}$ independently of $i \in \{1, \ldots, N\}$ for some fixed a^* isolated from $A \setminus \{a^*\}$. Moreover, let us assume that $s^* \in S^i$ for $i = 1, \ldots, N$. Then the *total reward* of a player from population i using policy $\pi^i \in \mathcal{M}^i$ when other players apply policies $\overline{\sigma} = (\sigma^1, \ldots, \sigma^N)$ and the initial global state is $\overline{\mu_0}$, with the initial individual state of player α being s_0^i, is defined in the following way:

$$J_*^i(s_0^i, \overline{\mu_0}, \pi^i, \overline{\sigma}) = \mathbb{E}^{s_0^i, \overline{\mu_0}, \overline{Q}, \pi^i, \overline{\sigma}} \sum_{t=0}^{\mathcal{T}^i - 1} r^i(s_t^i, a_t^i, \overline{\tau_t}),$$

where $\tau_0^j = \Pi_0^j(\sigma^j, \mu_0^j)$, $\tau_{t+1}^j = \Pi_t^j(\sigma^j, \Phi^j(\cdot \mid \overline{\tau_t}))$ for $t = 1, 2, \ldots$ and $j = 1, \ldots, N$, while $\mathcal{T}^i$ is the moment of the first arrival of the process $\{s_t^i\}$ to s^*. The total reward is interpreted as the reward accumulated by the player over the whole of his lifetime. State s^* is an artificial state (so is action a^*) denoting that a player is dead. $\overline{\mu_0}$ corresponds to the distribution of the states across the population when he is born, while s_0^i is his own state when he is born. The fact that after some time the state of a player can become again different from s^* should be interpreted as that after some time the player is replaced by some new-born one.

Finally we define the solutions we will be looking for:

Definition 1 Stationary strategies $f^1 \in \mathcal{F}^1, \ldots f^N \in \mathcal{F}^N$ and a global state $\overline{\mu} \in \Pi_{i=1}^N \Delta(S^i)$ form a *stationary mean-field equilibrium* in the β-discounted reward game if for any $i, s_0^i \in S^i$, and every other stationary strategy of a player from population i, $g^i \in \mathcal{F}^i$

$$J_\beta^i(s_0^i, \overline{\mu}, f^i, \overline{f}) \geq J_\beta^i(s_0^i, \overline{\mu}, g^i, \overline{f})$$

and if $\overline{\mu}_0 = \overline{\mu}$, then $\overline{\mu}_t = \overline{\mu}$ for every $t \geq 1$ if strategies $f^1, \ldots, f^N$ are used by all the players.

Markov strategies $\pi^1 \in \mathcal{M}^1, \ldots \pi^N \in \mathcal{M}^N$ and a global state flow $(\overline{\mu}_0^*, \overline{\mu}_1^*, \ldots) \in (\Pi_{i=1}^N \Delta(S^i))^\infty$ form a *Markov mean-field equilibrium* in the β-discounted reward game if for any $i, s_0^i \in S^i$, and every other Markov strategy of a player from population i, $\sigma^i \in \mathcal{M}^i$

$$J_\beta^i(s_0^i, \overline{\mu_0}, \pi^i, \overline{\pi}) \geq J_\beta^i(s_0^i, \overline{\mu_0}, \sigma^i, \overline{\pi})$$

and if $\overline{\mu}_0 = \overline{\mu}_0^*$ implies $\overline{\mu}_t = \overline{\mu}_t^*$ for every $t \geq 1$ if strategies $\pi^1, \ldots, \pi^N$ are used by all the players.

Similarly,

[3] Here we replace the superscript α used to define the measure $\mathbb{P}^{s, \overline{\mu_0}, \overline{Q}, \pi^\alpha, \overline{\sigma}}$ by i, as the situation is symmetric within the population.

Definition 2 Stationary strategies $f^1 \in \mathcal{F}^1, \dots f^N \in \mathcal{F}^N$ and a global state $\overline{\mu} \in \Pi_{i=1}^N \Delta(S^i)$ form a *stationary mean-field equilibrium* in the total reward game if for any i, $s_i^0 \in S^i$, and every other stationary strategy of a player from population i, $g^i \in \mathcal{F}^i$

$$J_*^i(s_0^i, \overline{\mu}, f^i, \overline{f}) \geq J_*^i(s_0^i, \overline{\mu}, g^i, \overline{f}).$$

Moreover, if $\overline{\mu}_0 = \overline{\mu}$, then $\overline{\mu}_t = \overline{\mu}$ for every $t \geq 1$ if strategies $f^1, \dots, f^N$ are used by all the players.

Markov strategies $\pi^1 \in \mathcal{M}^1, \dots \pi^N \in \mathcal{M}^N$ and a global state flow $(\overline{\mu}_0^*, \overline{\mu}_1^*, \dots) \in (\Pi_{i=1}^N \Delta(S^i))^\infty$ form a *Markov mean-field equilibrium* in the total reward game if for any i, t, $s_i^t \in S^i$ and every other Markov strategy of a player from population i, $\sigma^i \in \mathcal{M}^i$,

$$J_*^i(s_t^i, \overline{\mu}_t^*, {}^t\pi^i, {}^t\overline{\pi}) \geq J_*^i(s_t^i, \overline{\mu}_t^*, {}^t\sigma^i, {}^t\overline{\pi}),$$

with ${}^t a$ denoting for any infinite vector $a = (a_0, a_1, \dots)$, the vector $(a_t, a_{t+1}, \dots)$. Moreover, if $\overline{\mu}_0 = \overline{\mu}_0^*$, then $\overline{\mu}_t = \overline{\mu}_t^*$ for every $t \geq 1$ if strategies $\pi^1, \dots, \pi^N$ are used by all the players.

2.2 *n*-Person Counterparts of a Mean-Field Game

The n-person games that will be approximated by our model are discrete-time n-person stochastic games as defined in [6]. Below we define *n-person stochastic counterparts* of the mean-field game for the multi-population case.

- There are n players in the game belonging to N populations. The number of players in population i is denoted by n_i, with $\sum_{i=1}^N n_i = n$. Hence, the state space is $\Pi_{i=1}^N (S^i)^{n_i}$ while an arbitrary state in the game can be denoted by $\overline{s} = (\overline{s}^1, \dots, \overline{s}^N)$ with $\overline{s}^i = (s_1^i, \dots, s_{n_i}^i)$ for $i = 1, \dots, N$. We shall also use the notation $\mathbf{n} := (n_1, \dots, n_N)$ with $\mathbf{n} \to \infty$ standing for $n_i \to \infty$ for $i = 1, \dots, N$. Similarly as in the case of the mean-field game, the set of actions available to the kth player in population i in state $\overline{s}$ is given by $A^i\left(s_k^i\right)$. An arbitrary action of the kth player in population i will be denoted by a_k^i and an arbitrary profile of actions of all the players by $\overline{a} = (\overline{a}^1, \dots, \overline{a}^N)$ with $\overline{a}^i = (a_1^i, \dots, a_{n_i}^i)$ for $i = 1, \dots, N$.
- We assume that for each i the initial values in the vector $\overline{s}^i$ are i.i.d. vectors coming from an arbitrary known distribution μ_0^{*i}. To simplify the notation we will write that the vector of initial distributions of states for each player is $\overline{\mu}_0^*$ or simply that $\overline{s_0} \sim \overline{\mu}_0^*$ to denote that.
- Empirical state-action distribution in the game is defined as

$$\overline{\tau}(\overline{s}, \overline{a}) = (\tau^1(\overline{s}, \overline{a}), \dots, \tau^N(\overline{s}, \overline{a}))$$

  with $\tau^i(\overline{s}, \overline{a}) = \frac{1}{n_i} \sum_{k=1}^{n_i} \delta_{(s_k^i, a_k^i)}$, $i = 1, \dots, N$.
- Individual immediate reward of kth player from population i, $r_n^{i,k} : \Pi_{j=1}^N (D^i)^{n_i} \to \mathbb{R}$, $i = 1, \dots, N, k = 1, \dots, n_i$ is defined for any profile of players' states $\overline{s}$ and any profile of players' actions $\overline{a}$ by

$$r_n^{i,k}(\overline{s}, \overline{a}) := r^i\left(s_k^i, a_k^i, \overline{\tau}(\overline{s}, \overline{a})\right).$$

- The transition probability $Q_n : \Pi_{j=1}^N (D^i)^{n_i} \to \Delta(S^n)$ can be defined for any $\overline{s}$ and $\overline{a}$ by the formula (for the clarity of exposition we write it only for Borel rectangles, which obviously defines the product measure):

$$Q_n(B_1^1 \times \ldots \times B_{n_1}^1 \ldots \times B_1^N \times \ldots \times B_{n_N}^N \mid \overline{s}, \overline{a})$$
$$:= Q^1\left(B_1^1 \mid s_1^1, a_1^1, \overline{\tau}(\overline{s}, \overline{a})\right) \ldots Q^1\left(B_{n_1}^1 \mid s_{n_1}^1, a_{n_1}^1, \overline{\tau}(\overline{s}, \overline{a})\right)$$
$$\ldots Q^N\left(B_1^N \mid s_1^N, a_1^N, \overline{\tau}(\overline{s}, \overline{a})\right) \ldots Q^N\left(B_{n_N}^N \mid s_{n_N}^N, a_{n_N}^N, \overline{\tau}(\overline{s}, \overline{a})\right).$$

- In n-person game we assume that the players are limited to policies depending on their own state, that is each player from population i uses Markov policies from the set $\mathcal{M}^i$ or (more specifically) stationary policies from $\mathcal{F}^i$.
- The functional maximized by each player is either his β-discounted reward reward or his total reward. The definitions of both are slight modifications of those for mean-field models. The β-discounted reward of kth player in population i is defined for any initial state $\overline{s_0}$ and any profile of policies of all the players $\overline{\pi}$ as

$$J_{\beta,n}^{k,i}(\overline{s_0}, \overline{\pi}) = \mathbb{E}^{\overline{s_0}, Q_n, \overline{\pi}} \sum_{t=0}^{\infty} \beta^t r_n^{i,k}(\overline{s_t}, \overline{a_t}),$$

with $\mathbb{P}^{\overline{s_0}, Q_n, \overline{\pi}}$ denoting the measure on the set of all infinite histories of the game corresponding to $\overline{s_0}$, Q_n and $\overline{\pi}$ defined with the help of the Ionescu-Tulcea theorem similarly as in case of the mean-field game.
Similarly, the total reward of kth player in population i is defined for any initial state $\overline{s_0}$ and any profile of policies of all the players $\overline{\pi}$ as

$$J_{*n}^{i,k}(\overline{s_0}, \overline{\pi}) = \mathbb{E}^{\overline{s_0}, Q_n, \overline{\pi}} \sum_{t=0}^{\mathcal{T}_k^i - 1} r_n^{i,k}(\overline{s_t}, \overline{a_t}),$$

with $\mathcal{T}_k^i$ denoting the moment of the first arrival of the process $\{s_{k,t}^i\}$ to s^*.
- Finally, the solution we will be looking for in n-person counterparts of the stochastic game is a variant of Nash equilibrium, the standard solution concept used in the stochastic game literature:

Definition 3 A profile of strategies $\overline{\pi} \in \Pi_{i=1}^N \left(\mathcal{M}^i\right)^{n_i}$ is a *Markov–Nash equilibrium* in the n-person discounted-reward game if

$$\mathbb{E}\left[J_{\beta,n}^{k,i}(\overline{s_0}, \overline{\pi}) \mid \overline{s_0} \sim \overline{\mu}_0^*\right] \geq \mathbb{E}\left[J_{\beta,n}^{k,i}(\overline{s_0}, [\overline{\pi}_{-i,k}, \widehat{\pi_k^i}]) \mid \overline{s_0} \sim \overline{\mu}_0^*\right] \quad (3)$$

for any $\widehat{\pi_k^i} \in \mathcal{M}^i$, and $i \in \{1, \ldots, N\}$, $k \in \{1, \ldots, n_i\}$.

The notation $[\overline{\pi}_{-i,k}, \widehat{\pi_k^i}]$ denotes here and in the sequel the profile of policies $\overline{\pi}$ with the policy of kth player in population i replaced by $\widehat{\pi_k^i}$. If (3) is true up to some $\varepsilon > 0$, we say that $\overline{\pi}$ is an ε-*Markov–Nash equilibrium*.

In case of the total reward, we will further reduce the requirements for our approximate solution. We will say that a profile of strategies $\overline{\pi} \in \Pi_{i=1}^N \left(\mathcal{M}^i\right)^{n_i}$ is an (ε, T)-*Markov–Nash equilibrium* in the n-person total-reward game if for $t \in \{0, \ldots, T\}$, $i \in \{1, \ldots, N\}$, $k \in \{1, \ldots, n_i\}$ and $\widehat{\pi_k^i} \in \mathcal{M}^i$:

$$\mathbb{E}\left[J_{*n}^{k,i}(\overline{\mu_t}, {}^t\overline{\pi}) \mid \overline{s_0} \sim \overline{\mu}_0^*\right] \geq \mathbb{E}\left[J_{*n}^{k,i}(\overline{\mu_t}, [{}^t\overline{\pi}_{-i,k}, {}^t\widehat{\pi_k^i}]) \mid \overline{s_0} \sim \overline{\mu}_0^*\right] - \varepsilon$$

3 Preliminaries

As we have written, we assume that S and A are metric spaces. The metric on S will be denoted by d_S while that on A by d_A. Whenever we relate to a metric on a product space, we mean the sum of the metrics on its coordinates. Some of the assumptions presented below will be given with respect to the *moment function* $w_0 : S \to [1, \infty)$, that is a continuous function satisfying

$$\lim_{n\to\infty} \inf_{s\in S\setminus K_n} w_0(s) = \infty$$

for some sequence $\{K_n\}_{n\geq 1}$ of compact subsets of S. Moreover,

$$w_0(s) \geq 1 + d_S(s, s_0)^p \tag{4}$$

for some $p \geq 1$ and $s_0 \in S$.

