SIAM J. CONTROL OPTIM.
Vol. 62, No. 3, pp. 1437–1464

NASH EQUILIBRIA FOR EXCHANGEABLE TEAM-AGAINST-TEAM GAMES, THEIR MEAN-FIELD LIMIT, AND THE ROLE OF COMMON RANDOMNESS*

SINA SANJARI†, NACI SALDI‡, AND SERDAR YÜKSEL§

Abstract. We study stochastic exchangeable games among a finite number of teams consisting of a large but finite number of decision makers as well as their mean-field limit with infinite number of decision makers in each team. For this class of games within static and dynamic settings, we introduce sets of randomized policies under various decentralized information structures with privately independent or common randomness for decision makers within each team. (i) For a general class of exchangeable stochastic games with a finite number of decision makers, we first establish the existence of a Nash equilibrium under randomized policies (with common randomness) within each team that are exchangeable (but not necessarily symmetric, i.e., identical) among decision makers within each team. (ii) As the number of decision makers within each team goes to infinity (that is, for the mean-field limit game among teams), we show that a Nash equilibrium exists under randomized policies within each team that are independently randomized and symmetric among decision makers within each team (that is, there is no common randomness). (iii) Finally, we establish that a Nash equilibrium for a class of mean-field games among teams under independently randomized symmetric policies constitutes an approximate Nash equilibrium for the corresponding prelimit (exchangeable) game among teams with finite but large numbers of decision makers. (iv) We thus establish a rigorous connection between agent-based-modeling and team-against-team games, via the representative agents defining the game played in equilibrium, and we furthermore show that common randomness is not necessary for large team-against-team games, unlike the case with small-sized ones.

Key words. stochastic games among teams, mean-field games, information structure, Nash equilibrium

MSC codes. 91A13, 93E99, 93A99

DOI. 10.1137/22M1534055

1. Introduction.

Stochastic teams entail a collection of decision makers (DMs) acting together to optimize a common cost function, but not necessarily sharing all the available information. At each time stage, each DM only has partial access to the global information, which is defined as the *information structure* (IS) of the team [39]. If there is a predefined order according to which the DMs act, then the team is called a *sequential team.* For sequential teams, if each DM's information depends only on primitive random variables, the team is *static.* If at least one DM's information is affected by an action of another DM, the team is said to be *dynamic.* If the cost or the probabilistic model are not equal or equivalent, such a setup is considered as a game. In this paper, we study a class of stochastic games where each player in the game entails a sequential team with finite (but large) as well as countably infinite

*Received by the editors November 10, 2022; accepted for publication (in revised form) January 14, 2024; published electronically May 16, 2024. A summary of some of the results presented here is included in American Control Conference (ACC), IEEE, 2023, pp. 1104–1109.
https://doi.org/10.1137/22M1534055
Funding: The work of the third author was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.
†Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, ON, Canada (sanjari@rmc.ca).
‡Department of Mathematics, Bilkent University, Ankara, Turkey (naci.saldi@bilkent.edu.tr).
§Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada (yuksel@queensu.ca).

number of DMs. For this class of games among teams, we characterize existence and structural properties (i.e., symmetry) of a Nash equilibrium (NE).

Mean-field games are limit models of symmetric non-zero-sum noncooperative finite DM games with a mean-field interaction (see, e.g., [21, 20, 26]). Mean-field games have many applications in financial engineering, economics, and pricing in markets (see [11] for a thorough review on some applications of mean-field games). The existence of an NE for mean-field games has been established in [26, 5, 12, 27, 23, 28]. Also, there have been several studies for mean-field games where the limits of sequences of Nash equilibria have been investigated as the number of DMs goes to infinity (see, e.g., [16, 25, 6, 26, 4, 24]). In contrast to the setting of mean-field games studied in the literature, in this paper, we study stochastic mean-field games among teams. Such problems exhibit additional nuances due to lack of convexity in team policies under decentralized ISs; see, e.g., [3, 19, 29]. Such models have several applications as it combines mean-field game and team theory. A natural application area is in the sensor network where a large collection of decentralized sensors (acting as large teams) shares their information (by their actions) to a fusion center in the presence of jamming in the system entailing a large collection of decentralized jammers. As another application, consider information sharing across a medium via interference channels involving a collection of multiterminal encoder and decoder pairs where each encoder-decoder pair can be viewed as a team whose information serves as noise for the other teams.

Unlike games where NEs are often of interest, the notion of optimality among each team is global optimality. In general, NEs (in the language of teams, person-by-person optimal solutions) only arrive at local optima and not global optima. This gap is especially significant for stochastic teams with a countably infinite number of DMs (or mean-field teams). This is because any perturbation of finitely many policies fails to significantly deviate from the value of the expected cost. In view of this observation, the results on games may be inconclusive regarding global optimality for teams without uniqueness (e.g., see [26] for a strong sufficient condition of the monotonicity condition for establishing uniqueness). Nonuniqueness of NEs has also been studied; see, e.g., [5, 14, 10, 17, 7, 13, 25].

Mean-field teams under decentralized IS have been studied in [34, 33, 32]. For such models, the existence and convergence of a globally optimal solution have been established in [34, 33, 32]. In particular, it has been shown that under continuity of the cost function, the mean-field team admits a symmetric (possibly randomized) globally optimal solution. Since we are studying games among teams, the NE concept must take into account the global optimality notion among DMs within each team. In this paper, we study a class of mean-field games among a finite number of teams with infinite DMs under the concept of NEs which is globally optimal among DMs within each team but NE among teams (which are viewed as players). For this class of mean-field games, under sufficient continuity of cost functions, we show that an NE exists that exhibits symmetry among DMs within teams. The existence of such a symmetric NE does not a priori follow directly from a fixed point argument involving individual agents since restricting finding the best response policies to symmetric policies might result in a loss of global optimality for the DMs within each team. This is because, by fixing policies of other teams (players) to symmetric policies, each DM within a deviating team (player) faces a team problem that might not admit a symmetric optimal solution; e.g., see [32, Example 1]. Related to our setting of games among large teams, in [42], a mean-field-type rank-based reward competition game was studied among teams with a continuum of players and centralized IS.

In addition, for zero-sum games involving finite teams against finite teams, the existence of a randomized saddle-point equilibrium with common and private randomness has been established in [19]. We also note that a class of games among finite teams with a delayed information sharing structure has been studied in [38], analyzing connections between compression of information, perseverance of an NE payoff under compressed policies, and computation of NEs via sequential decomposition. In this paper, our focus is on exchangeable games among teams with both large and infinite numbers of DMs to establish the existence of an NE that is exchangeable for the finite case and symmetric in the mean-field limit. In [19, 3], convexity of the set of randomized policies with common and private randomness is required for the existence of an equilibrium, which can be realized by considering an arbitrary distribution over all possible common randomness variables. In practice, however, this may be too restrictive due to the arbitrary nature of distributions and since randomness must be externally provided.

1.1. Contributions.

1. For general games among teams with a finite number of DMs, in Theorem 3.4, we establish the existence of an NE under randomized policies with common and independent randomness. For the special case of exchangeable games among teams with a finite number of DMs, in Theorems 3.5 and 4.1 (for static and dynamic settings, respectively), we establish the existence of an NE and show that it is exchangeable among DMs within each team (player). This NE might be asymmetric (i.e., nonidentical) with correlated randomization among DMs within each team (player). Toward this goal, we show that the set of exchangeable randomized policies is convex and compact in Lemma 3.3.
2. As the number of DMs within each team goes to infinity (that is, for the limiting mean-field game among teams), in Theorems 3.6 and 4.2 (for static and dynamic settings, respectively), we establish the existence of an NE and show that it is symmetric (i.e., identical) and independently randomized among DMs within each team (player). We establish that the game problem can be relaxed to that where representative DMs' behavior can be regarded as the behavior of their corresponding symmetric population. This relaxes, in particular, the need for common randomness required for the finite population regime.
3. In Theorems 3.7 and 4.3 (for static and dynamic settings, respectively), we show that an NE for a mean-field game among teams (which is symmetric and independent) constitutes an approximate NE for the corresponding game among teams with the mean-field interaction and a finite but large number of DMs.
4. As a final contribution, in view of our results noted above, our paper presents a rigorous connection between agent-based-modeling and team-against-team games, via the representative agents defining the game played in equilibrium. As noted in [8, section 4], agent-based-modeling is often used in economics, mathematical biology, and other sciences to refer to a model in which macroscopic phenomena are captured by aggregating individual actions.

2. Problem formulation.

2.1. A generalized intrinsic model. Consider a class of stochastic games where DMs act in a predefined order (i.e., a sequential game [39]). Under the intrinsic model for sequential games (described in the discrete time setting), any action applied at any given time is regarded as applied by an individual DM, who acts only once. However, depending on the desired equilibrium concepts, IS, and cost functions, it is suitable to consider a collection of DMs as a single player acting as a team. In this setting, teams take part in a game. For this setting, Witsenhausen's intrinsic model is inadequate as it views each DM individually, which does not capture a joint deviation of a collection of DMs acting as a team, required for our equilibrium concept. To formalize this class of games, we introduce a class of M-player games using the generalized intrinsic model in [30], where each player consists of a collection of (one-shot) DMs either with finite or infinite members. A formal description has the following components:

- Let $\mathcal{M} := \{1, 2, \ldots, M\}$ denote the set of players. For each $i \in \mathcal{M}$, let $\mathcal{N}_i := \{1, 2, \ldots, N_i\}$ denote the collection of DMs, DM_k for $k \in \mathcal{N}_i$, acting as player i (PL^i) (in other words, PL^i encapsulates the collection of DMs indexed by $\mathcal{N}_i$). We denote DM_k of PL^i by DM^i_k.
- There exists a collection of *measurable spaces* $\{(\Omega, \mathcal{F}), (\mathbb{U}^i_k, \mathcal{U}^i_k), (\mathbb{Y}^i_k, \mathcal{Y}^i_k)$ for $k \in \mathcal{N}_i$ and $i \in \mathcal{M}\}$, specifying the system's distinguishable events, DMs' control, and observation spaces. The observation and action spaces are standard Borel spaces and are described by the product spaces $\prod_{k \in \mathcal{N}^i} \mathbb{Y}^i_k$ and $\prod_{k \in \mathcal{N}_i} \mathbb{U}^i_k$, respectively, for each PL^i.
- The $\mathbb{Y}^i_k$-valued observation variables are given by $y^i_k = h^i_k(\omega, \{u^p_s\}_{(s,p) \in L^i_k})$, where h^i_k is a Borel measurable function and L^i_k denotes the set of all DMs acting before DM^i_k (i.e., $(s,p) \in L^i_k$ if DM^p_s acts before DM^i_k).
- An admissible policy for each PL^i is denoted by $\underline{\gamma}^i := \{\gamma^i_k\}_{k \in \mathcal{N}_i} \in \Gamma^i$ with $u^i_k = \gamma^i_k(y^i_k)$, and the set of admissible policies for each PL^i is described by the product space $\Gamma^i := \prod_{k \in \mathcal{N}_i} \Gamma^i_k$, where Γ^i_k is the set of all Borel measurable maps from $\mathbb{Y}^i_k$ to $\mathbb{U}^i_k$.
- There is a *probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$, making it a probability space. Let $\mathbb{E}$ denote the expectation with respect to $\mathbb{P}$. The prior probability measures in general can be subjective (for each DM or player); however, in this paper, we will not discuss this, assuming that DMs have access to the common correct prior $\mathbb{P}$.

Under the above formulation, a sequential game is *dynamic* if the information available to at least one DM is affected by the action of at least one other DM. A game is *static* if for every DM the information is only affected by exogenous disturbances. ISs can also be categorized as *classical, quasi-classical,* or *nonclassical.* An IS $\{y^i_k | i \in \mathcal{M}$ and $k \in \mathcal{N}_i\}$ is classical if y^i_k contains all of the information available to DM^p_s who acts precedent to DM^i_k (i.e., $(s,p) \in L^i_k$). An IS is quasi-classical (or *partially nested*) if whenever u^p_s affects y^i_k through the observation function h^i_k, y^i_k contains y^p_s (that is, $\sigma(y^p_s) \subset \sigma(y^i_k)$). An IS which is not partially nested is nonclassical.

We will also allow for randomized policies, where in addition to y^i_k, each DM^i_k has access to common and private randomization. This will be made precise in section 3.1.

2.2. Problems studied. We study a class of stochastic games with a finite number of players that involve a finite number of DMs, or a countably infinite number of DMs. For simplicity of notation, we consider only two players in the game, denoted by PL^1 and PL^2 (i.e., $\mathcal{M} = \{1, 2\}$). Our results remain valid for the finite player setting. We address the following questions:

(i) *Does there exist an NE within possibly randomized policies for games with a finite number of DMs? Is this NE exchangeable?* We address the first question for general static games in Theorem 3.4, and both questions for exchangeable static games in Theorem 3.5 and for exchangeable dynamic ones in Theorem 4.1.

(ii) *Does there exist an NE within possibly randomized policies for games with a countably infinite number of DMs? Is this NE symmetric and independently randomized within each team?* We address these questions for the limiting mean-field static games in Theorem 3.6 and for the dynamic ones in Theorem 4.2.

(iii) *Do symmetric NEs for games with countably infinite number of DMs constitute approximate NEs for the corresponding prelimiting games with finite but large number of DMs?* We address this question for static games in Theorem 3.7 and for dynamic ones in Theorem 4.3.

The NE concept for such a class of games among teams should take into account the fact that DMs within players face a team problem with the desired global optimality notion for fixed policies of the other players. This is because fixing policies of DMs within each player might lead to a local optimal and not globally optimal. We additionally note that since the NE concept for such games among teams requires global optimality within teams, establishing the existence of an NE that exhibits exchangeability or symmetry requires further analysis that is not required for the classical stochastic games where each player is a singleton. This is because, in order to utilize a fixed point theorem, convexity and compactness of a set of policies (for each player, entailing all its DMs) are required. To address this difficulty, we endow a suitable topology on the set of randomized policies which leads to convexity and compactness of a set of exchangeable randomized policies. As the number of DMs goes to infinity, we use a de Finetti representation theorem and an argument used in [32] to establish symmetry, and to address the aforementioned questions.

2.3. Static games. Our focus is on exchangeable stochastic games, and hence we suppose that action and observation spaces are identical through DMs of each player and are subsets of appropriate dimensional Euclidean spaces, i.e., $\mathbb{U}^i_k := \mathbb{U}^i \subseteq \mathbb{R}^{n_i}$ and $\mathbb{Y}^i_k := \mathbb{Y}^i \subseteq \mathbb{R}^{m_i}$ for all $i \in \{1,2\}$, where n_i and m_i are positive integers. In the following, we first introduce a class of exchangeable static games with a finite number of DMs, and then we introduce their limiting static mean-field games.

2.3.1. Finite DM static game $\boldsymbol{\mathcal{P}_N}$. Consider a stochastic game with a finite number of DMs within each player (team), i.e., each PL^i (for $i \in \{1,2\}$) involves a finite number of DMs indexed by $\mathcal{N}_i = \{1,\dots,N_i\}$. Each PL^i's admissible policies $\underline{\gamma}^i_N := (\gamma^i_1,\dots,\gamma^i_{N_i})$ with $\Gamma^i_N := \prod_{k=1}^{N_i}\Gamma^i_k$ are measurable functions so that $u^i_k = \gamma^i_k(y^i_k)$ for $k \in \mathcal{N}_i$. Let the expected cost function for each PL^i under a policy profile $\underline{\gamma}^{1:2}_N := (\underline{\gamma}^1_N, \underline{\gamma}^2_N)$ be given by

$$J^i_N(\underline{\gamma}^{1:2}_N) = \mathbb{E}^{\underline{\gamma}^{1:2}_N}\left[c^i(\omega_0, \underline{u}^{1:2}_N)\right] \tag{2.1}$$

for some Borel measurable cost function $c^i : \Omega_0 \times \prod_{j=1}^2 \prod_{k=1}^{N_j} \mathbb{U}^j \to \mathbb{R}_+$. We define ω_0 as the Ω_0-valued, cost function relevant, exogenous random variable, taking values from a Borel space Ω_0 with its Borel σ-field $\mathcal{F}_0$. In the above, we used the notation $k:j$ to denote $\{k,\dots,j\}$, and $\underline{u}^i_N := u^i_{1:N_i}$, and $\mathbb{E}^{\underline{\gamma}^{1:2}_N}$ to denote the expectation with respect to $\mathbb{P}$ when actions u^i_k are induced by policies γ^i_k given the private observations y^i_k for each DM^i_k.

DEFINITION 2.1 (ϵ-Nash equilibrium (NE) for $\mathcal{P}_N$). *Let $\epsilon \geq 0$. A policy profile $\underline{\gamma}^{1*:2*}_N$ constitutes an ϵ-NE if and only if the following inequalities hold for all $i \in \{1,2\}$:*

$$J_N^i(\underline{\gamma}_N^{1*:2*}) \le \inf_{\underline{\gamma}_N^i \in \Gamma_N^i} J_N^i(\underline{\gamma}_N^{-i*}, \underline{\gamma}_N^i) + \epsilon, \tag{2.2}$$

where $-i := \{1,2\} \backslash \{i\}$*. If* $\epsilon = 0$*, the policy profile* $\underline{\gamma}_N^{1*:2*}$ *constitutes an NE.*

We emphasize that in the above, DMs within a player (team) optimize a common cost function but may have a decentralized IS. Hence, our notion of (player-wise) NEs is a suitable equilibrium notion for such games since (global) optimality is desirable among DMs within players, which takes into account the fact that for each player, DMs face a team problem when the policies of other players are fixed. In section 3, we allow DMs to apply randomized policies, and hence we rewrite our game formulation and the equilibrium notion to incorporate randomized policies; see (3.6) (also (3.7)) and (3.8).

For general exchangeable static games, the cost functions of the players satisfy the following exchangeability condition.

Assumption 2.1. c^i is (separately) exchangeable with respect to actions of DMs for all ω_0 and $i \in \{1,2\}$, i.e., for any permutations σ and τ of $\{1,\dots,N_1\}$ and $\{1,\dots,N_2\}$, respectively,

$$c^i\left(\omega_0, u^1_{1:N_1}, u^2_{1:N_2}\right) = c^i\left(\omega_0, u^1_{\sigma(1):\sigma(N_1)}, u^2_{\tau(1):\tau(N_2)}\right) \quad \forall \omega_0,$$

where $u^1_{\sigma(1):\sigma(N_1)} = u^1_{\sigma(1)},\dots,u^1_{\sigma(N_1)}$ and $u^2_{\tau(1):\tau(N_2)} = u^2_{\tau(1)},\dots,u^2_{\tau(N_2)}$.

Let $\mathcal{P}(\cdot)$ denote the space of probability measures, and let $\delta_{\{\cdot\}}$ denote the Dirac delta. A special case of the preceding game is when the costs satisfy Assumption 2.1 and are given by

$$c^i\left(\omega_0, u^1_{1:N_1}, u^2_{1:N_2}\right) = \frac{1}{N_i}\sum_{k=1}^{N_i} c_k^i\left(\omega_0, u_k^i, \Xi^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u_p^1}\right), \Xi^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_p^2}\right)\right) \tag{2.3}$$

for some weakly continuous functions $\Xi^i : \mathcal{P}(\mathbb{U}^i) \to \widetilde{\mathbb{U}}^i$ with $\widetilde{\mathbb{U}}^i$, which is a subset of appropriate dimensional Euclidean space, and for some Borel measurable cost functions $c_k^i : \Omega_0 \times \mathbb{U}^i \times \prod_{j=1}^2 \widetilde{\mathbb{U}}^j \to \mathbb{R}_+$ that satisfy

$$\begin{aligned}
&\sum_{k=1}^{N_i} c_k^i\left(\omega_0, u_k^i, \Xi^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u_p^1}\right), \Xi^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_p^2}\right)\right)\\
&\quad = \sum_{k=1}^{N_i} c_{\beta(k)}^i\left(\omega_0, u_{\beta(k)}^i, \Xi^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u_p^1}\right), \Xi^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_p^2}\right)\right)
\end{aligned} \tag{2.4}$$

for every permutation $\beta = \sigma$ or $\beta = \tau$. Clearly, if $c_k^i = \tilde{c}^i$, then the above exchangeability condition holds. When the number of DMs goes to infinity, we consider the special case of $c_k^i = \tilde{c}^i$. We note that weak continuity of Ξ^i is not required for the exchangeability condition, but it is required for our main theorems.

2.3.2. Limiting mean-field static game $\mathcal{P}_\infty$. Consider a stochastic game with a countably infinite number of DMs within each player (team), i.e., each PL^i involves a countably infinite number of DMs indexed by $\mathcal{N}_i = \mathbb{N}$. The set of admissible policies $\underline{\gamma}^i := (\gamma_1^i, \gamma_2^i, \dots)$ for each PL^i is denoted by $\Gamma^i := \prod_{k\in\mathbb{N}} \Gamma_k^i$, where each γ_k^i is a measurable function so that $u_k^i = \gamma_k^i(y_k^i)$ for $k \in \mathcal{N}_i$. Let the expected cost of PL^i under a policy profile $\underline{\gamma}^{1:2}$ be given by

$$J^i_\infty(\underline{\gamma}^{1:2}) = \limsup_{N_1,N_2\to\infty} \mathbb{E}^{\underline{\gamma}^{1:2}}\left[\frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N^1}\sum_{p=1}^{N_1}\delta_{u^1_p}\right), \Xi^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u^2_p}\right)\right)\right] \tag{2.5}$$

for some weakly continuous functions $\Xi^i : \mathcal{P}(\mathbb{U}^i) \to \tilde{\mathbb{U}}^i$ and Borel measurable cost functions $c^i : \Omega_0 \times \mathbb{U}^i \times \prod_{j=1}^2 \tilde{\mathbb{U}}^j \to \mathbb{R}_+$.

Definition 2.2 (Nash equilibrium (NE) for $\mathcal{P}_\infty$). *A policy profile $\underline{\gamma}^{1*:2*}$ is NE (or mean-field equilibrium) for $\mathcal{P}_\infty$ if and only if the following inequalities hold for all $i \in \{1,2\}$:*

$$J^i_\infty(\underline{\gamma}^{1*:2*}) \le J^i_\infty(\underline{\gamma}^{-i,*}, \underline{\gamma}^i) \qquad \forall \underline{\gamma}^i \in \Gamma^i. \tag{2.6}$$

Our first main goal is to establish the existence of a symmetric (identical) randomized NE for static mean-field games $\mathcal{P}_\infty$. Our results require the following absolute continuity condition.

Assumption 2.2. For every $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$, let $\tilde{\mu}^N$ be the conditional distribution of observations $\underline{y}^{1:2}_N$ given ω_0. There exists a probability measure Q^i_k on $\mathbb{Y}^i$ and a bounded function $f^i_k : \mathbb{Y}^i \times \Omega_0 \times \prod_{j=1,\neq k}^{N_i} \mathbb{Y}^i \times \prod_{j=1}^{N_{-i}} \mathbb{Y}^{-i} \to \mathbb{R}_+$ for all $i \in \{1,2\}$ and $k \in \mathcal{N}_i$ such that for every Borel set B^i_k in $\mathbb{Y}^i$ (with $B := \prod_{i=1}^2 B^i_1 \times \cdots \times B^i_{N_i}$)

$$\tilde{\mu}^N(B|\omega_0) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{B^i_k} f^i_k\left(y^i_k, \omega_0, y^i_{-k}, \underline{y}^{-i}\right) Q^i_k(dy^i_k), \tag{2.7}$$

where $y^i_{-k} = \{y^i_{1:N_i}\} \backslash \{y^i_k\}$.

Under the change of measure in (2.7), probability measures on observations y^i_k are Q^i_k that are independent of observations y^i_{-k}, $\underline{y}^{-i}$, and also ω_0. Hence, under this assumption, via a change of measure argument (see, e.g., [40, 30]), we can equivalently view the observations of each DM as independent and also independent of ω_0. Under the new probability measure, the dependency of the random variables is incorporated into the expected cost function via the terms f^i_k's. The above allows us to introduce a suitable topology under which the space of randomized policies is Borel (see section 3.1). In addition, our main results (i.e., Theorems 3.5 and 3.6) for static games impose the following assumptions on the observations, action spaces, and cost functions.

Assumption 2.3.
(i) $(y^1_k)_{k\in\mathcal{N}_1}$ and $(y^2_k)_{k\in\mathcal{N}_2}$ are independent, conditioned on ω_0;
(ii) for all $i \in \{1,2\}$, $(y^i_k)_{k\in\mathcal{N}_i}$ have an identical distribution, conditioned on ω_0.

Assumption 2.4. For $i \in \{1,2\}$,
(i) $\mathbb{U}^i$ is compact;
(ii) $c^i(\omega_0, \cdot, \cdot, \cdot)$ in (2.3) is continuous for all ω_0.

We note that under Assumption 2.3, we can rewrite (2.7) as follows:

$$\tilde{\mu}^N(B|\omega_0) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{B^i_k} \hat{f}^i\left(y^i_k, \omega_0\right) Q^i(dy^i_k), \tag{2.8}$$

where $\hat{f}^i : \mathbb{Y}^i \times \Omega_0 \to \mathbb{R}_+$ and Q^i are identical for all DMs within player i for $i \in \{1,2\}$.

In section 3, we establish the existence of a randomized NE for $\mathcal{P}_\infty$, and we show that this NE is symmetric (identical among DMs within players) and independent. To this end, we first establish the existence of a randomized NE for $\mathcal{P}_N$ that is exchangeable (possibly correlated and asymmetric), and then we use this result to arrive at the existence result for $\mathcal{P}_\infty$. For our results in section 3, we let Assumptions 2.1 and 2.2 hold. For exchangeable games with a finite number of DMs, we consider Assumption 2.4, but we relax Assumption 2.3, where we allow DMs to have (finite-) exchangeable observations that can be correlated (see Assumption 3.1). When the number of DMs goes to infinity, we consider mean-field interaction among them, and we let Assumptions 2.3 and 2.4 hold.

2.4. Dynamic games. We again let action, observation, and state spaces, respectively, be identical through DMs $k \in \mathcal{N}_i$, also identical over time $t \in \{0,\dots,T-1\}$ for simplicity, and subsets of appropriate dimensional Euclidean spaces $\mathbb{U}^i_{k,t} = \mathbb{U}^i \subseteq \mathbb{R}^{n_i}$, $\mathbb{Y}^i_{k,t} = \mathbb{Y}^i \subseteq \mathbb{R}^{n'_i}$, $\mathbb{X}^i_{k,t} = \mathbb{X}^i \subseteq \mathbb{R}^{n''_i}$, where n_i, n'_i, and n''_i are positive integers.

Let state dynamics and observations of DM^i_k for $i \in \{1,2\}$ and $k \in \mathcal{N}_i$ be given by

$$x^i_{k,t+1} = f^i_t\Bigg(x^i_{k,t}, u^i_{k,t}, \Xi^1_x\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{x^1_{p,t}}\right), \Xi^2_x\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{x^2_{p,t}}\right), \Xi^1_u\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u^1_{p,t}}\right), \Xi^2_u\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u^2_{p,t}}\right), w^i_{k,t}\Bigg), \tag{2.9}$$

$$y^i_{k,t} = h^i_t\left(x^i_{k,0:t}, u^i_{k,0:t-1}, v^i_{k,0:t}\right) \tag{2.10}$$

for some weakly continuous functions $\Xi^i_x : \mathcal{P}(\mathbb{X}^i) \to \tilde{\mathbb{X}}^i$ and $\Xi^i_u : \mathcal{P}(\mathbb{U}^i) \to \tilde{\mathbb{U}}^i$, where $\tilde{\mathbb{X}}^i$ and $\tilde{\mathbb{U}}^i$ are subsets of appropriate dimensional Euclidean spaces. In the above, functions f^i_t and h^i_t are measurable functions, and $v^i_{k,t}$ and $w^i_{k,t}$ are exogenous random variables, representing uncertainties in state dynamics and observations.

In the following, we first introduce a class of games with a finite number of DMs within each player, and then we introduce their limiting dynamic mean-field games.

2.4.1. Finite DM dynamic game $\boldsymbol{\mathcal{P}^T_N}$. Consider a stochastic dynamic game with a finite number of DMs. Each PL^i's admissible policies $\underline{\boldsymbol{\gamma}}^{\boldsymbol{i}}_{\boldsymbol{N}} := \underline{\gamma}^i_{N,0:T-1}$ with $\underline{\gamma}^i_{N,t} := \gamma^i_{1:N_i,t}$ are measurable functions so that $u^i_{k,t} = \gamma^i_{k,t}(y^i_{k,t})$ for all $k \in \mathcal{N}_i$ and $t \in \{0,\dots,T-1\}$. We use bold letters for variables over time $t \in \{0,\dots,T-1\}$, and underline variables to denote variables of DM^i_k's for $k \in \mathcal{N}_i$ within PL^i. Denote the space of policies for each player with $\boldsymbol{\Gamma}^{\boldsymbol{i}}_{\boldsymbol{N}} = \prod_{t=0}^{T-1}\prod_{k=1}^{N_i}\Gamma^i_{k,t}$. Let PL^i's expected cost function under $\underline{\boldsymbol{\gamma}}^{\boldsymbol{1:2}}_{\boldsymbol{N}} \in \prod_{i=1}^2 \boldsymbol{\Gamma}^{\boldsymbol{i}}_{\boldsymbol{N}}$ be given by

$$J^{i,T}_N(\underline{\boldsymbol{\gamma}}^{\boldsymbol{1:2}}_{\boldsymbol{N}}) = \mathbb{E}^{\underline{\boldsymbol{\gamma}}^{\boldsymbol{1:2}}_{\boldsymbol{N}}}\left[\frac{1}{N_i}\sum_{k=1}^{N_i}\sum_{t=0}^{T-1} c^i_k\Bigg(\omega_0, x^i_{k,t}, u^i_{k,t}, \Xi^1_x\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{x^1_{p,t}}\right), \Xi^2_x\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{x^2_{p,t}}\right), \Xi^1_u\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u^1_{p,t}}\right), \Xi^2_u\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u^2_{p,t}}\right)\Bigg)\right] \tag{2.11}$$

for some weakly continuous functions $\Xi^i_x : \mathcal{P}(\mathbb{X}^i) \to \widetilde{\mathbb{X}}^i$ and $\Xi^i_u : \mathcal{P}(\mathbb{U}^i) \to \widetilde{\mathbb{U}}^i$, where $\widetilde{\mathbb{X}}^i$ and $\widetilde{\mathbb{U}}^i$ are subsets of Euclidean spaces. In the above, $c^i_k : \Omega_0 \times \mathbb{X}^i \times \mathbb{U}^i \times \prod_{j=1}^2 \widetilde{\mathbb{X}}^j \times \widetilde{\mathbb{U}}^j$ is a Borel measurable cost function that satisfies

$$\sum_{k=1}^{N_i} c_k^i(\cdot, x_{k,t}^i, u_{k,t}^i, \cdot, \ldots, \cdot) = \sum_{k=1}^{N_i} c_{\beta(k)}^i(\cdot, x_{\beta(k),t}^i, u_{\beta(k),t}^i, \cdot, \ldots, \cdot) \tag{2.12}$$

for any permutations $\beta = \sigma$ or $\beta = \tau$. Random variable ω_0 is again a cost-related random variable. If $c_k^i = \tilde{c}^i$, then (2.12) holds since

$$\sum_{k=1}^{N_i} \tilde{c}^i\left(\omega_0, x_{k,t}^i, u_{k,t}^i, \Xi_x^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{x_{p,t}^1}\right), \ldots, \Xi_u^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_{p,t}^2}\right)\right)$$
$$= \sum_{k=1}^{N_i} \tilde{c}^i\left(\omega_0, x_{\beta(k),t}^i, u_{\beta(k),t}^i, \Xi_x^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{x_{p,t}^1}\right), \ldots, \Xi_u^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_{p,t}^2}\right)\right).$$

When the number of DMs goes to infinity, we consider the special case $c_k^i = \tilde{c}^i$.