In order to study both bounded and unbounded one-stage reward functions, we define the following function:

$$w := \begin{cases} 1, & \text{if each } r_i \text{ is bounded} \\ w_0, & \text{otherwise} \end{cases}$$

For any function $h : S \to \mathbb{R}$ we define its w-norm as

$$\|h\|_w := \sup_{s\in S} \left|\frac{h(s)}{w(s)}\right|.$$

Whenever we speak of functions defined on a product of S and some other space, their w-norm is defined similarly, with the help of the same function w.

By $B_w(S)$ we denote the space of all real-valued measurable functions from S to $\mathbb{R}$ with finite w-norm. and by $C_w(S)$ – the space of all continuous functions in $B_w(S)$. Clearly, both $B_w(S)$ and $C_w(S)$ are Banach spaces. The same can be said of $B_w(S\times A)$ and $C_w(S\times A)$ – the spaces of bounded and bounded continuous functions from $S\times A$ to $\mathbb{R}$ with finite w-norm.

Analogously, for any finite signed measure μ on S, we define the w-norm of μ as

$$\|\mu\|_w = \sup_{g\in B_w(S), \|g\|_w\leq 1} \left|\int_S g(s)\mu(ds)\right|.$$

It should be noted that in case $w \equiv 1$, $\|\mu\|_w$ is the total variation distance (see e.g. [8], Section 7.2).

There are two standard types of convergence of probability measures which are used in the paper: the weak convergence denoted by $\Rightarrow$ and the strong (or setwise) convergence denoted by $\to$ and defined (for any Borel space $(X, \mathcal{B}(X))$) by

$$\mu_n \to \mu \iff \mu_n(B) \to \mu(B) \text{ for any } B \in \mathcal{B}(X).$$

It is known (see e.g. [9], Theorem 6.6) that the weak topology can be metrized using the metric

$$\rho(\mu, \nu) := \sum_{m=1}^{\infty} 2^{-m} \left|\int_S \phi_m(s)\mu(ds) - \int_S \phi_m(s)\nu(ds)\right|,$$

where $\{\phi_i\}_{i\geq 1}$ is a sequence of continuous bounded functions from S to $\mathbb{R}$ whose elements form a dense subset of the unit ball in $C(S)$. Strong convergence topology is in general not metric.

Next, let

$$\Delta_w(S) := \left\{ \mu \in \Delta(S) : \int_S w(s)\mu(ds) < \infty \right\}.$$

It has been shown in [11] that $\Delta_w(S)$ can be metrized using the metric

$$\rho_w(\mu, \nu) := \rho(\mu, \nu) + \left| \int_S w(s)\mu(ds) - \int_S w(s)\nu(ds) \right|$$

It can be shown that under (4) $\Delta_w(S)$ with metric ρ_w is a Polish space (see [3, 11] for more on that). We will use the topology defined by this metric (called w-topology in the sequel) as the standard topology on $\Delta_w(S)$.

We will also use the notation

$$\Delta_w(S \times A) := \left\{ \tau \in \Delta(S \times A) : \int_{S\times A} w(s)\tau(ds \times da) < \infty \right\}$$

with analogously defined metrics also denoted by ρ (metric defining weak convergence) and ρ_w (w-metric) as well as similar notation for subsets of S or $S \times A$.

Whenever we speak about continuity of correspondences, we refer to the following definitions:
Let X and Y be two metric spaces and $F : X \to Y$, a correspondence. Let $F^{-1}(G) = \{x \in X : F(x) \cap G \neq \emptyset\}$. We say that F is upper semicontinuous iff $F^{-1}(G)$ is closed for any closed $G \subset Y$. F is lower semicontinuous iff $F^{-1}(G)$ is open for any open $G \subset Y$. F is said to be continuous iff it is both upper and lower semicontinuous. For more on (semi)continuity of correspondences see [7], Appendix D or [1], Chapter 17.2.

4 Assumptions

In this section, we present the set of assumptions used in our results. It contains the assumptions from [18] used there to prove the existence of Markov mean-field equilibria in games with either discounted or total payoff and new assumption (A5) necessary to show that these equilibria are approximate equilibria for games with large finite number of players. The numbering of assumptions (including prime symbols) is consistent with that used in [18], where the basic versions of the assumptions were the strongest ones, used to prove the existence of stationary equilibria in mean-field games. We start by the assumptions used in the discounted case.

(A1') For $i = 1, \ldots, N$, r^i is continuous and bounded above by some constant R on $D^i \times \Pi_{i=1}^N \Delta(D^i)$. Moreover, there exist non-negative constants α, γ, M satisfying $\alpha\beta\gamma < 1$ and

$$\int_S w(s)\mu_0^i(ds) \leq M \quad \text{for } i = 1, \ldots, N,$$

and such that for $i = 1, \ldots, N$, $s \in S^i$ and $t = 0, 1, 2, \ldots$,

$$\inf_{(a,\overline{\tau})\in A^i(s)\times\Pi_{i=1}^N \Delta_w^{(t)}(D^i)} r^i(s, a, \overline{\tau}) \geq -R\gamma^t w(s)$$

with $\Delta_w^{(t)}(D^i) := \left\{ \tau^i \in \Delta_w(D^i) : \int_{D^i} w(s)\tau^i(ds \times da) \leq \alpha^t M \right\}$.

(A2') For $i = 1, \ldots, N$ and any sequence $\{s_n, a_n, \overline{\tau}_n\} \subset D^i \times \Pi_{i=1}^N \Delta_w(D^i)$ such that $s_n \to s_*$, $a_n \to a_*$ and $\overline{\tau}_n \Rightarrow \overline{\tau}^*$, $Q^i(\cdot \mid s_n, a_n, \overline{\tau}_n) \Rightarrow Q(\cdot \mid s_*, a_*, \overline{\tau}^*)$. Moreover,

(a) for $i = 1, \ldots, N$ the functions

$$\int_S w(s')Q^i(ds' \mid s, a, \overline{\tau})$$

are continuous in $(s, a, \overline{\tau})$,

(b) for $i = 1, \ldots, N$ and $s \in S^i$

$$\sup_{(a,\overline{\tau}) \in A^i(s) \times \Pi_{i=1}^N \Delta(D^i)} \int_S w(s')Q^i(s' \mid s, a, \overline{\tau}) \leq \alpha w(s).$$

(A3) For $i = 1, \ldots, N$, correspondences A^i are continuous.

Assumptions (A1') and (A2') are modified for the total payoff case and complemented by new assumption (A4"). Their formulation requires defining for $i = 1, \ldots, N$, $s \in S^i$, $a \in A^i(s)$ and $\overline{\tau} \in \Pi_{j=1}^N \Delta(D^j)$ the modified transition probabilities Q^i_*:

$$Q^i_*(\cdot \mid s, a, \overline{\tau}) := \begin{cases} Q^i(\cdot \mid s, a, \overline{\tau}), & \text{if } s \neq s^* \\ \delta_{s^*}, & \text{if } s = s^* \end{cases}$$

(A1") For $i = 1, \ldots, N$, r^i is continuous and bounded above by some constant R on $D^i \times \Pi_{i=1}^N \Delta(D^i)$. Moreover, there exist non-negative constants α, γ, M satisfying $\alpha \leq \gamma$, $\alpha\gamma < 1$ and

$$\int_S w(s)\mu_0^i(ds) \leq M \quad \text{for } i = 1, \ldots, N,$$

and such that for $i = 1, \ldots, N$, $s \in S^i$ and $t = 0, 1, 2, \ldots$,

$$\inf_{(a,\overline{\tau}) \in A^i(s) \times \Pi_{i=1}^N \Delta_w^{(t)}(D^i)} r^i(s, a, \overline{\tau}) \geq -R\gamma^t w(s)$$

with $\Delta_w^{(t)}(D^i) := \left\{\tau^i \in \Delta_w(D^i) : \int_{D^i} w(s)\tau^i(ds \times da) \leq \alpha^t M\right\}$.

(A2") For $i = 1, \ldots, N$ and any sequence $\{s_n, a_n, \overline{\tau}_n\} \subset D^i \times \Pi_{i=1}^N \Delta_w(D^i)$ such that $s_n \to s_*$, $a_n \to a_*$ and $\overline{\tau}_n \Rightarrow \tau^*$, $Q^i(\cdot \mid s_n, a_n, \overline{\tau}_n) \to Q(\cdot \mid s_*, a_*, \tau^*)$. Moreover,

(a) for $i = 1, \ldots, N$ the functions

$$\int_S w(s')Q^i(ds' \mid s, a, \overline{\tau})$$

are continuous in $(s, a, \overline{\tau})$,

(b) for $i = 1, \ldots, N$ and $s \in S^i$

$$\sup_{(a,\overline{\tau}) \in A^i(s) \times \Pi_{i=1}^N \Delta(D^i)} \int_S w(s')Q^i(s' \mid s, a, \overline{\tau}) \leq \alpha w(s).$$

(A4") For $i = 1, \ldots, N$,

$$\lim_{T \to \infty} \sup_{\substack{\pi^i \in \mathcal{M}^i, \\ (\overline{\tau}) \in \Pi_{t=0}^\infty \Pi_{j=1}^N \Delta(D^j)}} \left\| \sum_{t=T}^\infty \int_{S^i \setminus \{s^*\}} w(s')\alpha^{-t} \left(Q^i_*\right)^t (ds' \mid s, \pi^i, (\overline{\tau})) \right\|_w = 0.$$

Finally, we present additional assumptions used to prove the approximation theorems. Beforehand, for $i = 1, \ldots, N$ let us define the following moduli of continuity:

$$\omega_{Q^i}(\delta) := \sup_{(s,a)\in D^i} \sup_{\overline{\tau},\overline{\eta}:\sum_{j=1}^N \widetilde{\rho}_w(\tau^j,\eta^j)\leq\delta} \left\| Q^i(\cdot \mid s, a, \overline{\tau}) - Q^i(\cdot \mid s, a, \overline{\eta}) \right\|_w,$$

$$\omega_{r^i}(\delta) := \sup_{(s,a)\in D^i} \sup_{\overline{\tau},\overline{\eta}:\sum_{j=1}^N \widetilde{\rho}_w(\tau^j,\eta^j)\leq\delta} \left| r^i(s, a, \overline{\tau}) - r^i(s, a, \overline{\eta}) \right|,$$

where $\widetilde{\rho}_w = \rho_w$ if any r^i is unbounded, or $\widetilde{\rho}_w = \rho$ otherwise.

For any function $g : \Pi_{i=1}^N \Delta_w(D^i) \to \mathbb{R}$, $i = 1, \ldots, N$ we next define its w-norm as follows:

$$\|g\|_w^* := \sup_{\overline{\tau}\in\Pi_{i=1}^N \Delta_w(D^i)} \frac{|g(\overline{\tau})|}{\sum_{i=1}^N \int_{D^i} w(s)\tau^i(ds \times da)}.$$

Now we can formulate our additional assumptions. They adapt the assumptions used in [11] to our multi-population case.

(A5) (a) We assume that $\omega_{Q^i}(\delta) \to 0$ and $\omega_{r^i}(\delta) \to 0$ for $i = 1, \ldots, N$ as $\delta \to 0$. Moreover, the following real-valued functions defined on $\Pi_{i=1}^N \Delta_w(D^i)$:

$$\Omega_Q^{\overline{\tau}}(\overline{\eta}) := \max_i \omega_{Q^i}\Big(\sum_{j=1}^N \widetilde{\rho}_w(\eta^j, \tau^j)\Big) \text{ and } \Omega_r^{\overline{\tau}}(\overline{\eta}) := \max_i \omega_{r^i}\Big(\sum_{j=1}^N \widetilde{\rho}_w(\eta^j, \tau^j)\Big)$$

have finite w-norm.

(b) There exist non-negative real numbers B and B_0 such that for $i = 1, \ldots, N$ and $s \in S^i$

$$\sup_{(a,\overline{\tau})\in A^i(s)\times\Pi_{i=1}^N \Delta(D^i)} \int_S w^2(s')Q^i(s' \mid s, a, \overline{\tau}) \leq Bw^2(s)$$

and $\int_S w^2(s)\mu_0^{*i}(ds) \leq B_0$.

5 Main Results

5.1 Results for the Discounted Payoff Case

In the first of our main results we address the case of discounted reward game.

Theorem 1 *Suppose assumptions (A1'), (A2'), (A3) and (A5) hold and suppose $\overline{\pi}$ and $(\overline{\mu}_0^*, \overline{\mu}_1^*, \ldots)$ form a Markov mean-field equilibrium in the multi-population discrete-time mean-field game existing by Theorem 4 in [18]. If in addition, for each $t \geq 0$ and $i = 1, \ldots, N$, π_t^i is weakly continuous, then for any $\varepsilon > 0$ there exist positive integers $n_i(\varepsilon)$, $i = 1, \ldots, N$ such that the vector of strategies where each player from population i uses policy π^i is an ε-Markov–Nash equilibrium in any n-person stochastic counterpart of the β-discounted mean-field game if $n^i \geq n^i(\varepsilon)$, $i = 1, \ldots, N$.*

Remark 1 Note that stationary mean-field equilibrium existing according to Theorem 1 in [18] is a specific case of Markov mean-field equilibrium with stationarity condition imposed on global states of the game at subsequent stages. Hence, the result provided by Theorem 1 also holds in this case.

The proof of Theorem 1 will adapt the techniques used in [11] to our model. It will require introducing some additional notation. Recall the Polish space $(\Delta_w(S\times A),\rho_w)$. Define the Wasserstein distance of order 1 on the set of probability measures over $\Delta_w(S\times A)$, $\Delta(\Delta_w(S\times A))$ by the formula:

$$W_1(\Phi,\Psi):=\inf_{\gamma\in\Gamma(\Phi,\Psi)}\int_{\Delta(\Delta_w(S\times A))\times\Delta(\Delta_w(S\times A))} d_S(x,y)\,d\gamma(x,y)\},$$

where $\Gamma(\Phi,\Psi)$ denotes the collection of all measures on $\Delta(\Delta_w(S\times A))\times\Delta(\Delta_w(S\times A))$ with marginals Φ and Ψ on the 1st and the 2nd coordinate.

Next, for $i=1,\ldots,N$, define the following spaces:

$$\Delta_1^i(\Delta_w(S\times A)):=\left\{\Phi\in\Delta(\Delta_w(S\times A)):\int_{\Delta_w(S\times A)}\rho_w(\tau,\Pi_0^i(\pi^i,\mu_0^{*i}))\Phi(d\tau)<\infty\right\},$$

$$C_w(\Pi_{i=1}^N\Delta_w(D^i)):=\left\{g:\Pi_{i=1}^N\Delta_w(D^i)\to\mathbb{R}:\ g\text{ is continuous and }\|g\|_w^*<\infty\right\}.$$

Before we get to the actual proof of the theorem, note that the game is symmetric, hence, proving only that the inequality defining ε-Nash equilibrium holds for the first player in the first population will be enough to verify the theorem.

In our proof we shall use the following notation:

- We will use the notation $\overline{\pi}$ also to denote the vector of strategies in the n-person counterpart of the mean-field game where each player from population i uses strategy π^i. The strategy vector where the first player in the first population changes his strategy to an arbitrary weakly continuous Markov strategy $\widehat{\pi_1^1}$ will be denoted by $\widehat{\pi}$.
- The vector of states at time t in n-person game with n_i players in population i, $i=1,\ldots,N$ will be denoted by $\overline{s}_t^{\mathbf{n}}=(\overline{s}_t^{\mathbf{n},1},\ldots,\overline{s}_t^{\mathbf{n},N})$ with $\overline{s}_t^{\mathbf{n},i}=(s_{t,1}^{\mathbf{n},i},\ldots,s_{t,n_i}^{\mathbf{n},i})$ for $i=1,\ldots,N$. Similarly, the vector of actions at time t will be denoted by $\overline{a}_t^{\mathbf{n}}=(\overline{a}_t^{\mathbf{n},1},\ldots,\overline{a}_t^{\mathbf{n},N})$ with $\overline{a}_t^{\mathbf{n},i}=(a_{t,1}^{\mathbf{n},i},\ldots,a_{t,n_i}^{\mathbf{n},i})$ for $i=1,\ldots,N$. The corresponding empirical state-action distribution of i-th population will be denoted by $e_i^{\mathbf{n}}$.
- When we want to distinguish between what happens when strategy vector $\overline{\pi}$ is used and when $\widehat{\pi}$, an overline or a hat is added to a specific symbol, in particular the empirical state-action distributions at time t when the two strategy vectors are used will be denoted by $\overline{e}_{i,t}^{\mathbf{n}}$ and $\widehat{e}_{i,t}^{\mathbf{n}}$, respectively, with $\overline{\overline{e}_t^{\mathbf{n}}}=(\overline{e}_{1,t}^{\mathbf{n}},\ldots,\overline{e}_{N,t}^{\mathbf{n}})$ and $\overline{\widehat{e}_t^{\mathbf{n}}}=(\widehat{e}_{1,t}^{\mathbf{n}},\ldots,\widehat{e}_{N,t}^{\mathbf{n}})$, while the state and the action of the first player in the first population at time t by $\bar{s}_{t,1}^{\mathbf{n},1}$, $\bar{a}_{t,1}^{\mathbf{n},1}$, and by $\hat{s}_{t,1}^{\mathbf{n},1}$, $\hat{a}_{t,1}^{\mathbf{n},1}$, respectively.
- For any random element θ, its distribution will be denoted by $\mathcal{L}(\theta)$. In particular, the distributions of random elements $(\bar{s}_{t,1}^{\mathbf{n},1},\bar{a}_{t,1}^{\mathbf{n},1})$, $(\hat{s}_{t,1}^{\mathbf{n},1},\hat{a}_{t,1}^{\mathbf{n},1})$, $\overline{e}_{i,t}^{\mathbf{n}}$ and $\widehat{e}_{i,t}^{\mathbf{n}}$ will be denoted by $\mathcal{L}\left(\bar{s}_{t,1}^{\mathbf{n},1},\bar{a}_{t,1}^{\mathbf{n},1}\right)$, $\mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},1},\hat{a}_{t,1}^{\mathbf{n},1}\right)$, $\mathcal{L}\left(\overline{e}_{i,t}^{\mathbf{n}}\right)$ and $\mathcal{L}\left(\widehat{e}_{i,t}^{\mathbf{n}}\right)$, respectively.
- The equilibrium state distribution in the mean-field limit at time t will be denoted by μ_t^{*i}, $i=1,\ldots,N$ while equilibrium state-action distribution in the mean-field limit $\Pi_t^i(\pi^i,\mu_t^{*i})$ by τ_t^{*i} with $\overline{\mu_t^*}$ and $\overline{\tau_t^*}$ standing for their vectors. Finally, state and action of the first player in i-th population at time t if the first player in the first population uses policy $\widehat{\pi_1^1}$ while others stick to their policies in the mean-field equilibrium, will be denoted by $\hat{s}_{t,1}^i$ and $\hat{a}_{t,1}^i$ with $\mathcal{L}\left(\hat{s}_{t,1}^i,\hat{a}_{t,1}^i\right)$ their joint distribution.

In the first lemma, we adapt one of the results from Lemma 4.3 in [11] to our multidimensional case.

Lemma 2 *Let* $\Phi^i_{\mathbf{n}} \in \Delta^i_1(\Delta_w(S \times A))$ *and* $\delta_{\tau^*_i} \in \Delta^i_1(\Delta_w(S \times A))$ *for* $\mathbf{n} \in \mathbf{N}^N$ *and* $i = 1, \ldots, N$. *Then if* $W_1(\Phi^i_{\mathbf{n}}, \delta_{\tau^{*i}}) \to 0$ *for* $i = 1, \ldots, N$ *as* $\mathbf{n} \to \infty$, *then*

$$\mathbb{E}\left[|F(\overline{\tau}_{\mathbf{n}}) - F(\overline{\tau})|\right] \to 0 \quad as\ \mathbf{n} \to 0$$

for any $F \in C_w\left(\Pi^N_{i=1}\Delta_w(D^i)\right)$ *and any* $\Pi^N_{i=1}\Delta_w(D^i)$*-valued random elements* $\overline{\tau}_{\mathbf{n}} = \left(\tau^1_{\mathbf{n}}, \ldots, \tau^N_{\mathbf{n}}\right)$, $\mathbf{n} \in \mathbf{N}^N$, *such that* $\mathcal{L}\left(\tau^i_{\mathbf{n}}\right) = \Phi^i_{\mathbf{n}}$, $i = 1, \ldots, N$, *and a* $\Pi^N_{i=1}\Delta_w(D^i)$*-valued random element* $\overline{\tau} = \left(\tau^1, \ldots, \tau^N\right)$ *such that* $\mathcal{L}\left(\tau^i\right) = \delta_{\tau^{*i}}$, $i = 1, \ldots, N$

Proof In the proof we will inductively (with respect to N) verify a stronger result, stating that for any function F satisfying the assumptions of the lemma, the functions $\widetilde{F}_{\mathbf{n}} : \Delta_w(D^N) \to \mathbb{R}$, defined as

$$\widetilde{F}_{\mathbf{n}}(\tau^N) = \mathbb{E}\left[F(\tau^1_{\mathbf{n}}, \ldots, \tau^{N-1}_{\mathbf{n}}, \tau^N) \mid \mathcal{L}(\tau^1_{\mathbf{n}}) = \Phi^1_{\mathbf{n}}, \ldots, \mathcal{L}(\tau^{N-1}_{\mathbf{n}}) = \Phi^{N-1}_{\mathbf{n}}\right]$$

and $\widehat{F} : \Delta_w(D^N) \to \mathbb{R}$, defined as

$$\widehat{F}(\tau^N) = F(\tau^{*1}, \ldots, \tau^{*N-1}, \tau^N)$$

satisfy for any convergent sequence $\left\{\tau^N_k\right\}_{k\geq 1} \subset \Delta_w(D^N)$

$$\widetilde{F}_{\mathbf{n}}(\tau^N_k) \to_{\mathbf{n}\to\infty, k\to\infty} \widehat{F}(\lim_{k\to\infty} \tau^N_k). \tag{5}$$

We precede the main part of the proof by showing that for any $\mathbf{n} \in \mathbb{N}^N$ and any $F \in C_w\left(\Pi^N_{i=1}\Delta_w(D^i)\right)$, $\widetilde{F}_{\mathbf{n}} \in C_w(\Delta_w(D^N))$.

Note that from the definition of the $\|\cdot\|^*_w$ norm we know that for any $\overline{\tau} \in \Pi^N_{i=1}\Delta_w(D^i)$

$$\begin{aligned}\frac{|F(\overline{\tau})|}{\int_{D^N} w(s)\tau^N(ds \times da)} &\leq \|F\|^*_w \left(1 + \frac{\sum_{i=1}^{N-1}\int_{D^i} w(s)\tau^i(ds \times da)}{\int_{D^N} w(s)\tau^N(ds \times da)}\right) \\ &\leq \|F\|^*_w \left(1 + \sum_{i=1}^{N-1}\int_{D^i} w(s)\tau^i(ds \times da)\right),\end{aligned}$$

with the last inequality following from the fact that $w \geq 1$. Consequently,

$$\frac{\left|\widetilde{F}_{\mathbf{n}}(\tau^N)\right|}{\int_{D^N} w(s)\tau^N(ds \times da)} \leq \|F\|^*_w \left(1 + \sum_{i=1}^{N-1}\mathbb{E}\left[\int_{D^i} w(s)\tau^i(ds \times da) \mid \mathcal{L}(\tau^i_{\mathbf{n}}) = \Phi^i_{\mathbf{n}}\right]\right).$$

By Lemma 4.3 in [11] we know that the RHS of the above inequality converges to

$$\|F\|^*_w \left(1 + \sum_{i=1}^{N-1}\int_{D^i} w(s)\tau^{*i}(ds \times da)\right) < \infty$$

as $\mathbf{n} \to \infty$. This however implies that there exists a number W^N_* such that for every $\mathbf{n} \in \mathbb{N}^N$,

$$\frac{\left|\widetilde{F}_{\mathbf{n}}(\tau^N)\right|}{\int_{D^N} w(s)\tau^N(ds \times da)} \leq W^N_*,$$

which means that for each $\mathbf{n}$, $\left\|\widetilde{F}_{\mathbf{n}}\right\|^*_w \leq W^N_*$.

To show that for each $\mathbf{n}$, $\widetilde{F}_{\mathbf{n}}$ is continuous, we take a convergent sequence $\{\tau_k^N\}_{k\geq 1} \subset \Delta_w(D^N)$ and note that

$$\begin{aligned}
&\lim_{k\to\infty}\left|\widetilde{F}_{\mathbf{n}}(\tau_k^N) - \widehat{F}(\lim_{k\to\infty}\tau_k^N)\right| \\
&= \lim_{k\to\infty}\left|\mathbb{E}\left[F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\tau_k^N) - F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\lim_{k\to\infty}\tau_k^N)\right.\right. \\
&\quad\left.\left.\mid \mathcal{L}(\tau_{\mathbf{n}}^1) = \Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1}) = \Phi_{\mathbf{n}}^{N-1}\right]\right| \\
&\leq \lim_{k\to\infty}\mathbb{E}\left[\left|F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\tau_k^N) - F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\lim_{k\to\infty}\tau_k^N)\right|\right. \\
&\quad\left.\mid \mathcal{L}(\tau_{\mathbf{n}}^1) = \Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1}) = \Phi_{\mathbf{n}}^{N-1}\right]
\end{aligned} \tag{6}$$

Note that by the definition of the w-topology, the sequence of integrals $\int_{D^N} w(s)\tau_k^N(ds\times da)$ converges to $\int_{D^N} w(s)\lim_{k\to\infty}\tau_k^N(ds\times da)$, hence there exists a number $W^N > 0$ such that

$$\int_{D^N} w(s)\tau_k^N(ds\times da) \leq W^N \quad \text{for } k\geq 1 \tag{7}$$

Next, note that by the definition of the $\|\cdot\|_w^*$ norm,

$$\left|F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\tau_k^N)\right| \leq \|F\|_w^*\left(\sum_{i=1}^{N-1}\int_{D^i} w(s)\tau_{\mathbf{n}}^i(ds\times da) + \int_{D^N} w(s)\tau_k^N(ds\times da)\right)$$

The function on the RHS is the sum of a function of variables $\tau_{\mathbf{n}}^i$, $i=1,\ldots,N-1$ with a finite integral with respect to the measure $\Pi_{i=1}^{N-1}\Phi_{\mathbf{n}}^i$ by the assumption of the lemma and a bounded term independent of these variables (a function of τ_k^N bounded by $\|F\|_w^* W^N$ for any $k\geq 1$ by (7)). Hence, by the dominated convergence theorem the RHS of (6) equals zero for any $\mathbf{n}$, which implies that $\widetilde{F}_{\mathbf{n}}$ is continuous.