DEFINITION 2.3 (ϵ-NE for $\mathcal{P}_N^T$). *Let $\epsilon \geq 0$. A policy profile $\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{1*:2*}}$ constitutes an ϵ-NE if and only if the following inequalities hold for all $i \in \{1,2\}$:*

$$J_N^{i,T}(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{1*:2*}}) \leq \inf_{\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{i}} \in \boldsymbol{\Gamma}_{\boldsymbol{N}}^{\boldsymbol{i}}} J_N^{i,T}(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{-i*}}, \underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{i}}) + \epsilon. \tag{2.13}$$

If $\epsilon = 0$, then $\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{1*:2*}}$ constitutes an NE.

In section 4, we allow DMs to apply randomized policies, and we rewrite our game formulation and the equilibrium notion to incorporate randomized policies; see (4.6) and (4.7).

2.4.2. Limiting mean-field dynamic game $\boldsymbol{\mathcal{P}_\infty^T}$. Consider a stochastic dynamic game with a countably infinite number of DMs within each player (team), i.e., each PL^i involves a countably infinite number of DMs indexed by $\mathcal{N}_i = \mathbb{N}$. The set of admissible policies $\underline{\boldsymbol{\gamma}}^{\boldsymbol{i}} := \underline{\gamma}_{0:T-1}^i$ and $\underline{\gamma}_t^i := (\gamma_{1,t}^i, \gamma_{2,t}^i, \ldots)$ for each PL^i is denoted by $\boldsymbol{\Gamma}^{\boldsymbol{i}} := \prod_{t=0}^{T-1}\prod_{k\in\mathbb{N}}\Gamma_{k,t}^i$, where each $\gamma_{k,t}^i$ is a measurable function so that $u_{k,t}^i = \gamma_{k,t}^i(y_{k,t}^i)$ for $k \in \mathcal{N}_i$ and $t \in \{0, \ldots, T-1\}$. Let the expected cost of PL^i under a policy profile $\underline{\boldsymbol{\gamma}}^{\boldsymbol{1:2}}$ be given by

$$J_\infty^{i,T}(\underline{\boldsymbol{\gamma}}^{\boldsymbol{1:2}}) = \limsup_{N_1, N_2 \to \infty} J_N^{i,T}(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}^{\boldsymbol{1:2}}). \tag{2.14}$$

Similarly to (2.13), we can define NE (or mean-field equilibrium) for $\mathcal{P}_\infty^T$. Our solution approach for the dynamic setting is similar to that of the static one with additional technical arguments and assumptions. Our theorems for the dynamic games (Theorems 4.1 and 4.2) require that absolute continuity conditions in Assumption 4.1 hold. Assumption 4.1 allows us to endow a suitable topology for the space of randomized policies. Furthermore, Theorem 4.2 requires the following assumptions.

Assumption 2.5. For $i \in \{1,2\}$,

(i) $f_t^i(\cdot, \ldots, \cdot, w_{k,t}^i)$ and $h_t^i(\cdot, \cdot, v_{k,0:t}^i)$ are continuous and uniformly bounded for all $(w_{k,t}^i, v_{k,0:t}^i)$ and for $t \in \{0, \ldots, T-1\}$;

(ii) $c^i(\omega_0, \cdot, \ldots, \cdot)$ is continuous and uniformly bounded for all ω_0;

(iii) $\mathbb{U}^i$ is compact.

Assumption 2.6. For $i \in \{1,2\}$,

(i) $(x_{k,0}^i)_{k\in\mathcal{N}_i}$ are i.i.d., conditioned on ω_0;

(ii) $(w_{k,t}^i)_{k\in\mathcal{N}_i}$ are i.i.d. for $t \in \{0, \ldots, T-1\}$, and $(w_{k,0:T-1}^i)$ are mutually independent, and independent of ω_0 and $(x_{k,0}^i)_{k\in\mathcal{N}_i}$; random variables $(v_{k,t}^i)_{k\in\mathcal{N}_i}$ are i.i.d. for $t \in \{0, \ldots, T-1\}$, and $(v_{k,0:T-1}^i)$ are mutually independent, and independent of ω_0, $(x_{k,0}^i)_{k\in\mathcal{N}_i}$, and $(w_{k,t}^i)_{k\in\mathcal{N}_i}$ for $t \in \{0, \ldots, T-1\}$.

In section 4, we establish the existence of a randomized NE for $\mathcal{P}_{\infty}^{T}$, and we show that it is symmetric (identical among DMs within players) and independent. Similarly to the static case, for a more general formulation of $\mathcal{P}_{N}^{T}$, we first establish the existence of a randomized NE that is exchangeable, and then we use the result to arrive at the existence result for $\mathcal{P}_{\infty}^{T}$. For our results in section 4, we let Assumption 4.1 hold, but we let Assumptions 4.4, 2.5, and 2.6 hold only when they are needed.

3. Exchangeable static games. In this section, we study exchangeable static stochastic games. We first introduce randomized policies equipped with a suitable topology, and then we establish the existence of an NE for $\mathcal{P}_N$ that is exchangeable, and an NE for $\mathcal{P}_{\infty}$ that is symmetric.

3.1. Topology on control policies for static games. In the following, we introduce randomized policies as Borel probability measures, equipped with a suitable topology. Following [44, 40], via a change of measure argument in Assumption 2.2, we make observations of DMs of players independent and also independent of ω_0. This enables us to separate the DMs' policy spaces (both across the player and DMs) and equip each of them with a topology introduced below. Let

$$\Theta_k^i := \Bigg\{ P \in \mathcal{P}(\mathbb{U}^i \times \mathbb{Y}^i) \Bigg| \; P(B) = \int_B \delta_{\{g_k^i(y_k^i)\}}(du_k^i) Q_k^i(dy_k^i),\; g_k^i : \mathbb{Y}^i \to \mathbb{U}^i,\; B \in \mathcal{B}(\mathbb{U}^i \times \mathbb{Y}^i) \Bigg\}. \tag{3.1}$$

Also, let

$$\mathcal{R}_k^i := \left\{ P \in \mathcal{P}(\mathbb{U}^i \times \mathbb{Y}^i) \middle| P(B) = \int_B \Pi_k^i(du_k^i|y_k^i) Q_k^i(dy_k^i),\; B \in \mathcal{B}(\mathbb{U}^i \times \mathbb{Y}^i) \right\}, \tag{3.2}$$

where Π_k^i is a stochastic kernel from $\mathbb{Y}^i$ to $\mathbb{U}^i$. The set Θ_k^i is the set of extreme points of $\mathcal{R}_k^i$. Hence, Θ_k^i inherits Borel measurability and topological properties of the Borel measurable set $\mathcal{R}_k^i$ under the weak convergence topology [9]. We identify the set of relaxed policies[1] for each DM_k^i, Γ_k^i, by $\mathcal{R}_k^i$, with the convergence in policies defined by

$$\gamma_{n,k}^i \xrightarrow[n\to\infty]{} \gamma_k^i \iff \gamma_{k,n}^i(du_k^i|y_k^i)Q_k^i(dy_k^i) \xrightarrow[n\to\infty]{\text{weakly}} \gamma_k^i(du_k^i|y_k^i)Q_k^i(dy_k^i). \tag{3.3}$$

We equip $\mathcal{P}(\mathbb{Y}^i \times \mathbb{U}^i)$ with the w-s topology, which is the coarsest topology under which the map $\int \kappa(y_k^i, u_k^i) P(dy_k^i, du_k^i) : \mathcal{P}(\mathbb{Y}^i \times \mathbb{U}^i) \to \mathbb{R}$ is continuous for every measurable and bounded function κ, which is continuous in u_k^i for every y_k^i. Unlike the weak convergence topology, κ need not be continuous in y_k^i (see [35]). Since the marginals on observations are fixed, under the w-s topology, the convergence coincides with that in the weak convergence topology; i.e., the convergence of probability measures is the weak convergence (see [35, Theorem 3.10]). Hence. we can view convergence in (3.3) in terms of w-s topology without any loss of generality.

3.1.1. Randomized policies for $\mathcal{P}_N$. The above formulation of relaxed policies for $\Gamma_N^i := \prod_{k=1}^{N_i} \Gamma_k^i$ allows us to introduce the set randomized policies $L_N^i := \mathcal{P}(\Gamma_N^i)$ for each PL^i as a collection of Borel probability measures on Γ_N^i, where Borel σ-field $\mathcal{B}(\Gamma_k^i)$ is induced by the introduced topology in (3.3) and Γ_N^i is equipped with the

[1]This set corresponds to the set of Young measures introduced in [41].

product topology. In the following, we introduce different sets of randomized policies by allowing common and individual randomness among DMs. Let

$$L^i_{\mathsf{CO},N} := \left\{ P_\pi \in L^i_N \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : P_\pi(\gamma^i_1 \in A_1, \dots, \gamma^i_{N_i} \in A_{N_i}) \right.$$
$$\left. = \int_{z^i \in [0,1]} \prod_{k=1}^{N_i} P_{\pi,k}(\gamma^i_k \in A_k | z^i) \eta^i(dz^i), \quad \eta^i \in \mathcal{P}([0,1]) \right\},$$

where η^i is the distribution of common randomness (independent of the intrinsic exogenous system random variables). In the above, for every fixed z^i, $P_{\pi,k} \in \mathcal{P}(\Gamma^i_s)$ corresponds to an independent randomized policy for each DM^i_k ($k \in \mathcal{N}_i$ and $i \in \{1,2\}$). Since $L^i_{\mathsf{CO},N}$ and L^i_N are equal (see [32, Theorem A.1]), L^i_N corresponds to randomized policies induced by individual and common randomness. We next recall the definition of *exchangeability* for random variables.

Definition 3.1. *Random variables $x^{1:m}$ defined on a common probability space are m-exchangeable if for any permutation σ of $\{1, \dots, m\}$, $\mathcal{L}\left(x^{\sigma(1):\sigma(m)}\right) = \mathcal{L}\left(x^{1:m}\right)$, where $\mathcal{L}(\cdot)$ denotes the (joint) law of random variables. Random variables $(x^1, x^2, \dots)$ are infinitely exchangeable if finite distributions of $(x^1, x^2, \dots)$ and $(x^{\sigma(1)}, x^{\sigma(2)}, \dots)$ are identical for any finite permutation (affecting only finitely many elements) of $\mathbb{N}$.*

The set of exchangeable randomized policies $L^i_{\mathsf{EX},N} \subseteq L^i_N$ for each PL^i is given by

$$L^i_{\mathsf{EX},N} := \left\{ P_\pi \in L^i_N \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) \text{ and } \forall \sigma \in S_{N_i} : \right.$$
$$\left. P_\pi(\gamma^i_1 \in A_1, \dots, \gamma^i_{N_i} \in A_{N_i}) = P_\pi(\gamma^i_{\sigma(1)} \in A_1, \dots, \gamma^i_{\sigma(N_i)} \in A_{N_i}) \right\}, \tag{3.4}$$

where S_{N_i} is the set of all permutations of $\{1, \dots, N_i\}$. We note that $L^i_{\mathsf{EX},N}$ is a convex subset of L^i_N. Let $L^i_{\mathsf{CO,SYM},N}$ be the set of all symmetric (identical) randomized policies induced by common and individual randomness:

$$L^i_{\mathsf{CO,SYM},N} := \left\{ P_\pi \in L^i_N \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : P_\pi(\gamma^i_1 \in A_1, \dots, \gamma^i_{N_i} \in A_{N_i}) \right.$$
$$\left. = \int_{z^i \in [0,1]} \prod_{k=1}^{N_i} \widetilde{P}_\pi(\gamma^i_k \in A_k | z) \eta^i(dz^i), \quad \eta^i \in \mathcal{P}([0,1]) \right\},$$

where for all $k \in \mathcal{N}_i$, conditioned on z^i, $\widetilde{P}_\pi \in \mathcal{P}(\Gamma^i_k)$ corresponds to an independent randomized policy for each DM^i_k, which is symmetric among PL^i's DMs. Note that $L^i_{\mathsf{CO,SYM},N} \subset L^i_{\mathsf{EX},N}$.

3.1.2. Randomized policies for $\mathcal{P}_\infty$. When the number of DMs for each player is countably infinite, we define the corresponding sets of randomized policies $L^i, L^i_{\mathsf{CO}}, L^i_{\mathsf{EX}}$, and $L^i_{\mathsf{CO,SYM}}$ using the Ionescu–Tulcea extension theorem and by iteratively adding new coordinates for our probability measures (see, e.g., [2, 18]). The set of randomized policies L on the infinite product Borel spaces $\Gamma^i := \prod_{k\in\mathbb{N}} \Gamma^i_k$ is defined as $L^i := \mathcal{P}(\Gamma^i)$. Let the set of all randomized policies with common and independent randomness be given by

$$L^i_{\mathsf{CO}} := \left\{ P_\pi \in L^i \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : P_\pi(\gamma^i_1 \in A_1, \gamma^i_2 \in A_2, \dots) \right.$$
$$\left. = \int_{z^i \in [0,1]} \prod_{k\in\mathbb{N}} P_{\pi,k}(\gamma^i_k \in A_k | z^i) \eta^i(dz^i), \quad \eta^i \in \mathcal{P}([0,1]) \right\}.$$

Let the set of all infinitely exchangeable randomized policies L^i_{EX} be given by

$$L^i_{\mathsf{EX}} := \left\{ P_\pi \in L^i \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k),\ \forall N_i \in \mathbb{N},\text{ and } \forall \sigma \in S_{N_i}: \right.$$
$$\left. P_\pi(\gamma^i_1 \in A_1, \ldots, \gamma^i_{N_i} \in A_{N_i}) = P_\pi(\gamma^i_{\sigma(1)} \in A_1, \ldots, \gamma^i_{\sigma(N_i)} \in A_{N_i}) \right\}.$$

Let the symmetric policies with common and independent randomness be given by

$$L^i_{\mathsf{CO,SYM}} := \left\{ P_\pi \in L^i \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : P_\pi(\gamma^i_1 \in A_1, \gamma^i_2 \in A_2, \ldots) \right.$$
$$\left. = \int_{z^i \in [0,1]} \prod_{k \in \mathbb{N}} \widetilde{P}_\pi(\gamma^i_k \in A_k | z^i) \eta^i(dz^i), \quad \eta^i \in \mathcal{P}([0,1]) \right\}.$$

For the infinite population setting, we also define the following set of randomized policies with only private independent randomness:

$$L^i_{\mathsf{PR}} := \left\{ P_\pi \in L^i \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : \right.$$
$$\left. P_\pi(\gamma^i_1 \in A_1, \gamma^i_2 \in A_2, \ldots) = \prod_{k \in \mathbb{N}} P_{\pi,k}(\gamma^i_k \in A_k) \text{ for } P_{\pi,k} \in \mathcal{P}(\Gamma^i_k) \right\}. \tag{3.5}$$

Define the set of symmetric randomized policies with private independent randomness as

$$L^i_{\mathsf{PR,SYM}} := \left\{ P_\pi \in L^i \,\middle|\, \forall A_k \in \mathcal{B}(\Gamma^i_k) : \right.$$
$$\left. P_\pi(\gamma^i_1 \in A_1, \gamma^i_2 \in A_2, \ldots) = \prod_{k \in \mathbb{N}} \widetilde{P}_\pi(\gamma^i_k \in A_k) \text{ for } \tilde{P}_\pi \in \mathcal{P}(\Gamma^i_k) \right\}.$$

We note that L^i_{PR} and $L^i_{\mathsf{PR,SYM}}$ are not convex sets; however, L^i_{PR} ($L^i_{\mathsf{PR,SYM}}$) contains the set of extreme points of the convex set L^i_{CO} ($L^i_{\mathsf{CO,SYM}}$). This fact plays a pivotal role later on in our analysis.

3.2. Some properties of randomized policies for $\mathcal{P}_N$ and $\mathcal{P}_\infty$. In the following, we first present two lemmas on convexity, compactness, and relationships between sets of exchangeable randomized policies. We use the following lemmas for our main results in Theorems 3.5–4.3.