Next, we turn to the inductive proof of (5). It is obvious that it holds for $N=1$. Suppose that (5) holds for some $N-1$ for any function satisfying the assumptions of the lemma. We will show it is also true for N. First, note that

$$\begin{aligned}
&\lim_{\mathbf{n}\to\infty,k\to\infty}\left|\widetilde{F}_{\mathbf{n}}(\tau_k^N) - \widehat{F}(\lim_{k\to\infty}\tau_k^N)\right| \\
&= \lim_{\mathbf{n}\to\infty,k\to\infty}\left|\mathbb{E}\left[F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\tau_k^N)\mid \mathcal{L}(\tau_{\mathbf{n}}^1) = \Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1}) = \Phi_{\mathbf{n}}^{N-1}\right]\right. \\
&\quad\left.-F(\tau^{*1},\ldots,\tau^{*N-1},\lim_{k\to\infty}\tau_k^N)\right| \\
&\leq \lim_{\mathbf{n}\to\infty,k\to\infty}\mathbb{E}\left[\left|F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\tau_k^N) - F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\lim_{k\to\infty}\tau_k^N)\right|\right. \\
&\quad\left.\mid \mathcal{L}(\tau_{\mathbf{n}}^1) = \Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1}) = \Phi_{\mathbf{n}}^{N-1}\right] \\
&\quad+\lim_{\mathbf{n}\to\infty}\left|\mathbb{E}\left[F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\lim_{k\to\infty}\tau_k^N) - F(\tau^{*1},\ldots,\tau^{*N-1},\lim_{k\to\infty}\tau_k^N)\right.\right. \\
&\quad\left.\left.\mid \mathcal{L}(\tau_{\mathbf{n}}^1) = \Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1}) = \Phi_{\mathbf{n}}^{N-1}\right]\right|
\end{aligned} \tag{8}$$

To show that the first term on the RHS of (8) equals zero, we define functions $\theta_k^N, \theta^N : \Pi_{i=1}^{N-1}\Delta_w(D^i) \to \mathbb{R}$ as

$$\theta_k^N(\tau^1,\ldots,\tau^{N-1}) = F(\tau^1,\ldots,\tau^{N-1},\tau_k^N),$$
$$\theta^N(\tau^1,\ldots,\tau^{N-1}) = F(\tau^1,\ldots,\tau^{N-1},\lim_{k\to\infty}\tau_k^N).$$

The definition of the $\|\cdot\|_w^*$ norm, (7) and the fact that $w \geq 1$ imply then that

$$\left\|\theta_k^N\right\|_w^* \leq \|F\|_w^*(1+W^N) \text{ for } k \geq 1 \quad \text{and} \quad \left\|\theta^N\right\|_w^* \leq \|F\|_w^*(1+W^N).$$

As θ_k^N converges continuously to θ^N, by Theorem 3.3 in [15], the first term on the RHS of (8) is zero.

To show that the same is true for the second term, we first rewrite it as follows:

$$\begin{aligned}
&\lim_{\mathbf{n}\to\infty}\left|\mathbb{E}\left[F(\tau_{\mathbf{n}}^1,\ldots,\tau_{\mathbf{n}}^{N-1},\lim_{k\to\infty}\tau_k^N) - F(\tau^{*1},\ldots,\tau^{*N-1},\lim_{k\to\infty}\tau_k^N)\right.\right.\\
&\quad\left.\left.\mid \mathcal{L}(\tau_{\mathbf{n}}^1)=\Phi_{\mathbf{n}}^1,\ldots,\mathcal{L}(\tau_{\mathbf{n}}^{N-1})=\Phi_{\mathbf{n}}^{N-1}\right]\right|\\
&= \lim_{\mathbf{n}\to\infty}\left|\int_{\Delta_w(S\times A)}\left(\widetilde{\left(\theta^N\right)}_{\mathbf{n}}(\tau_{\mathbf{n}}^{N-1}) - \widehat{\theta^N}(\tau^{*N-1})\right)\Phi_{\mathbf{n}}^{N-1}\left(d\tau_{\mathbf{n}}^{N-1}\right)\right|\\
&\leq \lim_{\mathbf{n}\to\infty}\left[\int_{\Delta_w(S\times A)}\left|\widetilde{\left(\theta^N\right)}_{\mathbf{n}}(\tau_{\mathbf{n}}^{N-1}) - \widetilde{\left(\theta^N\right)}_{\mathbf{n}}(\tau^{*N-1})\right|\Phi_{\mathbf{n}}^{N-1}\left(d\tau_{\mathbf{n}}^{N-1}\right)\right.\\
&\quad\left.+\left|\widetilde{\left(\theta^N\right)}_{\mathbf{n}}(\tau^{*N-1}) - \widehat{\theta^N}(\tau^{*N-1})\right|\right]
\end{aligned}$$

The second term goes to zero by the inductive hypothesis. To show that the same is true for the first one, note that θ^N is a continuous function of $N-1$ variables with a finite $\|\cdot\|_w^*$ norm, hence, for any $\mathbf{n} \in \mathbb{N}^N$, $\widetilde{\left(\theta^N\right)}_{\mathbf{n}} \in C_w(\Delta_w(D^{N-1}))$. Now we can apply Lemma 4.3 from [11] to show that the first term also goes to zero, ending the proof. □

In the second lemma, we show that the sequences of random measures $\widehat{e}_i^{\mathbf{n}}$ converge in some sense to the mean-field equilibrium state-action distributions τ_t^{*i} as $\mathbf{n} \to \infty$.

Lemma 3 *For $i = 1,\ldots,N$ and any $t \geq 0$,*

$$\lim_{\mathbf{n}\to\infty} W_1\left(\mathcal{L}\left(\widehat{e}_{i,t}^{\mathbf{n}}\right), \delta_{\tau_t^{*i}}\right) = 0$$

in $\Delta_1^i(\Delta_w(S\times A))$.

Proof By Lemma 4.3 in [11], to prove the thesis we only need to show that for any i and t,

$$\lim_{\mathbf{n}\to\infty}\mathbb{E}\left[\int_{S\times A} f(s,a)\widehat{e}_{i,t}^{\mathbf{n}}(ds\times da) - \int_{S\times A} f(s,a)\tau_t^{*i}(ds\times da)\right] = 0$$

for any $f \in C_w(S\times A)$. We do it by induction on t.

Suppose $t = 0$. If $i \neq 1$, then $\{(\hat{s}_{0,k}^{\mathbf{n},i}, \hat{a}_{0,k}^{\mathbf{n},i})\}_{1\leq k\leq n_i} \sim \Pi_{k=1}^{n_i}\tau_0^{*i}$. As any $f \in C_w(S\times A)$ is w-integrable by assumption (A1'), the claim holds in this case. Next, suppose that $i = 1$. Then

$$\int_{S\times A} f(s,a)\widehat{e}_{i,t}^{\mathbf{n}}(ds\times da) = \frac{1}{n_1}f(\hat{s}_{0,1}^{\mathbf{n},1},\hat{a}_{0,1}^{\mathbf{n},1}) + \frac{n_1-1}{n_1}\int_{S\times A} f(s,a)\widehat{e}_{1,t}^{\mathbf{n}-\mathbf{e}_1}(ds\times da)$$

with[4] $\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t} = \frac{1}{n_1-1}\sum_{k=2}^{n_1}\delta_{(\hat{s}^{\mathbf{n},1}_{t,k},\hat{a}^{\mathbf{n},1}_{t,k})}$. Now note that the expectation of the first term converges to zero when $n_1 \to \infty$ as

$$\mathbb{E}\left[\frac{1}{n_1} f(\hat{s}^{\mathbf{n},1}_{0,1}, \hat{a}^{\mathbf{n},1}_{0,1})\right] \le \frac{1}{n_1}\|f\|_w M$$

by assumption (A1'). On the other hand, the expected value of the second term goes to $\int_{S\times A} f(s,a)\tau^{*1}_0(ds\times da)$ by the argument used for $i \neq 1$.

Now suppose the claim holds for t and consider $t+1$. The claim will only be proved for $i = 1$. The proof for $i \neq 1$ goes along the same lines (we only do not need to consider the first term on the RHS below in that case). Let us fix $f \in C_w(S\times A)$. Then

$$\begin{aligned}
&\left|\int_{S\times A} f(s,a)\widehat{e}^{\mathbf{n}}_{1,t+1}(ds\times da) - \int_{S\times A} f(s,a)\tau^{*1}_{t+1}(ds\times da)\right| \\
&\le \frac{1}{n_1}\left| f(\hat{s}^{\mathbf{n},1}_{t+1,1}, \hat{a}^{\mathbf{n},1}_{t+1,1}) - \int_{S\times A} f(s',a')\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid \hat{s}^{\mathbf{n},1}_{t,1}, \hat{a}^{\mathbf{n},1}_{t,1}, \overline{\widehat{e}^{\mathbf{n}}_t})\right| \\
&\quad + \frac{n_1-1}{n_1}\left|\int_{S\times A} f(s,a)\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t+1}(ds\times da)\right. \\
&\quad \left. - \int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid s,a,\overline{\widehat{e}^{\mathbf{n}}_t})\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t}(ds\times da)\right| \\
&\quad + \left|\int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid s,a,\overline{\widehat{e}^{\mathbf{n}}_t})\widehat{e}^{\mathbf{n}}_{1,t}(ds\times da)\right. \\
&\quad \left. - \int_{S\times A} f(s,a)\tau^{*1}_{t+1}(ds\times da)\right|
\end{aligned} \tag{9}$$

To finish the proof, we need to show that the expected values of each term on the RHS of (9) go to zero as $\mathbf{n} \to \infty$.

First, let us consider the first term.

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{n_1}\left| f(\hat{s}^{\mathbf{n},1}_{t+1,1}, \hat{a}^{\mathbf{n},1}_{t+1,1}) - \int_{S\times A} f(s',a')\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid \hat{s}^{\mathbf{n},1}_{t,1}, \hat{a}^{\mathbf{n},1}_{t,1}, \overline{\widehat{e}^{\mathbf{n}}_t})\right|\right] \\
&\le \frac{1}{n_1}\mathbb{E}\left[\left| f(\hat{s}^{\mathbf{n},1}_{t+1,1}, \hat{a}^{\mathbf{n},1}_{t+1,1})\right|\right] + \frac{1}{n_1}\mathbb{E}\left[\int_{S\times A} |f(s',a')|\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid \hat{s}^{\mathbf{n},1}_{t,1}, \hat{a}^{\mathbf{n},1}_{t,1}, \overline{\widehat{e}^{\mathbf{n}}_t})\right] \\
&\le \frac{\|f\|_w}{n_1}\mathbb{E}\left[w(\hat{s}^{\mathbf{n},1}_{t+1,1})\right] + \frac{\|f\|_w}{n_1}\mathbb{E}\left[\int_S w(s')Q^1(ds' \mid \hat{s}^{\mathbf{n},1}_{t,1}, \hat{a}^{\mathbf{n},1}_{t,1}, \overline{\widehat{e}^{\mathbf{n}}_t})\right] \le 2\frac{\|f\|_w}{n_1}\alpha^{t+1}M,
\end{aligned}$$

where the last inequality follows from assumption (A1') and a repeated application of (b) of assumption (A2'). Clearly, the last expression goes to zero as $n_1 \to \infty$.

Next, let us consider the expectation of the second term on the RHS of (9). We can write it as

$$\begin{aligned}
&\frac{n_1-1}{n_1}\mathbb{E}\left[\mathbb{E}\left[\left|\int_{S\times A} f(s,a)\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t+1}(ds\times da)\right.\right.\right. \\
&\quad \left.\left.\left. - \int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da' \mid s')Q^1(ds' \mid s,a,\overline{\widehat{e}^{\mathbf{n}}_t})\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t}(ds\times da)\right| \,\middle|\, \overline{\hat{s}^{\mathbf{n}}_t}, \overline{\hat{a}^{\mathbf{n}}_t}\right]\right]
\end{aligned}$$

[4] Here and in the sequel $\mathbf{e}_j$ stands for a versor with one on its j-th coordinate.

The square of it can be bounded above as follows:

$$\left(\frac{n_1-1}{n_1}\mathbb{E}\left[\mathbb{E}\left[\left|\int_{S\times A} f(s,a)\widehat{e}^{\mathbf{n}-\mathbf{e_1}}_{1,t+1}(ds\times da)\right.\right.\right.\right.$$
$$\left.\left.\left.\left.- \int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid s,a,\overline{\widehat{e}^{\mathbf{n}}_t})\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t}(ds\times da)\right|\,\middle|\,\overline{\hat{s}^{\mathbf{n}}_t},\overline{\hat{a}^{\mathbf{n}}_t}\right]\right]\right)^2$$
$$\le \frac{(n_1-1)^2}{n_1^2}\mathbb{E}\left[\left(\mathbb{E}\left[\left|\int_{S\times A} f(s,a)\widehat{e}^{\mathbf{n}-\mathbf{e_1}}_{1,t+1}(ds\times da)\right.\right.\right.\right.$$
$$\left.\left.\left.\left.- \int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid s,a,\overline{\widehat{e}^{\mathbf{n}}_t})\widehat{e}^{\mathbf{n}-\mathbf{e}_1}_{1,t}(ds\times da)\right|\,\middle|\,\overline{\hat{s}^{\mathbf{n}}_t},\overline{\hat{a}^{\mathbf{n}}_t}\right]\right)^2\right]$$
$$\le \frac{1}{n_1^2}\mathbb{E}\left[\sum_{k=2}^{n_1}\left[\int_{S\times A} f^2(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid \hat{s}^{\mathbf{n},1}_{t,k},\hat{a}^{\mathbf{n},1}_{t,k},\overline{\widehat{e}^{\mathbf{n}}_t})\right.\right.$$
$$\left.\left.+\left(\int_{S\times A} f(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid \hat{s}^{\mathbf{n},1}_{t,k},\hat{a}^{\mathbf{n},1}_{t,k},\overline{\widehat{e}^{\mathbf{n}}_t})\right)^2\right]\right]$$
$$\le \frac{\|f\|^2_w}{n_1^2}\mathbb{E}\left[\sum_{k=2}^{n_1}\left[\int_{S\times A} w^2(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid \hat{s}^{\mathbf{n},1}_{t,k},\hat{a}^{\mathbf{n},1}_{t,k},\overline{\widehat{e}^{\mathbf{n}}_t})\right.\right.$$
$$\left.\left.+\left(\int_{S\times A} w(s',a')\pi^1_{t+1}(da'\mid s')Q^1(ds'\mid \hat{s}^{\mathbf{n},1}_{t,k},\hat{a}^{\mathbf{n},1}_{t,k},\overline{\widehat{e}^{\mathbf{n}}_t})\right)^2\right]\right]$$
$$\le \frac{\|f\|^2_w}{n_1^2}\sum_{k=2}^{n_1}\mathbb{E}\left[Bw^2(\hat{s}^{\mathbf{n},1}_{t,k})+\alpha^2w^2(\hat{s}^{\mathbf{n},1}_{t,k})\right]$$

with the second inequality following from Lemma 6.2 in [11], the third one from the definition of the w-norm and the last one from (b) of assumption (A2') and (b) of assumption (A5). As by assumption (A1') and (b) of assumption (A5), for each k and any $\mathbf{n}$, $\mathbb{E}[w^2(\hat{s}^{\mathbf{n},1}_{t,k})]\le B^tB_0$, this implies that the expectaton of the second term on the RHS of (9) also converges to zero when $\mathbf{n}\to 0$.

We finish the proof by showing that the same is true for the third term. In order to do it, let us introduce the function $\phi_t:\Pi_{i=1}^N\Delta_w(D^i)\to\mathbb{R}$ with the formula

$$\phi_t(\overline{\tau}) := \int_{S\times A}\int_{S\times A} f(s',a')\pi^1_{t+1}(da'\mid s)Q^1(ds'\mid s,a,\overline{\tau})\tau^1(ds\times da).$$

Note that the third term on the RHS of (9) can be rewritten using ϕ_t as

$$\left|\phi_t\left(\overline{\widehat{e}^{\mathbf{n}}_t}\right)-\phi_t\left(\tau^{*1}_t\right)\right|. \tag{10}$$

As by the induction hypothesis $\lim_{\mathbf{n}\to\infty}W_1\left(\mathcal{L}\left(\widehat{e}^{\mathbf{n}}_{i,t}\right),\delta_{\tau^{*i}_t}\right)=0$, $i=1,\ldots,N$, by Lemma 2 showing that the expected value of (10) goes to zero as $\mathbf{n}\to\infty$ only requires proving that

$\phi_t \in C_w\left(\Pi_{i=1}^N \Delta_w(D^i)\right)$. We start by showing that it has a finite $\|\cdot\|_w^*$ norm:

$$\begin{aligned}
\|\phi_t\|_w^* &= \sup_{\overline{\tau}\in\Pi_{i=1}^N\Delta_w(D^i)} \frac{|\phi_t(\overline{\tau})|}{\sum_{i=1}^N \int_{S\times A} w(s)\tau^i(ds\times da)}\\
&= \sup_{\overline{\tau}\in\Pi_{i=1}^N\Delta_w(D^i)} \frac{\left|\int_{S\times A}\int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s,a,\overline{\tau})\tau^1(ds\times da)\right|}{\sum_{i=1}^N \int_{S\times A} w(s)\tau^i(ds\times da)}\\
&\le \sup_{\overline{\tau}\in\Pi_{i=1}^N\Delta_w(D^i)} \frac{\int_{S\times A}\int_S \|f\|_w w(s')Q^1(ds'\mid s,a,\overline{\tau})\tau^1(ds\times da)}{\sum_{i=1}^N \int_{S\times A} w(s)\tau^i(ds\times da)}\\
&< \sup_{\overline{\tau}\in\Pi_{i=1}^N\Delta_w(D^i)} \frac{\|f\|_w\alpha\int_{S\times A} w(s)\tau^1(ds\times da)}{\int_{S\times A} w(s)\tau^1(ds\times da)} = \|f\|_w\alpha < \infty,
\end{aligned}$$

with the penultimate inequality following from part (b) of the assumption (A2') and the fact that $w \ge 1$.

We next show that ϕ_t is continuous. Let $\{\overline{\tau}_k\}_{k\ge1} \subset \Pi_{i=1}^N \Delta_w(D^i)$ be a sequence converging to $\overline{\tau} \in \Pi_{i=1}^N \Delta_w(D^i)$. Let us further define the functions $\widetilde{\phi}_t^k : S\times A \to \mathbb{R}$ where $k = 1, 2, \ldots$ and $\widetilde{\phi}_t : S\times A \to \mathbb{R}$ by

$$\begin{aligned}
\widetilde{\phi}_t^k(s,a) &:= \int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s,a,\overline{\tau}_k)\\
\widetilde{\phi}_t(s,a) &:= \int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s,a,\overline{\tau})
\end{aligned}$$

We will show that $\widetilde{\phi}_t^k$ converges continuously to $\widetilde{\phi}_t$. Let $\{s_k\}_{k\ge1} \subset S$ and $\{a_k\}_{k\ge1} \subset A$ be sequences converging to s^* and a^* respectively. Then

$$\begin{aligned}
&\left|\widetilde{\phi}_t^k(s_k,a_k) - \widetilde{\phi}_t(s^*,a^*)\right|\\
&\quad= \left|\int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s_k,a_k,\overline{\tau}_k)\right.\\
&\qquad\left.-\int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s^*,a^*,\overline{\tau})\right|\\
&\quad\le \int_{S\times A} |f(s',a')\pi_{t+1}^1(da'\mid s)|\left|Q^1(ds'\mid s_k,a_k,\overline{\tau}_k) - Q^1(ds'\mid s^*,a^*,\overline{\tau})\right|\\
&\quad\le \|f\|_w^* \int_S w(s')\left|Q^1(ds'\mid s_k,a_k,\overline{\tau}_k) - Q^1(ds'\mid s^*,a^*,\overline{\tau})\right|,
\end{aligned}$$

but the last expression goes to zero as $k \to \infty$ by (a) of assumption (A2'), proving that $\widetilde{\phi}_t^k$ converges continuously to $\widetilde{\phi}_t$. Moreover, clearly, this time by part (b) of the assumption (A2'), for each k,

$$\begin{aligned}
\left|\widetilde{\phi}_t^k(s,a)\right| &= \int_{S\times A} f(s',a')\pi_{t+1}^1(da'\mid s)Q^1(ds'\mid s,a,\overline{\tau}_k)\\
&\le \|f\|_w^* \int_S w(s'Q^1(ds'\mid s,a,\overline{\tau}_k) \le \|f\|_w^*\alpha w(s),
\end{aligned}$$

which makes the absolute values of each $\widetilde{\phi}_t^k$ bounded above by a τ^1-integrable function.

Now we can apply Theorem 3.3 in [15] to the sequence of functions $\{\widetilde{\phi_t^k}\}_{k\geq 1}$ and the sequence of measures $\{\tau_k^1\}_{k\geq 1}$ obtaining

$$\phi_t(\overline{\tau}_k) = \int_{S\times A} \widetilde{\phi_t^k}(s,a)\tau_k^1(ds\times da) \to_{k\to\infty} \int_{S\times A} \widetilde{\phi_t}(s,a)\tau^1(ds\times da) = \phi_t(\overline{\tau})$$

ending the proof of the continuity of ϕ_t, which also ends the proof that the expectation of the third term on the RHS of (9) goes to zero. □

In the third lemma, we prove an auxiliary result used to show that the rewards of the first player in the first population from using strategy vector $\widehat{\pi}$ in the n-person counterparts of the mean-field game converge to that in the mean-field limit.

Lemma 4 *Fix any $t \geq 0$ and $i \in \{1,\ldots,N\}$ and suppose that*

$$\lim_{\mathbf{n}\to\infty}\left|\int_{D^i} g_{\mathbf{n}}(s,a)\mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i},\hat{a}_{t,1}^{\mathbf{n},i}\right)(ds\times da) - \int_{D^i} g_{\mathbf{n}}(s,a)\mathcal{L}\left(\hat{s}_{t,1}^{i},\hat{a}_{t,1}^{i}\right)(ds\times da)\right| = 0$$

for any function $g_{\mathbf{n}} \in C_w(D^i)$, $\mathbf{n}\in\mathbb{N}^N$ satisfying $\sup_{\mathbf{n}\in\mathbb{N}^N}\|g_{\mathbf{n}}\|_w < \infty$. Moreover, suppose that the family $\{h_{\mathbf{n}} : \mathbf{n}\in\mathbb{N}^N\}$ of real-valued functions defined on $D^i \times \Pi_{j=1}^N \Delta_w(D^j)$ satisfying

(a) *The family $\{h_{\mathbf{n}}(s^i,a^i,\cdot),\ (s^i,a^i)\in D^i, \mathbf{n}\in\mathbb{N}^N\}$ is equicontinuous with respect to the product w-topology.*
(b) *$h_{\mathbf{n}}(\cdot,\cdot,\overline{\tau}) \in C_w(D^i)$ for any $\overline{\tau}\in\Pi_{j=1}^N\Delta_w(D^j)$ and $\mathbf{n}\in\mathbb{N}^N$.*
(c) *$\sup_{\mathbf{n}\in\mathbb{N}^N}\|h_{\mathbf{n}}(\cdot,\cdot,\overline{\tau})\|_w < \infty$ for $\overline{\tau}\in\Pi_{j=1}^N\Delta_w(D^j)$.*
(d) *The function*
$$F_t^i(\overline{\tau}) := \sup_{(s,a)\in D^i,\mathbf{n}\in\mathbb{N}^N}\left|h_{\mathbf{n}}(s,a,\overline{\tau}) - h_{\mathbf{n}}(s,a,\overline{\tau_t^*})\right|$$

defined for $\overline{\tau}\in\Pi_{j=1}^N\Delta_w(D^j)$ is real-valued and $\left\|F_t^i\right\|_w^* < \infty$.

Then

$$\lim_{\mathbf{n}\to\infty}\left|\int_{D^i\times\Pi_{j=1}^N\Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau})\mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i},\hat{a}_{t,1}^{\mathbf{n},i},\overline{e_t^{\mathbf{n}}}\right)(ds\times da\times d\overline{\tau}) - \int_{D^i\times\Pi_{j=1}^N\Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau})\mathcal{L}\left(\hat{s}_{t,1}^{i},\hat{a}_{t,1}^{i},\delta_{\overline{\tau_t^*}}\right)(ds\times da\times d\overline{\tau})\right| = 0$$

Proof Let us fix an arbitrary family $\{h_{\mathbf{n}} : \mathbf{n} \in \mathbb{N}^N\}$ satisfying the hypothesis of the lemma. Then we have

$$\Bigg| \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \widehat{\overline{e}_t^{\mathbf{n}}}\right)(ds \times da \times d\overline{\tau}) - \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}, \delta_{\overline{\tau_t^*}}\right)(ds \times da \times d\overline{\tau}) \Bigg|$$
$$\leq \Bigg| \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \widehat{\overline{e}_t^{\mathbf{n}}}\right)(ds \times da \times d\overline{\tau}) - \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \delta_{\overline{\tau_t^*}}\right)(ds \times da \times d\overline{\tau}) \Bigg|$$
$$+ \Bigg| \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \delta_{\overline{\tau_t^*}}\right)(ds \times da \times d\overline{\tau}) - \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}, \delta_{\overline{\tau_t^*}}\right)(ds \times da \times d\overline{\tau}) \Bigg|$$

The second term on the RHS can be rewritten as

$$\left| \int_{D^i} h_{\mathbf{n}}(s,a,\overline{\tau_t^*}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}\right)(ds \times da) - \int_{D^i} h_{\mathbf{n}}(s,a,\overline{\tau_t^*}) \mathcal{L}\left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}\right)(ds \times da) \right|,$$

which goes to zero as $\mathbf{n} \to \infty$ by the assumptions of the lemma. Next, we show that the same is true for the first term. Note that

$$\Bigg| \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \widehat{\overline{e}_t^{\mathbf{n}}}\right)(ds \times da \times d\overline{\tau}) - \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s,a,\overline{\tau}) \mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \delta_{\overline{\tau_t^*}}\right)(ds \times da \times d\overline{\tau}) \Bigg|$$
$$\leq \mathbb{E}\left[\mathbb{E}\left[\left|h_{\mathbf{n}}(s,a,\widehat{\overline{e}_t^{\mathbf{n}}}) - h_{\mathbf{n}}(s,a,\overline{\tau_t^*})\right| \mid (s,a) = \left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}\right)\right]\right] \leq \mathbb{E}\left[F_t^i(\widehat{\overline{e}_t^{\mathbf{n}}})\right] \quad (11)$$

We next show that F_t^i is continuous with respect to the product w-topology. Suppose $\{\overline{\tau}_k\}_{k\geq 1} \subset \Pi_{i=1}^N \Delta_w(D^i)$ is a sequence converging to $\overline{\tau} \in \Pi_{i=1}^N \Delta_w(D^i)$. Then

$$\left| F_t^i(\overline{\tau}_k) - F_t^i(\overline{\tau}) \right| = \left| \sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} \left| h_{\mathbf{n}}(s,a,\overline{\tau}_k) - h_{\mathbf{n}}(s,a,\overline{\tau_t^*}) \right| - \sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} \left| h_{\mathbf{n}}(s,a,\overline{\tau}) - h_{\mathbf{n}}(s,a,\overline{\tau_t^*}) \right| \right|$$
$$\leq \sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} \left| h_{\mathbf{n}}(s,a,\overline{\tau}_k) - h_{\mathbf{n}}(s,a,\overline{\tau}) \right| \to_{k\to\infty} 0$$

by the equicontinuity of the family $\{h_{\mathbf{n}}(s^i, a^i, \cdot),\ (s^i, a^i) \in D^i, \mathbf{n} \in \mathbb{N}^N\}$. Since by the assumption of the lemma also $\|F_t^i\|_w^* < 0$, this implies that $F_t^i \in C_w\left(\Pi_{i=1}^N \Delta_w(D^i)\right)$, by Lemma 2 the RHS in (11) goes to zero as $\mathbf{n} \to \infty$, ending the proof of the lemma ◻

In the penultimate lemma we show that the expected rewards obtained by the first player in any population in n-person counterparts of the discounted-reward mean-field game converge to his expected reward in the mean-field game for all moments of time t.