LEMMA 3.2. *Consider $L^i_{\mathsf{EX},N}, L^i_{\mathsf{CO,SYM},N}, L^i_{\mathsf{EX}}$, and $L^i_{\mathsf{CO,SYM}}$. Then*
(i) *any $P_\pi \in L^i_{\mathsf{EX}}$ belongs also to $L^i_{\mathsf{CO,SYM}}$, i.e., $L^i_{\mathsf{CO,SYM}} = L^i_{\mathsf{EX}}$;*
(ii) *$L^i_{\mathsf{CO,SYM},N} \subseteq L^i_{\mathsf{EX},N}$; however, in general $L^i_{\mathsf{CO,SYM},N} \neq L^i_{\mathsf{EX},N}$.*

Proof. Part (i) has been established in [32, Theorem 1] using a de Finetti representation theorem [22, Theorem 1.1]. Part (ii) follows from the fact that finite exchangeable random variables are not necessarily conditionally i.i.d.; see, e.g., [36, Example 1.18] and [15, 1]. ∎

The proof of the following lemma is provided in Appendix A.1.

LEMMA 3.3. *Let $\mathbb{U}^i$ be compact. Then $L^i_{\mathsf{EX},N}$ and L^i_{EX} are convex and compact.*

3.3. Existence of a Nash equilibrium for $\mathcal{P}_N$: The role of common randomness. In this section, we establish an existence result for general stochastic game $\mathcal{P}_N$ without imposing exchangeability on the cost function (Assumption 2.1).

As in $\mathcal{P}_N$ when a team plays against another team, the existence of an NE may not hold in general. This is due to the lack of convexity of the space of policies for teams taking part in the game (see [3] for an example where such a situation arises). Considering the common randomness among DMs within teams is crucial for establishing the existence of an NE for these games since allowing the common randomness leads to the convexity of the policy space for each team.

First, the expected costs for $\mathcal{P}_N$ under a randomized policy profile $(P_\pi^1, P_\pi^2) \in L_N^1 \times L_N^2$ are given by

$$\begin{aligned} J_{\pi,N}^i(P_\pi^1, P_\pi^2) &:= \int P_\pi^1(d\underline{\gamma}^1) P_\pi^2(d\underline{\gamma}^2) \mu^N(d\omega_0, d\underline{y}) c_N^i(\underline{\gamma}, \underline{y}, \omega_0) \\ &:= \int \int c^i(\omega_0, u_{1:N_i}^1, u_{1:N_i}^2) \prod_{i=1}^2 \prod_{k=1}^{N_i} \gamma_k^i(du_k^i | y_k^i) \\ &\quad \times \prod_{i=1}^2 P_\pi^i(d\gamma_{1:N_i}^i) \mu^N(d\omega_0, dy_{1:N_1}^1, dy_{1:N_2}^2), \end{aligned} \tag{3.6}$$

where μ^N is the joint probability measure of $\omega_0, dy_{1:N_1}^1, dy_{1:N_2}^2$ and

$$c_N^i(\underline{\gamma}, \underline{y}, \omega_0) := \int c^i(\omega_0, u_{1:N_i}^1, u_{1:N_i}^2) \prod_{i=1}^2 \prod_{k=1}^{N_i} \gamma_k^i(du_k^i | y_k^i).$$

In the following theorem, we establish existence of an NE for $\mathcal{P}_N$. The proof of the theorem is provided in Appendix A.2.

THEOREM 3.4. *Consider the game $\mathcal{P}_N$ with a given IS. Let Assumptions* 2.2 *hold. Suppose further that $\mathbb{U}^i$ is compact and $c^i(\omega_0, \cdot, \dots, \cdot)$ is continuous and (uniformly) bounded for all ω_0. Then there exists a randomized policy profile $(P_\pi^{1*}, P_\pi^{2*}) \in L_{\mathsf{CO},N}^1 \times L_{\mathsf{CO},N}^2$ that constitutes an NE for $\mathcal{P}_N$.*

Note that allowing common and independent randomness in randomization of policies leads to the convexity of $L_{\mathsf{CO},N}^i$. Since $L_{\mathsf{CO},N}^i$ and L_N^i are identical, $L_{\mathsf{CO},N}^i$ is closed. As a result, continuity of the costs together with the compactness of the action spaces allows us to utilize the Kakutani–Fan–Glicksberg fixed point theorem [2, Corollary 17.55]. A similar analysis has been used for zero-sum games among two teams in [19] to establish the existence of a randomized saddle-point equilibrium with common and independent randomness. Allowing common randomness in policies may be too restrictive due to the arbitrary nature of distributions and since randomness must be externally provided. As the number of DMs goes to infinity for $\mathcal{P}_\infty$, we relax this assumption, which further motivates the study of exchangeable games among teams. Additionally, in [19], it has been shown that the absolute continuity in Assumption 2.2 among all DMs can be relaxed for zero-sum games by instead imposing Assumption 2.2 only on the marginal distributions of observations among DMs within each team, fixing the policies of other teams. Hence, this is also true for the zero-sum version of $\mathcal{P}_N$ studied here.

3.4. Existence of an exchangeable Nash Equilibrium for exchangeable games $\boldsymbol{\mathcal{P}_N}$. In the following, we show that exchangeable stochastic games $\mathcal{P}_N$ (those $\mathcal{P}_N$ with exchangeable costs as in Assumption 2.1) admits a (finite) exchangeable NE, belonging to $L_{\mathsf{EX},N}^1 \times L_{\mathsf{EX},N}^2$ (this NE is not necessarily symmetric and independent). We first introduce the following assumptions on DMs' action spaces, costs, and observations.

Assumption 3.1. For all $i \in \{1,2\}$,
(i) $\mathbb{U}^i$ is compact;
(ii) $c^i(\omega_0,\cdot,\ldots,\cdot)$ is continuous and (uniformly) bounded for all ω_0;
(iii) $(y_k^i)_{k\in\mathcal{N}_i}$ are N_i-exchangeable, conditioned on ω_0;
(iv) $(y_k^1)_{k\in\mathcal{N}_1}$ and $(y_k^2)_{k\in\mathcal{N}_2}$ are mutually independent, conditioned on ω_0.

We note that since i.i.d. random variables are (infinitely) exchangeable, Assumption 2.3 implies Assumptions 3.1 (iii) and (iv). The following theorem establishes existence of an exchangeable NE of $\mathcal{P}_N$. The proof of the theorem is provided in Appendix A.3.

THEOREM 3.5. *Consider the game $\mathcal{P}_N$ with a given IS. Let Assumptions* 2.1*,* 2.2*, and* 3.1 *hold. Then there exists an exchangeable randomized policy profile $(P_\pi^{1*},P_\pi^{2*}) \in L^1_{\mathsf{EX},N}\times L^2_{\mathsf{EX},N}$ that constitutes an NE for $\mathcal{P}_N$ among all policy profiles in $L_N^1\times L_N^2$.*

Since the costs are continuous by Assumption 3.1, the convexity and compactness of $L^i_{\mathsf{EX},N}$ established in Lemma 3.3 allow us to invoke the Kakutani–Fan–Glicksberg fixed point theorem [2, Corollary 17.55] to show the existence of an NE among all exchangeable policy profiles in $L_N^1\times L_N^2$. Then we use the exchangeability of the costs and observations in Assumption 3.1 to show that we can restrict the search for finding the best response policies to a fixed exchangeable policy of another player to exchangeable policies (those belong to $L^1_{\mathsf{EX},N}\times L^2_{\mathsf{EX},N}$).

Two remarks are in order: First, Theorem 3.5 only guarantees the existence of an exchangeable NE which might not be symmetric and independent since not all finite exchangeable random variables are i.i.d. (see Lemma 3.2 (ii)). We cannot guarantee the existence of a symmetric NE since restricting to symmetric independent policies in finding the best response policies might result in a loss for the DMs within the player. This is because fixing policies of the other player to symmetric or exchangeable policies, DMs within a deviating player face a team problem that might not admit a symmetric independently randomized optimal solution; e.g., see [32, Example 1]. Second, in contrast to [19], in Theorem 3.5, our focus is on exchangeable games among teams, and we establish exchangeability of an NE which still might require a common randomness among DMs with teams.

3.5. Existence of a symmetric Nash equilibrium for $\mathcal{P}_\infty$. The expected costs for $\mathcal{P}_N$ with the mean-field interaction under a randomized policy profile (P_π^1,P_π^2) $\in L_N^1\times L_N^2$ are given by

$$J^i_{\pi,N}(P_\pi^1,P_\pi^2) := \int P_\pi^1(d\underline{\gamma}^1)P_\pi^2(d\underline{\gamma}^2)\mu^N(d\omega_0,d\underline{y})c_N^i(\underline{\gamma},\underline{y},\omega_0) \tag{3.7}$$

with

$$c_N^i(\underline{\gamma},\underline{y},\omega_0)$$
$$:= \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0,u_k^i,\Xi^1\left(\frac{1}{N_1}\sum_{p=1}^{N_1}\delta_{u_p^1}\right),\Xi^2\left(\frac{1}{N_2}\sum_{p=1}^{N_2}\delta_{u_p^2}\right)\right)\prod_{i=1}^{2}\prod_{k=1}^{N_i}\gamma_k^i(du_k^i|y_k^i).$$

We note that the above cost is a special case of $\mathcal{P}_N$ introduced in (3.6) since we have a special structure for the cost function c_N^i which satisfies Assumption 2.1.

Next, the expected costs for $\mathcal{P}_\infty$ under a randomized policy profile $(P_\pi^1,P_\pi^2)\in L^1\times L^2$ are given by

$$J^i_{\pi,\infty}(P_\pi^1,P_\pi^2) := \limsup_{N_1,N_2\to\infty}\int P^1_{\pi,N_1}(d\underline{\gamma}^1)P^2_{\pi,N_2}(d\underline{\gamma}^2)\mu^N(d\omega_0,d\underline{y})c_N^i(\underline{\gamma},\underline{y},\omega_0), \tag{3.8}$$

where P^i_{π,N_i} is the marginal of $P^i_\pi \in L^i$ to the first N_i components, and μ^N is the marginal of the fixed probability measure on $(\omega_0, y^1_{1:N_1}, y^2_{1:N_2})$.

The following theorem establishes the existence of an NE for $\mathcal{P}_\infty$ that is symmetric and independent. The proof of the theorem is provided in Appendix A.4.

THEOREM 3.6. *Consider the game $\mathcal{P}_\infty$ with a given IS. Let Assumptions* 2.2, 2.3, *and* 2.4 *hold. Then there exists an independently randomized symmetric policy profile $(P^{1*}_\pi, P^{2*}_\pi) \in L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$ that constitutes an NE for $\mathcal{P}_\infty$ among all policy profiles in $L^1 \times L^2$.*

We now shed light on the proof of Theorem 3.6. First, since observations are conditionally i.i.d., and the costs are continuous (by Assumptions 2.3 and 2.4), by restricting the search to policies belonging to $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$, we show that without any loss we can consider a representative DM for each player (team), and then we use the Kakutani–Fan–Glicksberg fixed point theorem to show that there exists a symmetric NE for $\mathcal{P}_\infty$ among policies belonging to $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$. Then we use a de Finetti representation theorem, an argument used in [32, Lemma 2] based on the de Finetti theorem for finite exchangeable random variables, and lower semicontinuity of the expected cost functions to show that for $\mathcal{P}_\infty$ we can restrict the search in finding best response policies to fixed symmetric policies of the other player over only symmetric policies. This guarantees that an independently randomized symmetric NE for $\mathcal{P}_\infty$ constitutes an NE among all policies belonging to $L^1 \times L^2$.

Unlike Theorem 3.5 where we only established the existence of exchangeable randomized policies, in Theorem 3.6, we established the existence of a symmetric randomized NE with only independent randomness. The reason is that the set of randomized policies with only independent randomness is not a convex subset of L^i_N, and hence the Kakutani–Fan–Glicksberg fixed point theorem does not apply. However, in the limit as the number of DMs goes to infinity, by the optimality results, we can consider a single DM as a representative DM for each team with a countably infinite number of DMs, who all apply the same policy as the representative DM. This allows us to use the Kakutani–Fan–Glicksberg fixed point theorem on the corresponding set of randomized policies with independent randomization (for a single DM), which is convex and compact.

3.6. Approximations of a symmetric Nash equilibrium for $\mathcal{P}_N$. By Theorem 3.5, there exists an exchangeable NE for exchangeable games $\mathcal{P}_N$ (those with exchangeable costs in Assumption 2.1), e.g., those with the mean-field interaction. This NE is not necessarily symmetric and varies by the number of DMs in the game. On the other hand, Theorem 3.6 established the existence of an NE for $\mathcal{P}_\infty$ that is symmetric and independent. In view of this, we address the following question: *Does there exist a scalable (with respect to the number of DMs) symmetric independently randomized approximate NE for $\mathcal{P}_N$ with the mean-field interaction?* We answer this question in the affirmative by showing that a symmetric independently randomized NE for $\mathcal{P}_\infty$ constitutes an approximate NE for $\mathcal{P}_N$. The proof of the following theorem is provided in Appendix A.5.

THEOREM 3.7. *Consider the games $\mathcal{P}_N$ and $\mathcal{P}_\infty$ with a given IS. Let Assumptions* 2.2, 2.3, *and* 2.4 *hold. Then an independently randomized symmetric NE for $\mathcal{P}_\infty$ constitutes an $\epsilon_{N_{1:2}}$-NE for $\mathcal{P}_N$ among all policies in $L^1_N \times L^2_N$, where $\epsilon_{N_{1:2}} \to 0$ as $N_1, N_2 \to \infty$.*

Theorem 3.7 is a by-product of Theorem 3.5 and our analysis in the proof of Theorem 3.6. We first show that a symmetric independently randomized NE for $\mathcal{P}_\infty$

constitutes an $\epsilon_{N_{1:2}}$-NE for $\mathcal{P}_N$ among symmetric independent ones. Then, as an implication of analysis developed in the proof of Theorem 3.6, using Theorem 3.5 and Lemma 3.2, we establish that a symmetric independent NE for $\mathcal{P}_\infty$ constitutes an $\epsilon_{N_{1:2}}$-NE for $\mathcal{P}_N$ among all randomized policies.

4. Exchangeable dynamic games with symmetric information structures. In this section, we present our results for exchangeable dynamic stochastic games. We first introduce sets of randomized policies equipped with a suitable topology. Then we establish the existence of an NE for $\mathcal{P}_N^T$ that is exchangeable, and an NE for $\mathcal{P}_\infty^T$ that is symmetric and independent.

4.1. Topology on control policies for dynamic games. We first introduce two reduction conditions as variations of Assumption 2.2 for static games, adapted to the multistage dynamic game setting with a given IS, that allow us to define randomized policies for dynamic games as Borel spaces. Let $\zeta_{k,t}^i := (w_{k,t}^i, v_{k,t}^i)$ with $\zeta_{k,-1}^i := x_{k,0}^i$, $\underline{\zeta}_t^i := (\zeta_{1,t}^i, \dots, \zeta_{k,t}^i)$, $\underline{\zeta}_{-1:t-1}^{1:2} = (\underline{\zeta}_{-1}^1, \dots, \underline{\zeta}_{t-1}^1, \underline{\zeta}_{-1}^2, \dots, \underline{\zeta}_{t-1}^2)$, and $\underline{\boldsymbol{\zeta}}^{\boldsymbol{i}} := \underline{\zeta}_{-1:T-1}^i$. Also, define the information history at time t for any $t \in \{1, \dots, T-1\}$ by

$$H_t := \left\{\omega_0, \underline{\zeta}_{-1:t-1}^{1:2}, \underline{y}_{0:t-1}^{1:2}, \underline{u}_{0:t-1}^{1:2}\right\}. \tag{4.1}$$

Assumption 4.1. For every $N \in \mathbb{N} \cup \{\infty\}$, let ν_t^N be the distribution of $\underline{y}_t^{1:2}$, conditioned on the information history H_t in (4.1). One of the following conditions holds for every $N \in \mathbb{N} \cup \{\infty\}$ and for any $t \in \{0, \dots, T-1\}$:

(i) (Independent reduction) There exist a probability measure $\tau_{k,t}^i$ and a function $\psi_{k,t}^i$ such that for all Borel sets A_k^i on $\mathbb{Y}^i$ (with $A = \prod_{i=1}^2 A_1^i \times \cdots \times A_{N_i}^i$)

$$\nu_t^N(A|H_t) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{A_k^i} \psi_{k,t}^i(y_{k,t}^i, H_t) \tau_{k,t}^i(dy_{k,t}^i).$$

(ii) (Nested reduction) There exist a probability measure $\tilde{\eta}_{k,t}^i$ and a function $\phi_{k,t}^i$ such that for all Borel sets A_k^i on $\mathbb{Y}^i$

$$\nu_t^N(A|H_t) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{A_k^i} \phi_{k,t}^i(y_{k,t}^i, H_t) \tilde{\eta}_{k,t}^i(dy_{k,t}^i | H_{k,t}^i)$$

with $H_{k,t}^i := \zeta_{k,-1:t-1}^i, y_{k,0:t-1}^i, u_{k,0:t-1}^i$, and for each DM_k^i over time, there exists a static reduction under which the observations $y_{k,t}^i$ are expanding and distributed by fixed probability measures $\eta_{k,t}^i$, i.e., $\sigma(y_{k,t}^i) \subset \sigma(y_{k,t+1}^i)$.

For examples under which reductions in Assumption 4.1 hold, see [32, p. 15] and [30]. Let $\mathbb{T} := \{0, \dots, T-1\}$. Using Assumption 4.1 (i), we equip randomized policies with a topology under which the convergence in policies is defined by

$$\boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i,n}} \xrightarrow[n\to\infty]{} \boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i}} \iff \gamma_{k,t}^{i,n}(du_{k,t}^i|y_{k,t}^i)\tau_{k,t}^i(dy_{k,t}^i) \xrightarrow[n\to\infty]{\text{weakly}} \gamma_{k,t}^i(du_{k,t}^i|y_{k,t}^i)\tau_{k,t}^i(dy_{k,t}^i) \ \forall t \in \mathbb{T}.$$

Similarly, under Assumption 4.1(ii), the topology on randomized policies yields that

$$\boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i,n}} \xrightarrow[n\to\infty]{} \boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i}} \iff \gamma_{k,t}^{i,n}(du_{k,t}^i|y_{k,0:t}^i)\eta_{k,t}^i(dy_{k,0:t}^i) \xrightarrow[n\to\infty]{\text{weakly}} \gamma_{k,t}^i(du_{k,t}^i|y_{k,0:t}^i)\eta_{k,t}^i(dy_{k,0:t}^i).$$

That is, under Assumption 4.1, by considering $\boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i}}$, we can define sets of randomized policies $L_N^i, L_{\mathsf{CO},N}^i, L_{\mathsf{EX},N}^i, L_{\mathsf{CO,SYM},N}^i, L^i, L_{\mathsf{CO}}^i, L_{\mathsf{EX}}^i, L_{\mathsf{CO,SYM}}^i, L_{\mathsf{PR}}^i$, and $L_{\mathsf{PR,SYM}}^i$ similarly to those in section 3.1 adapted to the dynamic setting.

4.2. Existence of an exchangeable Nash equilibrium for $\mathcal{P}_N^T$. Here, we establish the existence of an exchangeable NE for exchangeable games more general than $\mathcal{P}_N^T$. The IS is given by $I_{k,t}^i := \{y_{k,t}^i\}$, where observations are symmetric and given by

$$y_{k,t}^i = h_t^i(\zeta_{k,-1:t}^i, u_{k,0:t-1}^i, M_{-k,0:t-1}^{1:2}), \tag{4.2}$$

where the functions h_t^i are identical for all $k \in \mathcal{N}_i$. In the above, we used $M_{-k,0:t-1}^{1:2} := (\zeta_{-k,-1:t}^i, u_{-k,0:t-1}^i, \underline{\zeta}_{-1:t}^{-i}, \underline{u}_{0:t-1}^{-i})$ to denote the information up to time t excluding $(\zeta_{k,-1:t}^i, u_{k,0:t-1}^i)$. We note that the ISs of $\mathcal{P}_N^T$ and $\mathcal{P}_\infty^T$, given by (2.9) and (2.10), are symmetric and correspond to special cases of (4.2).

The PL^i's expected cost under a randomized policy profile $(P_\pi^1, P_\pi^2) \in L_N^1 \times L_N^2$ are given by

$$\begin{aligned} J_{\pi,N}^{i,T}(P_\pi^1, P_\pi^2) &:= \int P_\pi^1(d\underline{\gamma}^{\mathbf{1}})P_\pi^2(d\underline{\gamma}^{\mathbf{2}})\mu^N(d\omega_0, d\underline{\zeta})c_N^i(\underline{\gamma}, \underline{\zeta}, \underline{y}, \omega_0)\nu^N(d\underline{y}|\underline{\zeta}, \underline{\gamma}, \omega_0) \\ &:= \int \left(\int c^i(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}, \underline{\boldsymbol{u}}^{\mathbf{1:2}}) \prod_{i=1}^{2}\prod_{k=1}^{N_i} \boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i}}(d\boldsymbol{u}^{\boldsymbol{i}}|\boldsymbol{y}^{\boldsymbol{i}})\right) \\ &\quad \times \prod_{i=1}^{2} P_\pi^i(d\underline{\gamma}^{\boldsymbol{i}})\mu^N(d\omega_0, d\underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}) \prod_{t=0}^{T-1} \nu_t^N\left(d\underline{y}_t^{1:2}|H_t\right), \end{aligned} \tag{4.3}$$

where μ^N denotes the joint distribution of $(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}})$, and

$$c_N^i(\underline{\gamma}, \underline{\zeta}, \underline{y}, \omega_0) := \int c^i(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}, \underline{\boldsymbol{u}}^{\mathbf{1:2}}) \prod_{i=1}^{2}\prod_{k=1}^{N_i} \boldsymbol{\gamma}_{\boldsymbol{k}}^{\boldsymbol{i}}(d\boldsymbol{u}^{\boldsymbol{i}}|\boldsymbol{y}^{\boldsymbol{i}}) \tag{4.4}$$

for some measurable function c^i satisfying the following exchangeability condition.

Assumption 4.2. c^i is (separately) exchangeable for any permutations σ and τ of $\{1, \dots, N_1\}$ and $\{1, \dots, N_2\}$, respectively, for all $\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}$

$$c^i\left(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}, \underline{\boldsymbol{u}}^{\mathbf{1:2}}\right) = c^i\left(\omega_0, \boldsymbol{\zeta}^{\mathbf{1}}_{\boldsymbol{\sigma(1):\sigma(N_i)}}, \boldsymbol{\zeta}^{\mathbf{2}}_{\boldsymbol{\tau(1):\tau(N_i)}}, \boldsymbol{u}^{\mathbf{1}}_{\boldsymbol{\sigma(1):\sigma(N_i)}}, \boldsymbol{u}^{\mathbf{2}}_{\boldsymbol{\tau(1):\tau(N_i)}}\right). \tag{4.5}$$

In addition, we let the following set of assumptions hold.

Assumption 4.3. For $i \in \{1, 2\}$,

(i) $\mathbb{U}^i$ is compact;
(ii) $c^i(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}, \cdot)$ in (4.4) is continuous and (uniformly) bounded for all $\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}$;
(iii) h_t^i in (4.2) is continuous in actions for $t = 0, \dots, T-1$;
(iv) $\underline{\boldsymbol{\zeta}}^{\mathbf{1:2}}$ are exchangeable, conditioned on ω_0;
(v) for all Borel sets A_k^i on $\mathbb{Y}^i$ (with $A^i = A_1^i \times \cdots \times A_{N_i}^i$)

$$\nu_t^N(A|H_t) = \prod_{i=1}^{2}\prod_{k=1}^{N_i} \nu_t^i\left(A^i|\omega_0, \zeta_{k,-1:t-1}^i, \underline{y}_{0:t-1}^{1:2}, \underline{u}_{0:t-1}^{1:2}\right) \quad \forall t \in \mathbb{T}.$$

The costs of $\mathcal{P}_N^T$ and $\mathcal{P}_\infty^T$, given in (2.11), satisfy Assumption 4.2. Also, for the state dynamics and observations of $\mathcal{P}_N^T$ and $\mathcal{P}_\infty^T$, Assumptions 2.5 and 2.6 imply Assumption 4.3. In the following theorem, we establish the existence of an exchangeable NE for exchangeable dynamic games $\mathcal{P}_N^T$. The proof of the theorem is provided in Appendix A.6.

THEOREM 4.1. *Consider the game $\mathcal{P}_N^T$ with a given symmetric IS. Let Assumptions 4.1, 4.2, and 4.3 hold. Then there exists an exchangeable policy profile $(P_\pi^{1*}, P_\pi^{2*}) \in L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$ that constitutes an NE for $\mathcal{P}_N^T$ among all policies in $L_N^1 \times L_N^2$.*

Similarly to Theorem 3.5, invoking the Kakutani–Fan–Glicksberg fixed point theorem [2, Corollary 17.55] establishes existence of an NE among all exchangeable policy profiles. Since the IS is symmetric, by Assumption 4.2 and Assumptions 4.3 (iv) and (v), we then show that restricting the search for finding the best response policies to exchangeable policies of others to exchangeable policies is without any loss.

4.2.1. Existence of a symmetric Nash equilibrium for $\mathcal{P}_\infty^T$. In this subsection, we establish existence of an independently randomized symmetric NE for $\mathcal{P}_\infty^T$. Define state dynamics and observations as (2.9) and (2.10), respectively. Let $I_{k,t}^i = \{y_{k,t}^i\}$.

Let PL^i's expected cost for $\mathcal{P}_N^T$ under a randomized policy profile $(P_\pi^1, P_\pi^2) \in L_N^1 \times L_N^2$ be given by

$$J_{\pi,N}^{i,T}(P_\pi^1, P_\pi^2) = \int P_\pi^1(d\underline{\boldsymbol{\gamma}}^{\mathbf{1}}) P_\pi^2(d\underline{\boldsymbol{\gamma}}^{\mathbf{2}}) \mu^N(d\omega_0, d\underline{\boldsymbol{\zeta}}) c_N^i(\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{y}}, \omega_0) \nu^N(d\underline{\boldsymbol{y}} | \underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\gamma}}, \boldsymbol{\omega_0}), \tag{4.6}$$

where

$$\begin{aligned} c_N^i(\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{y}}, \omega_0) := \int \frac{1}{N_i} \sum_{k=1}^{N_i} \sum_{t=0}^{T-1} c^i \Bigg(& \omega_0, x_{k,t}^i, u_{k,t}^i, \Xi_x^1 \left(\frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{x_{p,t}^1} \right), \\ & \Xi_x^2 \left(\frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,t}^2} \right), \Xi_u^1 \left(\frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,t}^1} \right), \Xi_u^2 \left(\frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,t}^2} \right) \Bigg) \\ & \times \prod_{i=1}^{2} \prod_{k=1}^{N_i} \boldsymbol{\gamma_k^i}(d\boldsymbol{u^i} | \boldsymbol{y^i}), \end{aligned}$$

where μ^N is the joint distribution of $(\omega_0, \underline{\boldsymbol{\zeta}}^{\mathbf{1:2}})$ and $\nu^N = \prod_{t=0}^{T-1} \nu_t^i$, with ν_t^i the joint distribution of $\boldsymbol{y_t^{1:2}}$ conditioned on H_t in (4.1).

Let PL^i's expected cost for $\mathcal{P}_\infty^T$ under a randomized policy profile $(P_\pi^1, P_\pi^2) \in L^1 \times L^2$ be given by

$$\begin{aligned} & J_{\pi,\infty}^{i,T}(P_\pi^1, P_\pi^2) \\ & \quad = \limsup_{N\to\infty} \int P_{\pi,N}^1(d\underline{\boldsymbol{\gamma}}^{\mathbf{1}}) P_{\pi,N}^2(d\underline{\boldsymbol{\gamma}}^{\mathbf{2}}) \mu^N(d\omega_0, d\underline{\boldsymbol{\zeta}}) c_N^i(\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{y}}, \omega_0) \nu^N(d\underline{\boldsymbol{y}} | \underline{\boldsymbol{\zeta}}, \underline{\boldsymbol{\gamma}}, \boldsymbol{\omega_0}), \end{aligned} \tag{4.7}$$

where $P_{\pi,N}^i$ is the marginal of $P_\pi^i \in L^i$ to its first N_i components.

Assumption 4.4. Assumption 4.1 holds with functions ψ_t^i and ϕ_t^i of the following forms for every $i \in \mathcal{N}$ and $t = 0, \ldots, T-1$:

$$\begin{aligned} & \psi_t^i \Bigg(y_{k,t}^i, \omega_0, \zeta_{k,-1:t-1}^i, y_{k,0:t-1}^i, u_{k,0:t-1}^i, \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}^1}, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}^2}, \\ & \qquad \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{x_{p,0:t-1}^1}, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,0:t-1}^2} \Bigg), \\ & \phi_t^i \left(y_{k,t}^i, \omega_0, \frac{1}{N_i} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}^1}, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}^2}, \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{x_{p,0:t-1}^1}, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,0:t-1}^2} \right), \end{aligned}$$

where $\psi^i_t(y^i_{k,t},\omega_0,\zeta^i_{k,-1:t-1},y^i_{k,0:t-1},\cdot,\dots,\cdot)$ and $\phi^i_t(y^i_{k,t},\omega_0,\cdot,\dots,\cdot)$ are continuous (weakly continuous in their four last arguments) and uniformly bounded for all $y^i_{k,t}$, ω_0, $\zeta^i_{k,-1:t-1}, y^i_{k,0:t-1}$.

In the following theorem, we establish the existence of a symmetric randomized NE for $\mathcal{P}^T_\infty$. The proof of the theorem is provided in Appendix A.7.

THEOREM 4.2. *Consider the game $\mathcal{P}^T_\infty$ with a given IS. Let Assumptions 2.5, 2.6, and 4.4 hold. Then there exists an independently randomized symmetric policy profile $(P^{1*}_\pi,P^{2*}_\pi)\in L^1_{\mathsf{PR,SYM}}\times L^2_{\mathsf{PR,SYM}}$ that constitutes an NE for $\mathcal{P}^T_\infty$ among all policies in $L^1\times L^2$.*

Theorem 4.2 follows from an argument similar to that used for Theorem 3.6 with an additional technical argument using continuity and exchangeability of the cost, state dynamics, and observations.

4.3. Approximations of a symmetric Nash equilibrium for $\mathcal{P}^T_N$. In the following, similarly to the static setting, we show that there exists an approximate independently randomized symmetric NE for $\mathcal{P}^T_N$ with the mean-field interaction.

THEOREM 4.3. *Consider the games $\mathcal{P}^T_N$ and $\mathcal{P}^T_\infty$ with a given IS. Let Assumptions 2.5, 2.6, and 4.4 hold. Then an independently randomized symmetric NE for $\mathcal{P}^T_\infty$ constitutes an $\epsilon_{N_{1:2}}$-NE for $\mathcal{P}^T_N$ among all policies in $L^1_N\times L^2_N$, where $\epsilon_{N_{1:2}}\to 0$ as $N_1,N_2\to\infty$.*

Proof. The proof follows from an argument similar to that used in the proof of Theorem 3.7. ∎

5. Conclusion. We have studied stochastic static and dynamic mean-field games among teams with a finite number of DMs and their mean-field limit with an infinite number of DMs. We have established existence of a randomized NE for exchangeable games with a finite number of DMs and have shown that this NE is exchangeable. For mean-field limiting games with an infinite number of DMs, we have established existence of a randomized NE and have shown that it is symmetric and independently randomized. Finally, we have established existence of an approximate independently randomized symmetric NE for games with a finite but large number of DMs using an independently randomized symmetric NE of the infinite population mean-field game.

Appendix A.