Lemma 5 *For any $t \geq 0$ and $i \in \{1, \ldots, N\}$ we have*

$$\lim_{\mathbf{n}\to\infty} \left| \mathbb{E}\left[r^i \left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \overline{e_t^{\mathbf{n}}} \right) \right] - \mathbb{E}\left[r^i \left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}, \overline{\tau_t^*} \right) \right] \right| = 0.$$

Proof We start by showing that for any family $\{g_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}^N} \subset C_w(D^i)$, satisfying $\sup_{\mathbf{n}\in\mathbb{N}^N} \|g_{\mathbf{n}}\|_w < \infty$,

$$\lim_{\mathbf{n}\to\infty} \left| \int_{D^i} g_{\mathbf{n}}(s,a)\mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}\right)(ds \times da) - \int_{D^i} g_{\mathbf{n}}(s,a)\mathcal{L}\left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}\right)(ds \times da) \right| = 0$$

We will show it by induction on t. The claim holds trivially for $t = 0$, as in this case $\mathcal{L}\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}\right) = \mathcal{L}\left(\hat{s}_{t,1}^{i}, \hat{a}_{t,1}^{i}\right) = \Pi_0^i(\pi^i, \mu_0^{*i})$. Suppose the claim holds for t and consider $t+1$. The first assumption of Lemma 4 is satisfied at time t by the induction hypothesis. Let us next define the family of real-valued functions defined on $D^i \times \Pi_{j=1}^N \Delta_w(D^j)$, $\{h_{\mathbf{n}} : \mathbf{n} \in \mathbb{N}^N\}$ using the formula

$$h_{\mathbf{n}}(s,a,\overline{\tau}) := \int_{D^i} g_{\mathbf{n}}(s',a')\widehat{\pi_{t+1,1}^i}(da' \mid s')Q^i(ds' \mid s,a,\overline{\tau}).$$

We will next show that it satisfies the assumptions of Lemma 4. Let $L := \sup_{\mathbf{n}\in\mathbb{N}^N} \|g_{\mathbf{n}}\|_w$. Then for any $\overline{\tau}, \overline{\eta} \in \Pi_{j=1}^N \Delta_w(D^j)$, we have

$$\sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} |h_{\mathbf{n}}(s,a,\overline{\tau}) - h_{\mathbf{n}}(s,a,\overline{\eta})|$$

$$\leq L \sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} \left\| Q^i(\cdot \mid s,a,\overline{\tau}) - Q^i(\cdot \mid s,a,\overline{\eta}) \right\|_w \leq L\omega_{Q^i}\left(\sum_{j=1}^N \tilde{\rho}_w(\tau^j, \eta^j) \right).$$

As by (a) of assumption (A5), $\omega_{Q^i}(\delta) \to 0$ when $\delta \to 0$, this implies that the family $\{h_{\mathbf{n}}(s,a,\cdot),\ (s,a) \in D^i, \mathbf{n} \in \mathbb{N}^N\}$ is equicontinuous. Moreover, the function

$$F_t^i(\overline{\tau}) = \sup_{(s,a)\in D^i, \mathbf{n}\in\mathbb{N}^N} \left| h_{\mathbf{n}}(s,a,\overline{\tau}) - h_{\mathbf{n}}(s,a,\overline{\tau_t^*}) \right|$$

$$\leq L\omega_{Q^i}\left(\sum_{j=1}^N \tilde{\rho}_w(\tau^j, \tau^{*j}) \right) \leq L\Omega_Q^{\overline{\tau_t^*}}(\overline{\eta}),$$

hence, it is real-valued and $\| F_t^i \|_w^* \leq L \left\| \Omega_Q^{\overline{\tau_t^*}} \right\|_w^* < \infty$, again by (a) of assumption (A5).

Next, note that for any $\overline{\tau} \in \Pi_{j=1}^N \Delta_w(D^j)$, $\sup_{\mathbf{n}\in\mathbb{N}^N} \|h_{\mathbf{n}}(\cdot,\cdot,\overline{\tau})\|_w \leq \alpha L$ by (b) of assumption (A2'). Finally, we need to check that for any $\overline{\tau}$, $h_{\mathbf{n}}(\cdot,\cdot,\overline{\tau})$ is continuous. To this end let us first define the function

$$l(s) := \int_A g_{\mathbf{n}}(s,a)\widehat{\pi_{t+1,1}^i}(da \mid s).$$

Clearly, l is a continuous function, given that for any sequence $\{s^k\}_{k\geq 1} \subset S$ converging to some s^*, we have

$$l(s^k) = \int_A g_{\mathbf{n}}(s^k,a)\widehat{\pi_{t+1,1}^i}(da \mid s^k) \to_{k\to\infty} \int_A g_{\mathbf{n}}(s^*,a)\widehat{\pi_{t+1,1}^i}(da \mid s^*) = l(s^*)$$

by Theorem 3.3 in [15] as $g_{\mathbf{n}}$ is continuous and $\widehat{\pi^i_{t,1}}$ is weakly continuous by the assumption we have made about the strategy $\widehat{\pi}$ (remember also that A is compact, so $g_{\mathbf{n}}$ is bounded on the set $\left(\{s^k\}_{k\geq 1} \cup \{s^*\}\right) \times A$). Moreover, $\|l\|_w \leq \|g_{\mathbf{n}}\|_w \leq L$.

Next, let us take a sequence $\{(s^k, a^k)\}_{k\geq 1} \subset D^i$ converging to some (s^*, a^*). Clearly,

$$h_{\mathbf{n}}(s^k, a^k, \overline{\tau}) = \int_{S^i} l(s')Q^i(ds' \mid s^k, a^k, \overline{\tau})$$

which, again by Theorem 3.3 in [15], converges to

$$h_{\mathbf{n}}(s^*, a^*, \overline{\tau}) = \int_{S^i} l(s')Q^i(ds' \mid s^*, a^*, \overline{\tau}),$$

as $l \in C_w(S^i)$ and Q^i is weakly continuous by assumption (A2').

As we have shown that the family $\{h_{\mathbf{n}} : \ \mathbf{n} \in \mathbb{N}^N\}$ satisfies all the assumptions given in Lemma 4, we can conclude as follows:

$$\begin{aligned}
&\lim_{\mathbf{n}\to\infty} \left| \int_{D^i} g_{\mathbf{n}}(s, a)\mathcal{L}\left(\hat{s}^{\mathbf{n},i}_{t+1,1}, \hat{a}^{\mathbf{n},i}_{t+1,1}\right)(ds \times da) \right. \\
&\quad \left. - \int_{D^i} g_{\mathbf{n}}(s, a)\mathcal{L}\left(\hat{s}^{i}_{t+1,1}, \hat{a}^{i}_{t+1,1}\right)(ds \times da) \right| \\
&= \lim_{\mathbf{n}\to\infty} \left| \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s, a, \overline{\tau})\mathcal{L}\left(\hat{s}^{\mathbf{n},i}_{t,1}, \hat{a}^{\mathbf{n},i}_{t,1}, \overline{e^{\mathbf{n}}_t}\right)(ds \times da \times d\overline{\tau}) \right. \\
&\quad \left. - \int_{D^i \times \Pi_{j=1}^N \Delta_w(D^j)} h_{\mathbf{n}}(s, a, \overline{\tau})\mathcal{L}\left(\hat{s}^{i}_{t,1}, \hat{a}^{i}_{t,1}, \delta_{\overline{\tau^*_t}}\right)(ds \times da \times d\overline{\tau}) \right| = 0,
\end{aligned}$$

showing the induction hypothesis.

The last step of the proof is showing that the function r^i satisfies all the assumptions of Lemma 4 (when taking $h_{\mathbf{n}} \equiv r^i$ for $\mathbf{n} \in \mathbb{N}^N$). Obviously, $r^i(\cdot, \cdot, \overline{\tau}) \in C_w(D^i)$ for any $\overline{\tau} \in \Pi_{j=1}^N \Delta_w(D^j)$ by assumption (A1'). Then for any $\overline{\tau}, \overline{\eta} \in \Pi_{j=1}^N \Delta_w(D^j)$, we have

$$\sup_{(s,a)\in D^i} \left| r^i(s, a, \overline{\tau}) - r^i(s, a, \overline{\eta}) \right| \leq \omega_{r^i}\left(\sum_{j=1}^N \tilde{\rho}_w(\tau^j, \eta^j)\right).$$

As by (a) of assumption (A5), $\omega_{r^i}(\delta) \to 0$ when $\delta \to 0$, this implies that the family $\{r^i(s, a, \cdot),\ (s, a) \in D^i\}$ is equicontinuous. Moreover, the function

$$\widetilde{F}^i_t(\overline{\tau}) := \sup_{(s,a)\in D^i} \left| r^i(s, a, \overline{\tau}) - r^i(s, a, \overline{\tau^*_t}) \right| \leq \omega_{r^i}\left(\sum_{j=1}^N \tilde{\rho}_w(\tau^j, \tau^{*j})\right) \leq \Omega_r^{\overline{\tau^*_t}}(\overline{\eta}),$$

hence, it is real-valued and $\|\widetilde{F}^i_t\|^*_w \leq \left\|\Omega_Q^{\overline{\tau^*_t}}\right\|^*_w < \infty$, again by (a) of assumption (A5). Now we can apply Lemma 4 to r^i, obtaining the thesis of the lemma. □

The last lemma can be treated as a counterpart of Theorem 2.3 in [11] tailored to our model. It shows that in our considerations we can restrict ourselves to the weakly continuous deviations from the mean-field equilibrium strategy.

Lemma 6 *For any i, under the assumptions of Theorem 1,*

$$\sup_{\widehat{\pi_1^i}\in\mathcal{M}^i} \mathbb{E}\left[J_{\beta,n}^{1,i}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi_1^i}])\mid \overline{s_0}\sim\overline{\mu}_0^*\right]$$

can be obtained using policies $\widehat{\pi_1^i}$ such that $\widehat{\pi_{1,t}^i}$ is weakly continuous for any time $t \geq 0$.

Proof Without loss of generality, we may show the result only for $i = 1$. Let us choose an arbitrary policy $\widetilde{\pi_1^1} \in \mathcal{M}^1$. We will show that for any $\varepsilon > 0$ there exists $\widehat{\pi_1^1} \in \mathcal{M}^1$ which is weakly continuous and satisfies

$$\mathbb{E}\left[J_{\beta,n}^{1,1}(\overline{s_0},[\overline{\pi}_{-1,1},\widehat{\pi_1^1}])\mid \overline{s_0}\sim\overline{\mu}_0^*\right] \geq \mathbb{E}\left[J_{\beta,n}^{1,1}(\overline{s_0},[\overline{\pi}_{-1,1},\widetilde{\pi_1^1}])\mid \overline{s_0}\sim\overline{\mu}_0^*\right] - \varepsilon, \tag{12}$$

which will complete the proof.

We start by noting that the problem the first player in first population is facing is that of finding an optimal policy in a Markov decision process with state space $\Pi_{j=1}^N (S^j)^{n_j}$, time-dependent transition probability

$$\widetilde{Q}_t(d\overline{x}\mid \overline{s},a_1^1) = \int_{A^1(s_2^1)}\!\!\ldots\int_{A^1(s_{n_1}^1)}\ldots\int_{A^N(s_1^N)}\!\!\ldots\int_{A^N(s_{n_N}^N)} \Pi_{j=1}^N\Pi_{k=1}^{n_j} Q^j\left(dx_k^j\mid s_k^j,a_k^j,\overline{\tau}(\overline{s},\overline{a})\right)$$
$$\Pi_{k=2}^{n_1}\pi_{t,k}^j(da_k^j\mid s_k^j)\Pi_{j=2}^N\Pi_{k=1}^{n_j}\pi_{t,k}^j(da_k^j\mid s_k^j)$$

and one-stage reward

$$\widetilde{r}_t(\overline{s},a_1^1) = \int_{A^1(s_2^1)}\!\!\ldots\int_{A^1(s_{n_1}^1)}\ldots\int_{A^N(s_1^N)}\!\!\ldots\int_{A^N(s_{n_N}^N)} r^i\left(s_1^1,a_1^1,\overline{\tau}(\overline{s},\overline{a})\right)$$
$$\Pi_{k=2}^{n_1}\pi_{t,k}^j(da_k^j\mid s_k^j)\Pi_{j=2}^N\Pi_{k=1}^{n_j}\pi_{t,k}^j(da_k^j\mid s_k^j).$$