A.1. Proof of Lemma 3.3. Convexity of $L^i_{\mathsf{EX},N}$ follows from the fact that the convex combination of exchangeable probability measures remains exchangeable. Hence, the set $L^i_{\mathsf{EX},N}$ is a nonempty convex subset of the locally convex set L^i_N. Since $\mathbb{U}^i$ is compact, the marginal of probability measures on $\mathbb{U}^i$ is tight. Since the probability measure on $\mathbb{Y}^i$ is fixed, the marginal on $\mathbb{Y}^i$ is also tight. Since marginals are tight, the collection of all measures on $\mathbb{U}^i\times\mathbb{Y}^i$ with these tight marginals is also tight (see, e.g., [43, Proof of Theorem 2.4]), and hence the set Γ^i_k is tight for each $i\in\{1,2\}$ and $k\in\mathcal{N}_i$. Hence, $L^i_{\mathsf{EX},N}$ is tight for $i\in\{1,2\}$. Next, we show that $L^i_{\mathsf{EX},N}$ is closed under the weak convergence topology. Suppose that $P^{i,\infty}_\pi$ is the limit, in the weak convergence topology, of a converging sequence of randomized policies $\{P^{i,n}_\pi\}_n$ as $n\to\infty$. Also, suppose that $P^{\sigma,i,\infty}_\pi$ is the limit, in the weak convergence topology, of a converging sequence of randomized policies $\{P^{\sigma,i,n}_\pi\}_n$ as $n\to\infty$, where for $A^i_k\in\mathcal{B}(\Gamma^i_k)$ and for all $\sigma\in S_{N_i}$

$$P^{\sigma,i,n}_\pi(\gamma^i_1\in A^i_1,\dots,\gamma^i_{N_i}\in A^i_{N_i})=P^{i,n}_\pi(\gamma^i_{\sigma(1)}\in A^i_1,\dots,\gamma^i_{\sigma(N_i)}\in A^i_{N_i}).$$

Let $\mathcal{T}$ be a countable measure-determining subset of the set of all real-valued continuous functions on $\prod_{k=1}^{N_i}\Gamma_k^i$. For a function $f\in\mathcal{T}$, we have

$$\text{(A.1)}\qquad \int f(\gamma_{1:N_i}^i)dP_\pi^{\sigma,i,\infty}=\lim_{n\to\infty}\int f(\gamma_{1:N_i}^i)dP_\pi^{i,n}=\int f(\gamma_{1:N_i}^i)dP_\pi^{i,\infty},$$

where (A.1) follows from the fact that $P_\pi^{i,n}$ is N_i-exchangeability for every n, and from convergence of $\{P_\pi^{\sigma,i,n}\}_n$ and $\{P_\pi^{i,n}\}_n$ in the weak convergence topology. Since $\mathcal{T}$ is countable, for all $f\in\mathcal{T}$, we get (A.1). Since $\mathcal{T}$ is measure-determining, we get that $P_\pi^{\sigma,i,\infty}$ and $P_\pi^{i,\infty}$ are the same, and hence $L_{\mathsf{EX},N}^i$ is compact since it is tight. An argument similar to the above establishes that L_{EX}^i is convex and compact.

A.2. Proof of Theorem 3.4. We use the Kakutani–Fan–Glicksberg fixed point theorem [2, Corollary 17.55]. By [32, Theorem A.1], the set $L_{\mathsf{CO},N}^i$ and L_N^i are identical. This is because $L_{\mathsf{CO},N}^i$ is a convex set and the set of extreme points of L_N^i is a subset of $L_{\mathsf{CO},N}^i$. This implies that the set $L_{\mathsf{CO},N}^i$ is nonempty, convex, and closed. An tightness argument analogous to that in the proof of Lemma 3.3 using the fact that the action spaces are compact yields that $L_{\mathsf{CO},N}^i$ is compact since it is tight.

Define the best response correspondence $\Phi:\prod_{i=1}^2 L_{\mathsf{CO},N}^i\to 2^{\prod_{i=1}^2 L_{\mathsf{CO},N}^i}$ as $\Phi(P_\pi^1,P_\pi^2):=\prod_{i=1}^2\mathrm{BR}(P_\pi^{-i})$, where for $i=1,2$,

$$\text{(A.2)}\qquad \mathrm{BR}(P_\pi^{-i}):=\left\{P_\pi^i\in L_{\mathsf{CO},N}^i\,\middle|\,J_\pi^i(P_\pi^i,P_\pi^{-i})\le J_\pi^i(\hat{P}_\pi^i,P_\pi^{-i})\ \forall\hat{P}_\pi^i\in L_{\mathsf{CO},N}^i\right\}.$$

In the following, we show that Φ admits a fixed point, which will be denoted by (P_π^{1*},P_π^{2*}).

Next, we show that the graph

$$G=\left\{\left((P_\pi^1,P_\pi^2),\Phi(P_\pi^1,P_\pi^2)\right)\,\middle|\,(P_\pi^1,P_\pi^2)\in\prod_{i=1}^2 L_{\mathsf{CO},N}^i\right\}$$

is closed. Consider sequences of policies $\{P_\pi^{-i,n}\}_n\subseteq L_{\mathsf{CO},N}^{-i}$ and $\{P_\pi^{i,n}\}_n\subseteq L_{\mathsf{CO},N}^i$ that converge weakly to P_π^{-i} and P_π^i, respectively. If $P_\pi^{i,n}\in\mathrm{BR}(P_\pi^{-i,n})$, we get $P_\pi^i\in\mathrm{BR}(P_\pi^{-i})$ since

$$\begin{aligned}
&\text{(A.3)}\qquad &\inf_{\hat{P}_\pi^i\in L_{\mathsf{CO},N}^i}J_\pi^i(\hat{P}_\pi^i,P_\pi^{-i})&=\inf_{\hat{P}_\pi^i\in L_{\mathsf{CO},N}^i}\lim_{n\to\infty}J_\pi^i(\hat{P}_\pi^i,P_\pi^{-i,n})\\
&\text{(A.4)}\qquad &&\ge J_\pi^i(P_\pi^i,P_\pi^{-i})\\
&\text{(A.5)}\qquad &&\ge\inf_{\hat{P}_\pi^i\in L_{\mathsf{CO},N}^i}J_\pi^i(\hat{P}_\pi^i,P_\pi^{-i}).
\end{aligned}$$

Equality (A.3) follows from the generalized dominated convergence theorem for varying measures [37, Theorem 3.5] since the bounded costs c^i converge continuously,[2] and $\{P_\pi^{-i,n}\}_n$ converges weakly to P_π^{-i}. Inequality (A.4) follows from exchanging limit and infimum, $P_\pi^{i,n}\in\mathrm{BR}(P_\pi^{-i,n})$, and an argument similar to that used for (A.3), and (A.5) follows from the fact that $P_\pi^i\in L_{\mathsf{CO},N}^i$ (since $L_{\mathsf{CO},N}^i$ is closed). This implies that the graph G is closed. Moreover, since $L_{\mathsf{CO},N}^i$ is compact, and c are continuous in actions, using [18, Proposition D.5(b)], the map

$$F^i:P_\pi^{-i}\mapsto\inf_{\hat{P}_\pi^i\in L_{\mathsf{CO},N}^i}J_\pi^i(\hat{P}_\pi^i,P_\pi^{-i})$$

[2]We recall that the sequence $\{f_n\}_n$ converges continuously to f if and only if $f_n(a_n)\to f(a)$ whenever $a_n\to a$ as $n\to\infty$.

is continuous. This implies that $\mathrm{BR}(P_\pi^{-i})$ is nonempty. Thus, by the Kakutani–Fan–Glicksberg fixed point theorem, Φ admits a fixed point, and this completes the proof since the sets $L^i_{\mathsf{CO},N}$ and L^i_N are identical.

A.3. Proof of Theorem 3.5. By Lemma 3.3, the set $L^i_{\mathsf{EX},N}$ is a nonempty convex compact subset of locally convex set L^i_N. Hence, using Kakutani–Fan–Glicksberg fixed point theorem [2, Corollary 17.55] similar to the proof of Theorem 3.4, we show that there exists an NE $(P_\pi^{1*}, P_\pi^{2*}) \in L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$ among all policies in $L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$. We only need to show that (P_π^{1*}, P_π^{2*}) constitutes an NE among all randomized policies in $L^1_N \times L^2_N$. It is sufficient to show that for any $P_\pi^{-i*} \in L^{-i}_{\mathsf{EX},N}$, we have

$$\inf_{\hat{P}^i_\pi \in L^i_N} J^i_\pi(\hat{P}^i_\pi, P_\pi^{-i*}) = \inf_{\hat{P}^i_\pi \in L^i_{\mathsf{EX},N}} J^i_\pi(\hat{P}^i_\pi, P_\pi^{-i*}). \tag{A.6}$$

This follows from [32, Lemma 1] using Assumptions 2.1 and 3.1 (iii) and (iv) by first showing that for any policy $\hat{P}^i_\pi \in L^i_N$, we have that $J^i_\pi(\hat{P}^i_\pi, P_\pi^{-i*}) = J^i_\pi(\hat{P}^{\sigma,i}_\pi, P_\pi^{-i*})$, where $\hat{P}^{\sigma,i}_\pi$ is a σ-permutation of $\hat{P}^i_\pi$, and then by showing that an exchangeable policy constructed as an average of all possible permutations of P^i_π does not perform worst for PL^i. This implies (A.6), and hence, under Assumption 2.2, there exists an NE profile (P_π^{1*}, P_π^{2*}) belonging to $L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$. We have that every NE policy profile for the game under the reduction of Assumption 2.2 constitutes an NE policy profile for the original game [30, Theorem 3.1] since policies do not change under the reduction and the expected costs are the same, and ISs are equivalent. Hence, (P_π^{1*}, P_π^{2*}) constitutes NE for $\mathcal{P}_N$, and this completes the proof.

A.4. Proof of Theorem 3.6. The proof proceeds in two steps. In Step 1, we show that there exists a policy profile $(P^{1*}_{\pi,\infty}, P^{2*}_{\pi,\infty})$ that constitutes an NE for $\mathcal{P}_\infty$, among all randomized policies in $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$. In Step 2, we show that the policy profile $(P^{1*}_{\pi,\infty}, P^{2*}_{\pi,\infty})$ constitutes an NE for $\mathcal{P}_\infty$, among all randomized policies in $L^1 \times L^2$.

Step 1. We first restrict the search for an NE for $\mathcal{P}_\infty$ to randomized policies belonging to $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$. We note that under Assumptions 2.2 and 2.3 (see (2.8)), observations of DMs within players are i.i.d. and also independent of ω_0 using a change of measure argument in (2.8). Since our policy space is restricted to $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$, by the strong law of large numbers and the generalized dominated convergence theorem for varying measures [37, Theorem 3.5] since the bounded costs c^i converge continuously, we have for any (P^1_π, P^2_π) in $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$

$$\begin{aligned}
&\limsup_{N_1,N_2\to\infty} \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N_1}\sum_{k=1}^{N_1} u^1_k\right), \Xi^2\left(\frac{1}{N_2}\sum_{k=1}^{N_2} u^2_k\right)\right)\\
&\qquad \times P^i_\pi(d\underline{\gamma}^i) P^{-i}_\pi(d\underline{\gamma}^{-i}) \mathbb{P}^0(d\omega_0) \prod_{i=1}^{2}\prod_{k=1}^{N_i} \hat{f}^i(dy^i_k, \omega_0) Q^i(dy^i_k) \gamma^i_k(du^i_k|y^i_k)\\
&\quad = \int c^i\left(\omega_0, u^i_{\mathsf{R}}, \Xi^1\left(\Lambda^1(du^1_{\mathsf{R}})\right), \Xi^2\left(\Lambda^2(du^2_{\mathsf{R}})\right)\right)\Lambda^i(du^i_{\mathsf{R}})\\
&\qquad \times \mathbb{P}^0(d\omega_0)\prod_{s=1}^{2} \hat{f}^s\left(y^s_{\mathsf{R}}, \omega_0\right) Q^s(dy^s_{\mathsf{R}}) P^s_{\pi,\mathsf{R}}(d\gamma^s_{\mathsf{R}})\gamma^s_{\mathsf{R}}(du^s_{\mathsf{R}}|y^s_{\mathsf{R}})\\
&\quad := \bar{J}^i_{\pi,\infty}(P^i_{\pi,\mathsf{R}}, P^{-i}_{\pi,\mathsf{R}}, \Lambda^i, \Lambda^{-i}),
\end{aligned}$$

where the subindex R denotes the representative DM for each player i. In the above,

$$P^i_\pi(\underline{\gamma}^i \in \cdot) = \prod_{k\in\mathbb{N}} P^i_{\pi,\mathsf{R}}(\gamma^i_k \in \cdot) \tag{A.7}$$

for some $P^i_{\pi,\mathsf{R}}(\gamma^i_k \in \cdot)$ belonging to the set $\mathcal{P}(\Gamma^i_{\mathsf{R}})$, and

$$\Lambda^i(\cdot) = \mathcal{L}(u^i_{\mathsf{R}}). \tag{A.8}$$

We note that $\Lambda^i(\cdot)$ depends on the randomized policy $P^i_{\pi,\mathsf{R}}$ as the law of u^i_{R} depends on the selection of randomized policies. For simplicity we use the notation $\Lambda^i(\cdot)$ without explicitly emphasizing the dependency of $\Lambda^i(\cdot)$ on $P^i_{\pi,\mathsf{R}}$. In view of the above costs, a policy profile (P^1_π, P^2_π) of the forms (A.7) constitutes an NE for $\mathcal{P}_\infty$ if $P^i_{\pi,\mathsf{R}} \in \mathrm{BR}(P^{-i}_{\pi,\mathsf{R}}, \Lambda^1, \Lambda^2)$ for $i = 1, 2$ with

$$\begin{aligned} \mathrm{BR}(P^{-i}_{\pi,\mathsf{R}}, \Lambda^1, \Lambda^2) := \Bigg\{ & P^i_{\pi,\mathsf{R}} \in \mathcal{P}(\Gamma^i_{\mathsf{R}}) \Bigg| \\ & \bar{J}^i_{\pi,\infty}(P^i_{\pi,\mathsf{R}}, P^{-i}_{\pi,R}, \Lambda^i, \Lambda^{-i}) \leq \bar{J}^i_{\pi,\infty}(\hat{P}^i_{\pi,\mathsf{R}}, P^{-i}_{\pi,\mathsf{R}}, \Lambda^i, \Lambda^{-i}) \quad \forall \hat{P}^i_{\pi,\mathsf{R}} \in \mathcal{P}(\Gamma^i_{\mathsf{R}}) \Bigg\}, \end{aligned} \tag{A.9}$$

and the consistency condition (A.8) holds. We note that the above best response map is different from that defined in (A.2); this is because, in (A.9), each representative DM policy of a player is a best response to the policy of the representative DM of the other player and the mean-field terms of both players.

Define the best response correspondence $\Phi : \prod_{i=1}^2 \mathcal{P}(\Gamma^i_{\mathsf{R}}) \to 2^{\prod_{i=1}^2 \mathcal{P}(\Gamma^i_{\mathsf{R}})}$ as $\Phi = \Psi \circ \Theta$ with $\Theta(P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}}) := (\Lambda^1, \Lambda^2)$ satisfying the consistency condition (A.8), and $\Psi(\Lambda^1, \Lambda^2) = \prod_{i=1}^2 \mathrm{BR}(P^{-i}_{\pi,\mathsf{R}}, \Lambda^1, \Lambda^2)$. We have that $\mathcal{P}(\Gamma^i_{\mathsf{R}})$ is nonempty, compact, and convex. Next, we show that the graph

$$G = \left\{ \left((P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}}), \Phi(P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}})\right) \middle| (P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}}) \in \prod_{i=1}^2 \mathcal{P}(\Gamma^i_{\mathsf{R}}) \right\}$$

is closed.

Suppose that sequences $\{P^{1,n}_{\pi,\mathsf{R}}\}_n$ and $\{P^{2,n}_{\pi,\mathsf{R}}\}_n$ converge to $P^{1,\infty}_{\pi,\mathsf{R}}$ and $P^{2,\infty}_{\pi,\mathsf{R}}$, respectively. Let $P^{i,n}_{\pi,\mathsf{R}} \in \mathrm{BR}(P^{-i,n}_{\pi,\mathsf{R}}, \Lambda^1_n, \Lambda^2_n)$ for $i = 1, 2$ and $\Theta(P^{1,n}_{\pi,\mathsf{R}}, P^{2,n}_{\pi,\mathsf{R}}) := (\Lambda^1_n, \Lambda^2_n)$. Since the distribution of $u^{i,n}_{\mathsf{R}}$, induced by the policy $P^{i,n}_{\pi,\mathsf{R}}$, converges to the distribution of $u^{i,\infty}_{\mathsf{R}}$, induced by the policy $P^{i,\infty}_{\pi,\mathsf{R}}$, using the fact that $\Theta(P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}})$ is singleton, we get that Θ is continuous. Next, we show that Ψ is upper-hemicontinuous. We have, for $i = 1, 2$,

$$\begin{aligned} & \inf_{\hat{P}^i_{\pi,\mathsf{R}} \in \mathcal{P}(\Gamma^i_{\mathsf{R}})} \int c^i\left(\omega_0, u^i_{\mathsf{R}}, \Xi^i\left(\Lambda^i_\infty(du^i_{\mathsf{R}})\right), \Xi^{-i}\left(\Lambda^{-i}_\infty(du^{-i}_{\mathsf{R}})\right)\right) \Lambda^i_\infty(du^i_{\mathsf{R}}) \\ & \qquad \times \hat{P}^i_{\pi,\mathsf{R}}(d\gamma^i_{\mathsf{R}}) P^{-i,\infty}_{\pi,\mathsf{R}}(d\gamma^{-i}_{\mathsf{R}}) \mathbb{P}^0(d\omega_0) \prod_{s=1}^2 \hat{f}^s\left(y^s_{\mathsf{R}}, \omega_0\right) Q^s(dy^s_{\mathsf{R}}) \gamma^s_{\mathsf{R}}(du^s_{\mathsf{R}} | y^s_{\mathsf{R}}) \\ & = \inf_{\hat{P}^i_{\pi,\mathsf{R}} \in \mathcal{P}(\Gamma^i_{\mathsf{R}})} \lim_{n\to\infty} \int c^i\left(\omega_0, u^i_{\mathsf{R}}, \Xi^i\left(\Lambda^i_n(du^i_{\mathsf{R}})\right), \Xi^{-i}\left(\Lambda^{-i}_n(du^{-i}_{\mathsf{R}})\right)\right) \Lambda^i_n(du^i_{\mathsf{R}}) \qquad \text{(A.10)} \\ & \qquad \times \hat{P}^i_{\pi,\mathsf{R}}(d\gamma^i_{\mathsf{R}}) P^{-i,n}_{\pi,\mathsf{R}}(d\gamma^{-i}_{\mathsf{R}}) \mathbb{P}^0(d\omega_0) \prod_{s=1}^2 \hat{f}^s\left(y^s_{\mathsf{R}}, \omega_0\right) Q^s(dy^s_{\mathsf{R}}) \gamma^s_{\mathsf{R}}(du^s_{\mathsf{R}} | y^s_{\mathsf{R}}) \\ & \geq \int c^i\left(\omega_0, u^i_{\mathsf{R}}, \Xi^i\left(\Lambda^i_\infty(du^i_{\mathsf{R}})\right), \Xi^{-i}\left(\Lambda^{-i}_\infty(du^{-i}_{\mathsf{R}})\right)\right) \Lambda^i_\infty(du^i_{\mathsf{R}}) \qquad \text{(A.11)} \\ & \qquad \times P^{i,\infty}_{\pi,\mathsf{R}}(d\gamma^i_{\mathsf{R}}) P^{-i,\infty}_{\pi,\mathsf{R}}(d\gamma^{-i}_{\mathsf{R}}) \mathbb{P}^0(d\omega_0) \prod_{s=1}^2 \hat{f}^s\left(y^s_{\mathsf{R}}, \omega_0\right) Q^s(dy^s_{\mathsf{R}}) \gamma^s_{\mathsf{R}}(du^s_{\mathsf{R}} | y^s_{\mathsf{R}}), \end{aligned}$$

where (A.10) follows from the generalized dominated convergence theorem for varying measures since the bounded costs c^i converge continuously and the actions spaces are compact. Inequality (A.11) follows from exchanging limit and infimum and since $P^{i,n}_{\pi,\mathsf{R}} \in \mathrm{BR}(P^{-i,n}_{\pi,\mathsf{R}}, \Lambda^1_n, \Lambda^2_n)$, and along the same lines as (A.10). Hence, the above chain of inequalities becomes the chain of equalities, and this implies that G is closed. Moreover, since $\mathcal{P}(\Gamma^i_{\mathsf{R}})$ is compact and $\bar{J}^i_{\pi,\infty}$ is upper-hemicontinuous, we have that $\mathrm{BR}(P^{-i}_{\pi,\mathsf{R}}, \Lambda^1, \Lambda^2)$ is nonempty. The Kakutani–Fan–Glicksberg fixed point theorem completes the proof of this step.

Step 2. It suffices to show that for fixed $P^{-i*}_{\pi,\infty}$ and Λ^{-i}_{∞}, the following equality holds:

$$\inf_{P^i_\pi \in L^i_{\mathsf{PR,SYM}}} J^i_{\pi,\infty}(P^i_\pi, P^{-i*}_{\pi,\infty}) = \inf_{P^i_\pi \in L^i} J^i_{\pi,\infty}(P^i_\pi, P^{-i*}_{\pi,\infty}). \tag{A.12}$$

By fixing $P^{-i*}_{\pi,\infty}$, PL^i faces a mean-field team problem, and hence (A.12) essentially follows from an argument used in [32, Theorem 2] using [32, Lemma 2 and Theorem 1]. For completeness, we included some details of the proof.

Let $L^i_{\mathsf{EX}}|_N$ be the set of randomized policies in L^i_{EX} restricted to their first N_i components. By Lemma 3.2, we have $L^i_{\mathsf{EX}}|_N = L^i_{\mathsf{CO,SYM},N}$, but $L^i_{\mathsf{EX}}|_N \neq L^i_{\mathsf{EX},N}$. We have

$$\begin{aligned}
&\inf_{\hat{P}^i_\pi \in L^i} \limsup_{N_1,N_2\to\infty} \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N_1}\sum_{k=1}^{N_1} u^1_k\right), \Xi^2\left(\frac{1}{N_2}\sum_{k=1}^{N_2} u^2_k\right)\right)\\
&\qquad \times \hat{P}^i_\pi(d\underline{\gamma}^i) P^{-i*,\infty}_{\pi,\mathsf{R}}(d\underline{\gamma}^{-i})\mathbb{P}^0(d\omega_0)\prod_{i=1}^{2}\prod_{k=1}^{N_i} \hat{f}^i(dy^i_k,\omega_0)Q^i(dy^i_k)\gamma^i_k(du^i_k|y^i_k)\\
&\text{(A.13)}\\
&\quad \geq \limsup_{N_1,N_2\to\infty} \inf_{\hat{P}^i_\pi \in L^i_{\mathsf{EX},N}} \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N_1}\sum_{k=1}^{N_1} u^1_k\right), \Xi^2\left(\frac{1}{N_2}\sum_{k=1}^{N_2} u^2_k\right)\right)\\
&\qquad \times \hat{P}^i_\pi(d\underline{\gamma}^i) P^{-i*,\infty}_{\pi,\mathsf{R}}(d\underline{\gamma}^{-i})\mathbb{P}^0(d\omega_0)\prod_{i=1}^{2}\prod_{k=1}^{N_i} \hat{f}^i(dy^i_k,\omega_0)Q^i(dy^i_k)\gamma^i_k(du^i_k|y^i_k)\\
&\text{(A.14)}\\
&\quad = \limsup_{N_1,N_2\to\infty} \inf_{\hat{P}^i_\pi \in L^i_{\mathsf{EX}}|_{N_i}} \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N_1}\sum_{k=1}^{N_1} u^1_k\right), \Xi^2\left(\frac{1}{N_2}\sum_{k=1}^{N_2} u^2_k\right)\right)\\
&\qquad \times \hat{P}^i_\pi(d\underline{\gamma}^i) P^{-i*,\infty}_{\pi,\mathsf{R}}(d\underline{\gamma}^{-i})\mathbb{P}^0(d\omega_0)\prod_{i=1}^{2}\prod_{k=1}^{N_i} \hat{f}^i(dy^i_k,\omega_0)Q^i(dy^i_k)\gamma^i_k(du^i_k|y^i_k)\\
&\text{(A.15)}\\
&\quad = \limsup_{N_1,N_2\to\infty} \inf_{\hat{P}^i_\pi \in L^i_{\mathsf{PR,SYM}}|_{N_i}} \int \frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u^i_k, \Xi^1\left(\frac{1}{N_1}\sum_{k=1}^{N_1} u^1_k\right), \Xi^2\left(\frac{1}{N_2}\sum_{k=1}^{N_2} u^2_k\right)\right)\\
&\qquad \times \hat{P}^i_\pi(d\underline{\gamma}^i) P^{-i*,\infty}_{\pi,\mathsf{R}}(d\underline{\gamma}^{-i})\mathbb{P}^0(d\omega_0)\prod_{i=1}^{2}\prod_{k=1}^{N_i} \hat{f}^i(dy^i_k,\omega_0)Q^i(dy^i_k)\gamma^i_k(du^i_k|y^i_k)\\
&\text{(A.16)}\\
&\quad \geq \inf_{P^i_\pi \in L^i_{\mathsf{PR,SYM}}} J^i_{\pi,\infty}(P^i_\pi, P^{-i*}_{\pi,\infty}),
\end{aligned}$$

where (A.13) follows from exchanging limsup and infimum and from Step 2 of the proof of Theorem 3.5. This is because, by fixing the policy of PL^{-i}, PL^i faces a team problem with an exchangeable cost function and IS. Hence, without loss of generality, we can restrict our search space to exchangeable policies. Equality (A.14) follows from [32, Lemma 2] for teams, showing that in the limit, without loss of optimality, the search for (globally) optimal policies of team problems can be restricted to the set of infinitely exchangeable randomized policies. In other words, although $L^i_{\mathsf{EX}}|_{N_i} \subset L^i_{\mathsf{EX},N}$, in the limit, the optimal costs become identical. Equality (A.15) follows from the de Finetti representation theorem for infinitely exchangeable policies [32, Theorem 1], and linearity of the expected cost in randomized policies $\hat{P}^i_\pi$ since $L_{\mathsf{PR,SYM}}|_{N_i}$ contains the set of extreme points of $L_{\mathsf{CO,SYM}}|_{N_i}$. Next, we show that (A.16) holds.

This is because, by fixing the policy of PL^{-i}, PL^i faces a team problem, and hence, since the cost function is continuous and actions space is compact, by [44, Theorem 5.2], a minimizer of the expected cost in (A.15) for every finite N_1 and N_2 exists. Denote such an optimal policy by $\{\tilde{P}^i_{\pi,N_{1,2}}\}_{N_{1,2}} \subseteq L^i_{\mathsf{PR,SYM}}|_{N_i}$. Since $L^i_{\mathsf{PR,SYM}}|_{N_i}$ is compact, there exists a subsequence $\{\tilde{P}^i_{\pi,n}\}_n \subseteq L^i_{\mathsf{PR,SYM}}|_{N_i}$ with index $n \in \mathbb{I} \subseteq \mathcal{N}_i$ that converges to $\tilde{P}^i_{\pi,\infty} \in L^i_{\mathsf{PR,SYM}}$. Hence, compactness of $L^i_{\mathsf{PR,SYM}}$ allows us to extract subsequences of policies $\{\tilde{P}^1_{\pi,n}\}_n$ and $\{\tilde{P}^2_{\pi,n}\}_n$ that converge to $\tilde{P}^1_{\pi,\infty}$ and $\tilde{P}^2_{\pi,\infty}$, respectively. Define the empirical measures of actions as

$$\tilde{\Lambda}^i_n(\cdot) := \frac{1}{n}\sum_{k=1}^{n} \delta_{\{\tilde{u}^i_{k,n}\}}(\cdot), \qquad \tilde{\Lambda}^i_{n,\infty}(\cdot) := \frac{1}{n}\sum_{k=1}^{n} \delta_{\{\tilde{u}^i_{k,\infty}\}}(\cdot), \tag{A.17}$$

where actions $\tilde{u}^i_{k,n}$ and $\tilde{u}^i_{k,\infty}$ are induced by randomized policies $\tilde{P}^i_{\pi,n}$ and $\tilde{P}^i_{\pi,\infty}$, respectively. Similarly, we define $\Lambda^{-i}_{n,\infty}$, where the actions $u^{-i}_{k,\infty}$ are induced by randomized policies $P^{-i*,\infty}_{\pi,\mathsf{R}}$. By Assumption 2.3 and under the change of measure in (2.8) (see Assumption 2.2), the observations $(y^i_k)_{k\in\mathcal{N}_i}$ are i.i.d. Since randomized policies $\tilde{P}^i_{\pi,n}$ and $\tilde{P}^i_{\pi,\infty}$ belong to $L^i_{\mathsf{PR,SYM}}|_n$ and $L^i_{\mathsf{PR,SYM}}$, respectively, we can conclude that $\tilde{u}^i_{k,\infty}$, $\tilde{u}^i_{k,\infty}$, and $u^{-i}_{k,\infty}$ are i.i.d. Hence, by the strong law of large numbers, along the same lines as in the proof of [33, Theorem 3.2] (see [33, equations (3.6)–(3.10)]), sequences $\{\tilde{\Lambda}^i_n\}_n$ and $\{\tilde{\Lambda}^i_{n,\infty}\}_n$ converge weakly to the same limit $\tilde{\Lambda}^i_\infty := \mathcal{L}(u^{i*}_{\mathsf{R},\infty})$. Also, $\Lambda^{-i}_{n,\infty}$ converges weakly to $\Lambda^{-i}_\infty := \mathcal{L}(u^{-i*}_{\mathsf{R},\infty})$. Using Assumption 2.4 and the generalized dominated convergence theorem, we get (A.16) and complete the proof.

A.5. Proof of Theorem 3.7. By Theorem 3.6, there exists an independently randomized symmetric NE $(P^{1*}_\pi, P^{2*}_\pi) \in L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$ for $\mathcal{P}_\infty$. Proceeding along the same lines as (A.15)–(A.16), we get

$$J^i_{\pi,\infty}(P^{1*}_\pi, P^{2*}_\pi) = \limsup_{N_1,N_2\to\infty} \inf_{P^i_{\pi,N} \in L^i_{\mathsf{PR,SYM}}|_{N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi) \quad \forall i \in \{1,2\}. \tag{A.18}$$

Following [44, Theorem 5.2], there exists a sequence of optimal policies as the best response to P^{-i*}_π, denoted by $\{\tilde{P}^i_{\pi,N_i}\}_{N_i}$. Again proceeding along the same lines as (A.15)–(A.16) with replacing the limsup with liminf, we get, for all $i \in \{1,2\}$,

$$\begin{aligned} \liminf_{N_1,N_2\to\infty} \inf_{P^i_{\pi,N} \in L^i_{\mathsf{PR,SYM}}|_{N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi) &\geq \inf_{P^i_\pi \in L^i_{\mathsf{PR,SYM}}} J^i_{\pi,\infty}(P^i_\pi, P^{-i*}_\pi) \\ &= J^i_{\pi,\infty}(P^{1*}_\pi, P^{2*}_\pi), \end{aligned} \tag{A.19}$$

where (A.19) follows from an argument similar to that used in (A.15)–(A.16). This together with (A.18) implies that

$$J^i_{\pi,\infty}(P^{1*}_\pi, P^{2*}_\pi) = \lim_{N_1,N_2\to\infty} \inf_{P^i_{\pi,N}\in L^i_{\mathsf{PR,SYM}}|_{N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi) \quad \forall i\in\{1,2\}. \tag{A.20}$$

By the strong law of large numbers since $(P^{1*}_\pi, P^{2*}_\pi)\in L^1_{\mathsf{PR,SYM}}\times L^2_{\mathsf{PR,SYM}}$, we get

$$J^i_{\pi,\infty}(P^{1*}_\pi, P^{2*}_\pi) = \lim_{N_1,N_2\to\infty} J^i_{\pi,N}(P^{1*}_\pi, P^{2*}_\pi) \quad \forall i\in\{1,2\}. \tag{A.21}$$

Equality (A.20) together with (A.21) implies that there exists $\hat{\epsilon}_{N_{1:2}}>0$, converging to zero as $N_1,N_2\to\infty$, such that

$$\left| J^i_{\pi,N}(P^{i*}_\pi, P^{-i*}_\pi) - \inf_{P^i_{\pi,N}\in L^i_{\mathsf{PR,SYM}}|_{N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi)\right| \le \hat{\epsilon}_{N_{1:2}} \quad \forall i\in\{1,2\}. \tag{A.22}$$

This implies that (P^{1*}_π, P^{2*}_π) constitutes an $\hat{\epsilon}_{N_{1:2}}$-NE for $\mathcal{P}_N$ among independently randomized symmetric policy profiles $L^1_{\mathsf{PR,SYM}}|_{N_1}\times L^2_{\mathsf{PR,SYM}}|_{N_2}$. Proceeding along the same lines as (A.13)–(A.16), we get, for $i\in\{1,2\}$,

$$\limsup_{N_1,N_2\to\infty} \inf_{P^i_{\pi,N}\in L^i_N} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi) = \limsup_{N_1,N_2\to\infty} \inf_{P^i_{\pi,N}\in L^i_{\mathsf{PR,SYM}}|_{N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi).$$

Hence, using (A.22), there exists $\epsilon_{N_{1:2}}>0$, converging to zero as $N_1,N_2\to\infty$ such that

$$\left| J^i_{\pi,N}(P^{i*}_\pi, P^{-i*}_\pi) - \inf_{P^i_{\pi,N}\in L^i_N} J^i_{\pi,N}(P^i_{\pi,N}, P^{-i*}_\pi)\right| \le \epsilon_{N_{1:2}} \quad \forall i\in\{1,2\}. \tag{A.23}$$

This implies that (P^{1*}_π, P^{2*}_π) constitutes an $\epsilon_{N_{1:2}}$-NE for $\mathcal{P}_N$ among all policy profiles in $L^1_N\times L^2_N$, and this completes the proof.