By Lusin's theorem (see Theorem 7.5.2 in [4]), for any $\delta > 0$, there exists a closed set $F_0^\delta \in S^1$ such that $\mu_0^{*1}(F_0^\delta) < \delta$ and $\widetilde{\pi_{0,1}^1}$ is weakly continuous on F_0^δ. As $\Delta(A)$ is a convex subset of a locally convex vector space of finite signed measures on A, by Dugundji's extension theorem (see Theorem 7.4 in [5]), we can extend $\widetilde{\pi_{0,1}^1}$ limited to F_0^δ continuously to S^1. Let $\widetilde{\pi_{0,1}^{1,\delta}}$ denote this extension. We then apply the same method to $\widetilde{\pi_{1,1}^1}$, that is, we define the measure $\widetilde{\mu}_1$ on S^1 with the formula

$$\widetilde{\mu}_1(B) := \int_B\int_{(S^1)^{n_1-1}}\ldots\int_{(S^N)^{n_N}}\int_A \widetilde{Q}_0\left(B\times(S^1)^{n_1-1}\times\Pi_{j=2}^N(S^j)^{n_j}\mid\overline{s},a_1^1\right)$$
$$\widetilde{\pi_{0,1}^1}(da_1^1\mid s_1^1)\mu_0^{*1}(ds_1^1)\ldots\mu_0^{*1}(ds_{n_1}^1)\ldots\mu_0^{*N}(ds_1^N)\ldots\mu_0^{*1}(ds_{n_N}^N)$$

and construct a continuous $\widetilde{\pi_{1,1}^{1,\delta}}$ that agrees with $\widetilde{\pi_{1,1}^1}$ on a closed subset F_1^δ of S^1 satisfying $\widetilde{\mu}_1(F_1^\delta) < \delta$. We continue in the same manner until time t^*, constructing measures $\widetilde{\mu}_t$ and weakly continuous stochastic kernels $\widetilde{\pi_{t,1}^{1,\delta}}$ for $t = 2,\ldots,t^*$ and define $\widehat{\pi_1^1(\delta,t^*)}$ as a Markov strategy for the first player in the first population of the form $\left(\widetilde{\pi_{0,1}^{1,\delta}},\widetilde{\pi_{1,1}^{1,\delta}},\ldots,\widetilde{\pi_{t^*,1}^{1,\delta}},\pi_{t^*+1,1}^1,\ldots\right)$ (remember that the kernel $\pi_{t,1}^1$ is weakly continuous for any t by the assumption of the theorem). Now we can conclude as follows:

$$\begin{aligned}
&\left|\mathbb{E}\left[J^{1,1}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-1,1},\widehat{\pi^1_1(\delta,t^*)}])\mid\overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^{1,1}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-1,1},\widetilde{\pi^1_1}])\mid\overline{s_0}\sim\overline{\mu^*_0}\right]\right|\\
&=\left|\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widehat{\pi^1_1(\delta,t^*)}]}\sum_{t=0}^{\infty}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})-\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widetilde{\pi^1_1}]}\sum_{t=0}^{\infty}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})\right|\\
&\le\left|\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widehat{\pi^1_1(\delta,t^*)}]}\sum_{t=0}^{t^*}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})-\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widetilde{\pi^1_1}]}\sum_{t=0}^{t^*}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})\right|\\
&\quad+\left|\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widehat{\pi^1_1(\delta,t^*)}]}\sum_{t=t^*+1}^{\infty}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})-\mathbb{E}^{\overline{\mu^*_0},Q_n,[\overline{\pi}_{-1,1},\widetilde{\pi^1_1}]}\sum_{t=t^*+1}^{\infty}\beta^t r^{1,1}_n(\overline{s_t},\overline{a_t})\right|
\end{aligned}\tag{13}$$

Note that by assumption (A1'), (b) of assumption (A2') and the definition of $r^{1,1}_n$, for any t we have

$$R\ge\mathbb{E}^{\overline{\mu^*_0},Q_n,\cdot}\left[r^{1,1}_n(\overline{s_t},\overline{a_t})\right]\ge-R\gamma^t\mathbb{E}^{\overline{\mu^*_0},Q_n,\cdot}\left[w(s^1_{t,1})\right]\ge-R\gamma^t\alpha^t M,\tag{14}$$

regardless of the strategy used. Therefore, the second term on the RHS of (13) can be bounded above by

$$\sum_{t=t^*+1}^{\infty}\beta^t\left(R+R\gamma^t\alpha^t M\right)=\beta^{t^*+1}\left(\frac{R}{1-\beta}+\frac{RM\alpha^{t^*}\gamma^{t^*}}{1-\alpha\beta\gamma}\right),$$

which goes to 0 as $t^*\to\infty$. Obviously, this implies that there exists a value of t^*, call it $\widehat{t^*}$, for which the second term on the RHS of (13) is smaller than $\frac{\varepsilon}{2}$.

Next, note that, by (14) and (b) of assumption (A2'), the first term on the RHS of (13) is bounded above by

$$\begin{aligned}
&\sum_{t=0}^{t^*}\beta^t\left|R\gamma^t\alpha^t\int_{F^\delta_0}w(s^1_{0,1})\mu^{*1}_0(ds^1_{0,1})+R\delta\right|\\
&\quad+\sum_{t=1}^{t^*}\beta^t\left|R\gamma^t\alpha^{t-1}\int_{F^\delta_0}w(s^1_{1,1})\widetilde{\mu}_1(ds^1_{1,1})+R\delta\right|\\
&\quad+\ldots+\sum_{t=t^*}^{t^*}\beta^t\left|R\gamma^t\alpha^{t-t^*}\int_{F^\delta_0}w(s^1_{t^*,1})\widetilde{\mu}_{t^*}(ds^1_{t^*,1})+R\delta\right|
\end{aligned}$$

This sum can be made arbitrarily small, say smaller than $\frac{\varepsilon}{2}$ for $t^*=\widehat{t^*}$ by taking appropriate $\delta=\delta^*$.

This however implies that $\widehat{\pi^1_k}=\widehat{\pi^1_1(\delta^*,\widehat{t^*})}$ satisfies (12). □

Proof of Theorem 1 Take $\varepsilon>0$ and note that by Lemma 6 for each population i there exists a weakly continuous policy $\widehat{\pi^i_1}\in\mathcal{M}^i$ such that

$$\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi^i_1}])\mid\overline{s_0}\sim\overline{\mu^*_0}\right]\ge\sup_{\widetilde{\pi^i_1}\in\mathcal{M}^i}\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widetilde{\pi^i_1}])\mid\overline{s_0}\sim\overline{\mu^*_0}\right]-\frac{\varepsilon}{5}.\tag{15}$$

Next note that for any $t^* \geq 0$ we have

$$\begin{aligned}&\left|\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi^i_1}])\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\widehat{\pi^i_1},\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\right|\\&\leq\sum_{t=0}^{t^*}\beta^t\left|\mathbb{E}\left[r^i\left(\hat{s}^{\mathbf{n},i}_{t,1},\hat{a}^{\mathbf{n},i}_{t,1},\overline{\widehat{e^{\mathbf{n}}_t}}\right)\right]-\mathbb{E}\left[r^i\left(\hat{s}^i_{t,1},\hat{a}^i_{t,1},\overline{\tau^*_t}\right)\right]\right|\\&\quad+\left|\mathbb{E}\left[\sum_{t=t^*+1}^{\infty}\beta^t\left(r^i\left(\hat{s}^{\mathbf{n},i}_{t,1},\hat{a}^{\mathbf{n},i}_{t,1},\overline{\widehat{e^{\mathbf{n}}_t}}\right)-r^i\left(\hat{s}^i_{t,1},\hat{a}^i_{t,1},\overline{\tau^*_t}\right)\right)\right]\right|\end{aligned}$$

The first t^* terms on the RHS go to zero as $\mathbf{n}\to\infty$ by Lemma 5, while (14) implies that the last term can be bounded above by

$$\sum_{t=t^*+1}^{\infty}\beta^t\left(R+R\gamma^t\alpha^tM\right)=\beta^{t^*+1}\left(\frac{R}{1-\beta}+\frac{RM\alpha^{t^*+1}\gamma^{t^*+1}}{1-\alpha\beta\gamma}\right)\to_{t^*\to\infty}0.$$

Hence, we can fix t^* such that $\beta^{t^*+1}\left(\frac{R}{1-\beta}+\frac{RM\alpha^{t^*+1}\gamma^{t^*+1}}{1-\alpha\beta\gamma}\right)<\frac{\varepsilon}{5}$ and $n^i_1(\varepsilon),\ldots,n^i_N(\varepsilon)$ such that for $n_1\geq n^i_1(\varepsilon),\ldots,n_N\geq n^i_N(\varepsilon)$, we have

$$\left|\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi^i_1}])\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\widehat{\pi^i_1},\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\right|<\frac{2\varepsilon}{5}. \tag{16}$$

Using similar reasoning, we may find $\widehat{n^i_1}(\varepsilon),\ldots,\widehat{n^i_N}(\varepsilon)$ such that for $n_1\geq\widehat{n^i_1}(\varepsilon),\ldots,n_N\geq\widehat{n^i_N}(\varepsilon)$,

$$\left|\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},\overline{\pi})\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\pi^i_1,\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\right|<\frac{2\varepsilon}{5}. \tag{17}$$

If we take $n_i(\varepsilon):=\max\{\max\{n^j_i(\varepsilon),j=1,\ldots,N\},\max\{\widehat{n^j_i}(\varepsilon),j=1,\ldots,N\}\}$, $i=1,\ldots,N$, the definition of the Markov mean-field equilibrium, (16) and (17) imply that

$$\begin{aligned}&\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},\overline{\pi})\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi^i_1}])\mid \overline{s_0}\sim\overline{\mu^*_0}\right]\\&\geq\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\pi^i_1,\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\widehat{\pi^i_1},\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\\&\quad-\left|\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widehat{\pi^i_1}])\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\widehat{\pi^i_1},\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\right|\\&\quad-\left|\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},\overline{\pi})\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\mathbb{E}\left[J^i_\beta(s^i_0,\overline{\mu}^*_0,\pi^i_1,\overline{\pi})\mid s^i_0\sim\mu^{*i}_0\right]\right|>-\frac{4\varepsilon}{5}\end{aligned}$$

for $n_1\geq n_1(\varepsilon)$,..., $n_1\geq n_1(\varepsilon)$. Combining it with (15) we get that

$$\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},\overline{\pi})\mid \overline{s_0}\sim\overline{\mu^*_0}\right]>\sup_{\widetilde{\pi^i_1}\in\mathcal{M}^i}\mathbb{E}\left[J^{1,i}_{\beta,n}(\overline{s_0},[\overline{\pi}_{-i,1},\widetilde{\pi^i_1}])\mid \overline{s_0}\sim\overline{\mu^*_0}\right]-\varepsilon$$

for any $i\in\{1,\ldots,N\}$ and $n_1\geq n_1(\varepsilon)$,..., $n_1\geq n_1(\varepsilon)$. As all the players within each population are symmetric, this implies that the profile of strategies $\overline{\pi}$ is an ε-equilibrium in the n-person counterpart of the discounted-payoff mean-field game in this case. □

5.2 Results for the Total Payoff Case

In the remaining results we address the n-person counterparts of total-payoff game.

Theorem 7 *Suppose assumptions (A1"), (A2"), (A3), (A4") and (A5) hold and suppose* $\overline{\pi}$ *and* $(\overline{\mu}_0^*, \overline{\mu}_1^*, \ldots)$ *form a Markov mean-field equilibrium in the multi-population discrete-time mean-field game existing by Theorem 8 in [18]. If in addition, for each* $t \geq 0$ *and* $i = 1, \ldots, N$, π_t^i *is weakly continuous, then for any* $\varepsilon > 0$ *and any* $T \geq 0$ *there exist positive integers* $n_i(\varepsilon, T)$, $i = 1, \ldots, N$ *such that the vector of strategies where each player from population* i *uses policy* π^i *is an* (ε, T)*-Markov–Nash equilibrium in any* n*-person stochastic counterpart of the total-payoff mean-field game if* $n^i \geq n^i(\varepsilon, T)$, $i = 1, \ldots, N$.

Remark 2 As in the case of discounted payoff, stationary mean-field equilibrium existing according to Theorem 5 in [18] is a specific case of Markov mean-field equilibria with stationarity condition imposed on global states of the game at subsequent stages. Hence, the result provided by Theorem 7 holds in this case as well.

Before we pass to the actual proof of Theorem 7, we present an auxiliary result that can be seen as a variant of Lemma 6 for the total payoff game.

Lemma 8 *For any* i *and any* $t_0 \in \mathbb{N}$ *under the assumptions of Theorem 7,*

$$\sup_{\widehat{\pi_1^i} \in \mathcal{M}^i} \mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [{}^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*} \right]$$

can be attained using policies $\widehat{\pi_1^i}$ *such that* $\widehat{\pi_{1,t}^i}$ *is weakly continuous for any time* $t \geq 0$. *Moreover, these policies do not depend on* t_0 *as long as* $t_0 \leq T$ *for some fixed* $T \in \mathbb{N}$.

Proof Without loss of generality we may only consider $i = 1$. As in the case of Lemma 6 what we need to prove is that for an arbitrary policy $\widetilde{\pi_1^1} \in \mathcal{M}^1$ and any $\varepsilon > 0$ there exists $\widehat{\pi_1^1} \in \mathcal{M}^1$ which is weakly continuous and satisfies for any fixed $t_0 \leq T$

$$\mathbb{E}\left[J_{*n}^{1,1}(\overline{\mu_{t_0}}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1}]) \mid \overline{s_0} \sim \overline{\mu_0^*} \right] \geq \mathbb{E}\left[J_{*n}^{1,1}(\overline{\mu_{t_0}}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]) \mid \overline{s_0} \sim \overline{\mu_0^*} \right] - \varepsilon. \tag{18}$$

The beginning of the proof is the same as for Lemma 6: we construct a Markov strategy for the first player in the first population $\widehat{\pi_1^1(\delta, t^*)}$ (with $t^* > T$) of the form $\left(\widetilde{\pi_{0,1}^{1,\delta}}, \widetilde{\pi_{1,1}^{1,\delta}}, \ldots, \widetilde{\pi_{t^*,1}^{1,\delta}}, \pi_{t^*+1,1}^1, \ldots\right)$, where each $\widetilde{\pi_{t,1}^{1,\delta}}$ is weakly continuous and agrees with $\widetilde{\pi_{t,1}^1}$ with probability $1 - \delta$.