A.6. Proof of Theorem 4.1. The proof is similar to that of Theorem 3.5. We first use the Kakutani–Fan–Glicksberg fixed point theorem to show that there exists an NE (P^{1*}_π, P^{2*}_π) among all policies in $L^1_{\mathsf{EX},N}\times L^2_{\mathsf{EX},N}$. This follows from an argument identical to that in Step 1 of the proof of Theorem 3.5 using continuity of the costs and observations in actions in Assumptions 4.3 (ii) and (iii). Then we can show that equality (A.6) adapted for $\mathcal{P}^T_N$ holds using Assumptions 4.3 (iv) and (v) and an argument similar to [32, Lemma 1] (see [32, Lemma 3] for details of this step). For brevity, the details are not included.

A.7. Proof of Theorem 4.2. The proof proceeds along the same lines as those in the proof of Theorem 3.6. In the following, we only present some additional technical arguments required for the dynamic setting. We first restrict the search for an NE for $\mathcal{P}^T_\infty$ to randomized policies belonging to $L^1_{\mathsf{PR,SYM}}\times L^2_{\mathsf{PR,SYM}}$. By the strong law of large numbers and under Assumptions 2.5, 2.6, 4.1 (i), and 4.4, we can rewrite the expected costs $J^i_{\pi,\infty}(P^1_\pi, P^2_\pi)$ in (4.7) for any (P^1_π, P^2_π) in $L^1_{\mathsf{PR,SYM}}\times L^2_{\mathsf{PR,SYM}}$ as $\bar{J}^{i,T}_{\pi,\infty}(P^i_{\pi,\mathsf{R}}, P^{-i}_{\pi,\mathsf{R}}, \Lambda^i, \Lambda^{-i})$ with $P^i_\pi(\underline{\gamma}^{\boldsymbol{i}}\in\cdot) = \prod_{k\in\mathbb{N}} P^i_{\pi,\mathsf{R}}(\boldsymbol{\gamma}^{\boldsymbol{i}}_{\boldsymbol{k}}\in\cdot)$ for some $P^i_{\pi,\mathsf{R}}(\boldsymbol{\gamma}^{\boldsymbol{i}}_{\boldsymbol{k}}\in\cdot)$ belonging to the set $\mathcal{P}(\prod_{t=0}^{T-1}\Gamma^i_{\mathsf{R},t})$, and

$$\Lambda^i\left(\cdot\times\prod_{t=0}^{T-1}\mathbb{X}^i\times\mathbb{Y}^i\times\mathbb{Z}^i\right) = \mathcal{L}(\boldsymbol{u}^{\boldsymbol{i}}_{\mathsf{R}}), \qquad \Lambda^i\left(\cdot\times\prod_{t=0}^{T-1}\mathbb{U}^i\times\mathbb{Y}^i\times\mathbb{Z}^i\right) = \mathcal{L}(\boldsymbol{x}^{\boldsymbol{i}}_{\mathsf{R}}). \tag{A.24}$$

Similarly, we can rewrite the expected cost under Assumption 4.1 (ii). Similarly to the proof of Theorem 3.6, we define $\Phi:\prod_{i=1}^2\mathcal{P}(\prod_{t=0}^{T-1}\Gamma^i_{\mathsf{R},t})\to 2^{\prod_{i=1}^2\mathcal{P}(\prod_{t=0}^{T-1}\Gamma^i_{\mathsf{R},t})}$ as the best response correspondence $\Phi = \Psi \circ \Theta$ with $\Theta(P^1_{\pi,\mathsf{R}}, P^2_{\pi,\mathsf{R}}) := (\Lambda^1, \Lambda^2)$ satisfying the consistency condition (A.24), and $\Psi(\Lambda^1,\Lambda^2) = \prod_{i=1}^2 \mathrm{BR}(P^{-i}_{\pi,\mathsf{R}}, \Lambda^1, \Lambda^2)$. The set $\mathcal{P}(\prod_{t=0}^{T-1} \Gamma^i_{\mathsf{R},t})$ is nonempty, compact, and convex. Next, we show that the graph G is closed.

Suppose that sequences $\{P^{1,n}_{\pi,\mathsf{R}}\}_n$ and $\{P^{2,n}_{\pi,\mathsf{R}}\}_n$ converge to $P^{1,\infty}_{\pi,\mathsf{R}}$ and $P^{2,\infty}_{\pi,\mathsf{R}}$, respectively. Let $P^{i,n}_{\pi,\mathsf{R}} \in \mathrm{BR}(P^{-i,n}_{\pi,\mathsf{R}}, \Lambda^1_n, \Lambda^2_n)$ for $i = 1, 2$ and $\Theta(P^{1,n}_{\pi,\mathsf{R}}, P^{2,n}_{\pi,\mathsf{R}}) := (\Lambda^1_n, \Lambda^2_n)$. Since the distribution of $u^{i,n}_{\mathsf{R},0}$, induced by the policy $P^{i,n}_{\pi,\mathsf{R}}$, converges to the distribution of $u^{i,\infty}_{\mathsf{R}.0}$, induced by the policy $P^{i,\infty}_{\pi,\mathsf{R}}$, the marginal of Λ^i_n on $u^{i,n}_{\mathsf{R},0}$ converges to the marginal of Λ^i_∞ on $u^{i,\infty}_{\mathsf{R},0}$. By continuity of f^i_t, Ξ^i_x, and Ξ^i_u, we get that $x^{i,n}_{\mathsf{R},1}$ converges weakly to $x^{i,\infty}_{\mathsf{R}.1}$. Hence, marginals of Λ^i_n on $x^{i,n}_{\mathsf{R},1}$ converge to marginals of Λ^i_∞ on $x^{i,\infty}_{\mathsf{R},1}$. Hence, we can recursively show that Λ^i_n converges weakly to Λ^i_∞ for $i \in \{1, 2\}$ (and $\mathbb{P}^0$-a.s.). This leads to the continuity of Θ. Along the same lines as (A.10)–(A.11), using the generalized dominated convergence theorem, it can be shown that Ψ is upper-hemicontinuous. Hence, by the Kakutani–Fan–Glicksberg fixed point theorem, there exists a policy profile $(P^{1*}_{\pi,\infty}, P^{2*}_{\pi,\infty})$ that constitutes an NE for $\mathcal{P}^T_\infty$, among all randomized policies in $L^1_{\mathsf{PR},\mathsf{SYM}} \times L^2_{\mathsf{PR},\mathsf{SYM}}$. By fixing $P^{-i*}_{\pi,\infty}$ and Λ^{-i}_∞, PL^i faces a team problem, and hence [32, Theorem 4] implies that (A.12) holds. Hence, $(P^{1*}_{\pi,\infty}, P^{2*}_{\pi,\infty})$ constitutes an NE for $\mathcal{P}^T_\infty$, among all randomized policy profiles in $L^1 \times L^2$, and completes the proof.

REFERENCES

[1] D. J. Aldous, I. A. Ibragimov, and J. Jacod, *Ecole d'Ete de Probabilites de Saint-Flour XIII, 1983*, Lecture Notes in Math. 1117, Springer, 1985.

[2] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd ed., Springer, Berlin, 2006.

[3] V. Anantharam and V. S. Borkar, *Common randomness and distributed control: A counterexample*, Systems Control Lett., 56 (2007), pp. 568–572.

[4] A. Arapostathis, A. Biswas, and J. Carroll, *On solutions of mean field games with ergodic cost*, J. Math. Pures Appl., 107 (2017), pp. 205–251.

[5] M. Bardi and M. Fischer, *On non-uniqueness and uniqueness of solutions in finite-horizon mean field games*, ESAIM Control Optim. Calc. Var., 25 (2019), 44.

[6] M. Bardi and F. S. Priuli, *Linear-quadratic N-person and mean-field games with ergodic cost*, SIAM J. Control Optim., 52 (2014), pp. 3022–3052, https://doi.org/10.1137/140951795.

[7] E. Bayraktar and X. Zhang, *On non-uniqueness in mean field games*, Proc. Amer. Math. Soc., 148 (2020), pp. 4091–4106.

[8] C. Bertucci, J. Lasry, and P. Lions, *Some remarks on mean field games*, Comm. Partial Differential Equations, 44 (2019), pp. 205–227.

[9] V. S. Borkar, *White-noise representations in stochastic realization theory*, SIAM J. Control Optim., 31 (1993), pp. 1093–1102, https://doi.org/10.1137/0331050.

[10] P. Cardaliaguet and C. Rainer, *An example of multiple mean field limits in ergodic differential games*, NoDEA Nonlinear Differential Equations Appl., 27 (2020), pp. 1–19.

[11] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer, 2018.

[12] R. Carmona, F. Delarue, and D. Lacker, *Mean field games with common noise*, Ann. Probab., 44 (2016), pp. 3740–3803.

[13] A. Cecchin, P. D. Pra, M. Fischer, and G. Pelino, *On the convergence problem in mean field games: A two state model without uniqueness*, SIAM J. Control Optim., 57 (2019), pp. 2443–2466, https://doi.org/10.1137/18M1222454.

[14] F. Delarue and R. Tchuendom, *Selection of equilibria in a linear quadratic mean-field game*, Stochastic Process. Appl., 130 (2020), pp. 1000–1040.

[15] P. Diaconis and D. Freedman, *Finite exchangeable sequences*, Ann. Probab., (1980), pp. 745–764.

[16] M. Fischer, *On the connection between symmetric N-player games and mean field games*, Ann. Appl. Probab., 27 (2017), pp. 757–810.

[17] B. Hajek and M. Livesay, *On non-unique solutions in mean field games*, in 2019 IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 1219–1224.
[18] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer, 1996.
[19] I. Hogeboom-Burr and S. Yüksel, *Zero-sum games involving teams against teams: Existence of equilibria, and comparison and regularity in information*, Systems Control Lett., 172 (2023), 105454.
[20] M. Huang, P. E. Caines, and R. P. Malhamé, *Large population stochastic dynamic games: Closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–251.
[21] M. Huang, P. E. Caines, and R. P. Malhamé, *Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria*, IEEE Trans. Automat. Control, 52 (2007), pp. 1560–1571.
[22] O. Kallenberg, *Probabilistic Symmetries and Invariance Principles*, Springer Science & Business Media, 2006.
[23] D. Lacker, *Mean field games via controlled martingale problems: Existence of Markovian equilibria*, Stochastic Process. Appl., 125 (2015), pp. 2856–2894.
[24] D. Lacker, *A general characterization of the mean field limit for stochastic differential games*, Probab. Theory Related Fields, 165 (2016), pp. 581–648.
[25] D. Lacker, *On the convergence of closed-loop Nash equilibria to the mean field game limit*, Ann. Appl. Probab., 30 (2020), pp. 1693–1761.
[26] J. M. Lasry and P. L. Lions, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.
[27] B. Light and G. Weintraub, *Mean field equilibrium: Uniqueness, existence, and comparative statics*, Oper. Res., 70 (2022), pp. 585–605.
[28] N. Saldi, T. Başar, and M. Raginsky, *Markov–Nash equilibria in mean-field games with discounted cost*, SIAM J. Control Optim., 56 (2018), pp. 4256–4287, https://doi.org/10.1137/17M1112583.
[29] N. Saldi and S. Yüksel, *Geometry of information structures, strategic measures and associated stochastic control topologies*, Probab. Surv., 19 (2022), pp. 450–532.
[30] S. Sanjari, T. Basar, and S. Yüksel, *Isomorphism properties of optimality and equilibrium solutions under equivalent information structure transformations: Stochastic dynamic games and teams*, SIAM J. Control Optim., 61 (2023), pp. 3102–3130, https://doi.org/10.1137/22M1521936.
[31] S. Sanjari, N. Saldi, and S. Yüksel, *Nash equilibria for exchangeable team against team games and their mean field limit*, in American Control Conference (ACC), IEEE, 2023, pp. 1104–1109.
[32] S. Sanjari, N. Saldi, and S. Yüksel, *Optimality of independently randomized symmetric policies for exchangeable stochastic teams with infinitely many decision makers*, Math. Oper. Res., 48 (2023), pp. 1213–1809.
[33] S. Sanjari and S. Yüksel, *Optimal policies for convex symmetric stochastic dynamic teams and their mean-field limit*, SIAM J. Control Optim., 59 (2021), pp. 777–804, https://doi.org/10.1137/19M1284294.
[34] S. Sanjari and S. Yüksel, *Optimal solutions to infinite-player stochastic teams and mean-field teams*, IEEE Trans. Automat. Control, 66 (2021), pp. 1071–1086.
[35] M. Schäl, *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Z. Wahrscheinlichkeitsth, 32 (1975), pp. 179–296.
[36] M. Schervish, *Theory of Statistics*, Springer Science & Business Media, 2012.
[37] R. Serfozo, *Convergence of Lebesgue integrals with varying measures*, Sankhyā A, (1982), pp. 380–402.
[38] D. Tang, H. Tavafoghi, V. Subramanian, A. Nayyar, and D. Teneketzis, *Dynamic games among teams with delayed intra-team information sharing*, Dyn. Games Appl., 13 (2023), pp. 353–411.
[39] H. Witsenhausen, *The intrinsic model for discrete stochastic control: Some open problems*, in Control Theory, Numerical Methods and Computer System Modelling, Lecture Notes in Economics and Mathematical Systems 107, A. Bensoussan and J. L. Lions, eds., Springer-Verlag, 1975, pp. 322–335.
[40] H. Witsenhausen, *Equivalent stochastic control problems*, Math. Control Signals Systems, 1 (1988), pp. 3–11.
[41] L. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, Comptes Rendus de la Societe des Sci. et des Lettres de Varsovie, 30 (1937), pp. 212–234.

[42] X. Yu, Y. Zhang, and Z. Zhou, *Teamwise mean field competitions*, Appl. Math. Optim., 84 (2021), pp. 903–942.

[43] S. Yüksel, *On stochastic stability of a class of non-Markovian processes and applications in quantization*, SIAM J. Control Optim., 55 (2017), pp. 1241–1260, https://doi.org/10.1137/140984154.

[44] S. Yüksel, *A universal dynamic program and refined existence results for decentralized stochastic control*, SIAM J. Control Optim., 58 (2020), pp. 2711–2739, https://doi.org/10.1137/18M1221382.