Next, we define modified (non-time homogeneous) transition probability $Q_n^{*t_0}$ as

$$Q_{n,t}^{*t_0}(\cdot \mid s, a, \overline{\tau}) := \begin{cases} Q_n(\cdot \mid \overline{s}, \overline{a}), & \text{if } s_1^1 \neq s^* \text{ or } t < t_0 \\ \delta_{(s^*)^n}, & \text{if } s = s^* \text{ and } t \geq t_0 \end{cases}$$

Let $\left(Q_n^{*t_0}\right)^t (\cdot \mid \overline{s}, \overline{\sigma})$ denote the transition in t steps when the initial state of the n-person game is $\overline{s}$ and the players use Markov strategy vector $\overline{\sigma}$. It can be checked that under assumption (A4"), $Q_n^{*t_0}$ satisfies

$$\lim_{t^* \to \infty} \sup_{\substack{\overline{\sigma} \in \Pi_{j=1}^N (\mathcal{M}^j)^{n_j}, \\ (\overline{s}_0, \overline{a}_0, \overline{s}_1, \overline{a}_1, \ldots) \in \Pi_{t=0}^\infty \Pi_{j=1}^N (D^j)^{n_j}}} \left\| \sum_{t=t^*+1}^{\infty} \int_{S^1 \setminus \{s^*\}} w(x_1^1) \alpha^{-t} \left(Q_n^{*t_0}\right)^t (d\overline{x} \mid \overline{s}, \overline{\sigma}) \right\|_w = 0,$$

which can shortly be written as

$$\lim_{t^* \to \infty} L_{t^*} = 0 \tag{19}$$

with L_{t^*} denoting the supremum under the limit.

Now we can proceed as follows:

$$\left|\mathbb{E}\left[J_{*n}^{1,1}(\overline{\mu_{t_0}}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1(\delta, t^*)}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_{*n}^{1,1}(\overline{\mu_{t_0}}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right]\right|$$
$$= \left|\mathbb{E}^{\overline{\mu_0}^*, Q_n, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1}(\delta,t^*)]} \sum_{t=t_0}^{\mathcal{T}_1^1-1} r_n^{1,1}(\overline{s_t}, \overline{a_t}) - \mathbb{E}^{\overline{\mu_0}^*, Q_n, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]} \sum_{t=t_0}^{\mathcal{T}_1^1-1} r_n^{1,1}(\overline{s_t}, \overline{a_t})\right|$$
$$= \left|\mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1}(\delta,t^*)]} \sum_{t=t_0}^{\infty} r_n^{1,1}(\overline{s_t}, \overline{a_t}) - \mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]} \sum_{t=t_0}^{\infty} r_n^{1,1}(\overline{s_t}, \overline{a_t})\right|$$
$$\leq \left|\mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1}(\delta,t^*)]} \sum_{t=t_0}^{t^*} r_n^{1,1}(\overline{s_t}, \overline{a_t}) - \mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]} \sum_{t=t_0}^{t^*} r_n^{1,1}(\overline{s_t}, \overline{a_t})\right|$$
$$+ \left|\mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widehat{\pi_1^1}(\delta,t^*)]} \sum_{t=t^*+1}^{\infty} r_n^{1,1}(\overline{s_t}, \overline{a_t}) - \mathbb{E}^{\overline{\mu_0}^*, Q_n^{*t_0}, [{}^{t_0}\overline{\pi}_{-1,1}, {}^{t_0}\widetilde{\pi_1^1}]} \sum_{t=t^*+1}^{\infty} r_n^{1,1}(\overline{s_t}, \overline{a_t})\right|$$

As assumptions (A1") and (A2") are stronger than (A1') and (A2'), bounds given in (14) still hold here. Hence, the first term on the RHS can be bounded above by

$$\sum_{t=\max\{0,t_0\}}^{t^*} \left|R\gamma^t\alpha^t \int_{F_0^\delta} w(s_{0,1}^1)\mu_0^{*1}(ds_{0,1}^1) + R\delta\right|$$
$$+ \sum_{t=\max\{1,t_0\}}^{t^*} \left|R\gamma^t\alpha^{t-1} \int_{F_0^\delta} w(s_{1,1}^1)\widetilde{\mu}_1(ds_{1,1}^1) + R\delta\right|$$
$$+ \ldots + \sum_{t=t^*}^{t^*} \left|R\gamma^t\alpha^{t-t^*} \int_{F_0^\delta} w(s_{t^*,1}^1)\widetilde{\mu}_{t^*}(ds_{t^*,1}^1) + R\delta\right|.$$

As far as the second term is concerned, (14) and the fact that $r_n^{1,1}(\overline{s}, \overline{a})$ equals zero whenever $s_1^1 = s^*$ imply that it can be bounded above by

$$\mathbb{E}\left[\sup_{\overline{\sigma}\in\Pi_{j=1}^N(\mathcal{M}^j)^{n_j}} \sum_{t=t^*+1}^{\infty} R\gamma^t\alpha^t \int_{S^1\setminus\{s^*\}} w(x_1^1)\alpha^{-t}\left(Q_n^{*t_0}\right)^t (d\overline{x} \mid \overline{s}, \overline{\sigma})\right.$$
$$\left. + \sup_{\overline{\sigma}\in\Pi_{j=1}^N(\mathcal{M}^j)^{n_j}} \sum_{t=t^*+1}^{\infty} R\left(Q_n^{*t_0}\right)^t \left(S^1 \setminus \{s^*\} \mid \overline{s}, \overline{\sigma}\right) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] \leq 2RML_{t^*}$$

with the last inequality following from the definition of L_{t^*}, (A1") (in particular the fact that $\alpha\gamma < 1$) and the inequality $w \geq 1$.

Both boundaries can be made arbitrarily small by taking t^* big enough (by (19)) in the second case and δ small enough in the first one. In particular, if both are less than $\frac{\varepsilon}{2}$, $\widehat{\pi_k^1} = \pi_1^1\widehat{(\delta^*, \widehat{t^*})}$ satisfies (18). As boundaries do not depend on t_0, the last statement of the lemma has also been proved. ◻

Proof of Theorem 7 Take $\varepsilon > 0$ and $T \in \mathbb{N}$. Note that by Lemma 8 for each population i there exists a weakly continuous policy $\widehat{\pi_1^i} \in \mathcal{M}^i$ such that

$$\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] \geq \sup_{\widetilde{\pi_1^i} \in \mathcal{M}^i} \mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widetilde{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \frac{\varepsilon}{5}. \tag{20}$$

for any $t_0 \leq T$. Next note that for any $t^* \geq T$ we have

$$\left|\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\widehat{\pi_1^i}, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]\right|$$
$$\leq \left|\mathbb{E}\left[\sum_{t=t_0}^{\min\{\mathcal{T}_1^i - 1, t^*\}} \left(r^i\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \overline{e_t^{\mathbf{n}}}\right) - r^i\left(\hat{s}_{t,1}^i, \hat{a}_{t,1}^i, \overline{\tau_t^*}\right)\right)\right]\right|$$
$$+ \left|\mathbb{E}\left[\sum_{t=t^*+1}^{\mathcal{T}_1^i} \left(r^i\left(\hat{s}_{t,1}^{\mathbf{n},i}, \hat{a}_{t,1}^{\mathbf{n},i}, \overline{e_t^{\mathbf{n}}}\right) - r^i\left(\hat{s}_{t,1}^i, \hat{a}_{t,1}^i, \overline{\tau_t^*}\right)\right)\right]\right|$$

The first term on the RHS goes to zero as $\mathbf{n} \to \infty$ by Lemma 5. As far as the second term is concerned, it can be bounded above by $2RML_{t^*}$ in a similar way as in the proof of Lemma 8. By (19) this can be made arbitrarily small by taking t^* big enough. In particular, we can find t^* such that $2RML_{t^*} < \frac{\varepsilon}{5}$ and $n_1^i(\varepsilon, T), \ldots, n_N^i(\varepsilon, T)$ such that for $n_1 \geq n_1^i(\varepsilon, T), \ldots, n_N \geq n_N^i(\varepsilon, T)$, for each $t_0 \geq T$ we have

$$\left|\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\widehat{\pi_1^i}, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]\right| < \frac{2\varepsilon}{5}. \tag{21}$$

Similarly we find $\widehat{n_1^i}(\varepsilon, T), \ldots, \widehat{n_N^i}(\varepsilon, T)$ such that for $n_1 \geq \widehat{n_1^i}(\varepsilon, T), \ldots, n_N \geq \widehat{n_N^i}(\varepsilon, T)$ and any $t_0 \leq T$,

$$\left|\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, {}^{t_0}\overline{\pi}) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\pi_1^i, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]\right| < \frac{2\varepsilon}{5}. \tag{22}$$

If we take $n_i(\varepsilon, T) := \max\{\max\{n_i^j(\varepsilon, T), j = 1, \ldots, N\}, \max\{\widehat{n_i^j}(\varepsilon, T), j = 1, \ldots, N\}\}$, $i = 1, \ldots, N$, the definition of the Markov mean-field equilibrium, (21) and (22) imply that

$$\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, {}^{t_0}\overline{\pi}) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right]$$
$$\geq \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\pi_1^i, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\widehat{\pi_1^i}, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]$$
$$- \left|\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widehat{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\widehat{\pi_1^i}, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]\right|$$
$$- \left|\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, {}^{t_0}\overline{\pi}) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \mathbb{E}\left[J_*^i(s_{t_0}^i, \overline{\mu}_{t_0}^*, {}^{t_0}\pi_1^i, {}^{t_0}\overline{\pi}) \mid s_0^i \sim \mu_0^{*i}\right]\right| > -\frac{4\varepsilon}{5}$$

for $n_1 \geq n_1(\varepsilon, T)$,..., $n_1 \geq n_1(\varepsilon, T)$ and $t_0 \leq T$. Combining it with (20) we get that

$$\mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, {}^{t_0}\overline{\pi}) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] > \sup_{\widetilde{\pi_1^i} \in \mathcal{M}^i} \mathbb{E}\left[J_{*n}^{1,i}(\overline{\mu_{t_0}}, [^{t_0}\overline{\pi}_{-i,1}, {}^{t_0}\widetilde{\pi_1^i}]) \mid \overline{s_0} \sim \overline{\mu_0^*}\right] - \varepsilon$$

for any $i \in \{1, \ldots, N\}$, $n_1 \geq n_1(\varepsilon, T)$,..., $n_1 \geq n_1(\varepsilon, T)$ and $t_0 \leq T$. As all the players within each population are symmetric, this implies that the profile of strategies $\overline{\pi}$ is an

(ε, T)-equilibrium in the n-person counterpart of the total-payoff mean-field game in this case. □

6 Concluding Remarks

The paper is the continuation of our previous article [18], where we have presented the conditions under which multiple-population discrete-time mean-field games admit Markov (or stationary) equilibria. These results were presented for two payoff criteria: β-discounted payoff and total expected payoff. In this article, we have presented the theorems showing that under some rather unrestrictive assumptions equilibria obtained in the mean-field models are approximate equilibria in their n-person counterparts when n is large enough. All of them are presented for both payoffs considered. As games with total payoff have only been studied in finite state space case, the approximation results presented here also extend those for total-payoff mean-field games with a single population. The article is a part of ongoing research on discrete-time mean-field games with multiple populations of players. The next step should be extending the results presented in this paper to the case of long-run average reward. This is an especially interesting case as standard ergodicity assumptions applied in this kind of models to show the existence of an equilibrium do not translate well to the case with multiple populations.

Author Contributions The single author did everything.

Funding This work was supported by the NCN Grant no 2016/23/B/ST1/00425.

Declarations

Conflict of interest The author declares that he has no Conflict of interest.

References

1. Aliprantis CD, Border KC (1999) Infinite dimensional analysis. A Hitchhiker's guide. Springer-Verlag, Berlin Heidelberg
2. Bertsekas DP, Shreve SE (1978) Stochastic optimal control: the discrete time case. Academic Press, New York
3. Bolley F. Applications du transport optimal à des problèmes de limites de champ moyen. Thèse de doctorat, Ecole Normale Supérieure de Lyon. Available at https://theses.hal.science/tel-00011462
4. Dudley RM (2004) Real analysis and probability. Cambridge University Press, Cambridge
5. Granas A, Dugundji J (2003) Fixed point theory. Springer, Berlin
6. Haurie A, Krawczyk JB, Zaccour G (2012) Games and dynamic games. World Scientific, Singapore
7. Hernández-Lerma O, Lasserre JB (1996) Discrete-time Markov control processes: basic optimality criteria. Springer, Berlin
8. Hernández-Lerma O, Lasserre JB (1999) Further topics on discrete-time Markov control processes. Springer, Berlin

9. Parthasarathy KR (1967) Probability measures on metric spaces. AMS Bookstore
10. Saldi N (2020) Discrete-time average-cost mean-field games on Polish spaces. Turk J Math 44:463–480
11. Saldi N, Başar T, Raginsky M (2018) Markov-Nash equilibria in mean-field games with discounted cost. SIAM J Control Optim 56(6):4256–4287
12. Saldi N, Başar T, Raginsky M (2019) Approximate Nash equilibria in partially observed stochastic games with mean-field interactions. Math Oper Res 44(3):1006–1033
13. Saldi N, Başar T, Raginsky M (2020) Approximate Markov-Nash equilibria for discrete-time risk-sensitive mean-field games. Math Oper Res 45(4):1596–1620
14. Saldi N, Başar T, Raginsky M (2022) Partially observed discrete-time risk-sensitive mean field games. Dyn Games Appl. https://doi.org/10.1007/s13235-022-00453-z
15. Serfozo R (1982) Convergence of Lebesgue integrals with varying measures. Sankhya Indian J Stat Ser A 44(3):380–402
16. Więcek P, Altman E (2015) Stationary anonymous sequential games with undiscounted rewards. J Optim Theory Appl 166(2):686–710
17. Więcek P (2020) Discrete-time ergodic mean field games with average reward on compact spaces. Dyn Games Appl 10:222–256
18. Więcek P (2024) Multiple-population discrete-time mean field games with discounted and total payoffs: the existence of equilibria. Dyn Games Appl. https://doi.org/10.1007/s13235-024-00567-6