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, 90E99, 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, $\text{DM}_k$ for $k \in \mathcal{N}_i$, acting as player $i$ (PL). In other words, PL encapsulates the collection of DMs indexed by $\mathcal{N}_i$. We denote DM$_k$ of PL by DM$_k$.
- There exists a collection of measurable spaces $\{(\Omega, \mathcal{F}), (\mathbb{U}_k, \mathcal{U}_k^i), (\mathbb{Y}_k^i, \mathcal{Y}_k^i)\}$ 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 $\prod_{k \in \mathcal{N}_i} \mathbb{Y}_k^i$ and $\prod_{k \in \mathcal{N}_i} \mathbb{U}_k^i$, respectively, for each PL.
- The $\mathbb{Y}_k^i$-valued observation variables are given by $y_k^i = h_k^i(\omega, \{u_p^k\}_{(s,p) \in L_k^i})$, where $h_k^i$ is a Borel measurable function and $L_k^i$ denotes the set of all DMs acting before DM$_k^i$ (i.e., $(s, p) \in L_k^i$ if DM$_k^i$ acts before DM$_k^i$).
- An admissible policy for each PL is denoted by $\gamma_k^i := \{\gamma_k^i\}_{k \in \mathcal{N}_i} \in \Gamma^i$ with $u_k^i = \gamma_k^i(y_k^i)$, and the set of admissible policies for each PL is described by the product space $\Gamma^i := \prod_{k \in \mathcal{N}_i} \Gamma_k^i$, where $\Gamma_k^i$ is the set of all Borel measurable maps from $\mathbb{Y}_k^i$ to $\mathbb{U}_k^i$.
- 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 $\{y_k^i | i \in \mathcal{M} \text{ and } k \in \mathcal{N}_i\}$ is classical if $y_k^i$ contains all of the information available to DM$_k^i$ who acts precedent to DM$_k^i$ (i.e., $(s, p) \in L_k^i$). An IS is quasi-classical (or partially nested) if whenever $u_p^k$ affects $y_k^i$ through the observation function $h_k^i$, $y_k^i$ contains $y_p^k$ (that is, $\sigma(y_k^i) \subset \sigma(y_k^i)$). An IS which is not partially nested is nonclassical.

We will also allow for randomized policies, where in addition to $y_k^i$, each DM$_k^i$ 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 approximate NEs for the corresponding prelimiting games with finite but large number of DMs. This is because, in order to utilize a de Finetti representation theorem, convexity and compactness of a set of exchangeable randomized policies are required. To address these questions for the number of DMs goes to infinity, we consider a class of exchangeable static games with a finite number of DMs, and then we introduce their limiting static mean-field games.

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., $U^i_k := U^i \subseteq \mathbb{R}^{n_i}$ and $Y^i_k := Y^i \subseteq \mathbb{R}^{m_i}$ 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 $\mathcal{P}_N$. 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 $N_i = \{1, \ldots, N_i\}$. Each PL $i$’s admissible policies $\gamma^i_N := (\gamma^i_1, \ldots, \gamma^i_{N_i})$ with $\Gamma^i_N := \prod_{k=1}^{N_i} \Gamma_k^i$ are measurable functions so that $u^i_k = \gamma^i_k(y^i_k)$ for $k \in N_i$. Let the expected cost function for each PL $i$ under a policy profile $\mathcal{\gamma}^{1:2}_N := (\gamma^1_N, \gamma^2_N)$ be given by

$$J^i_N(\mathcal{\gamma}^{1:2}_N) = \mathbb{E}^{\mathbb{X}^{1:2}}\left[c^i(\omega_0, y^i_N)\right]$$

for some Borel measurable cost function $c^i : \Omega_0 \times \prod_{j=1}^{2} \prod_{k=1}^{N_j} U^j \rightarrow \mathbb{R}_+$. We define $\omega_0$ as the $\Omega_0$-valued, cost function relevant, exogenous random variable, taking values from a Borel space $\Omega_0$ with its Borel $\sigma$-field $\mathcal{F}_0$. In the above, we used the notation $k : j$ to denote $\{k, \ldots, j\}$, and $U^i_N := u^i_{1:N_i}$, and $\mathbb{E}^{\mathbb{X}^{1:2}}$ to denote the expectation with respect to $\mathbb{P}$ when actions $u^i_k$ are induced by policies $\gamma^i_k$ given the private observations $y^i_k$ for each $DM^i_k$.

**Definition 2.1** ($\epsilon$-Nash equilibrium (NE) for $\mathcal{P}_N$). Let $\epsilon \geq 0$. A policy profile $\mathcal{\gamma}^{1:2}_N$ constitutes an $\epsilon$-NE if and only if the following inequalities hold for all $i \in \{1, 2\}$:
\begin{equation}
J_N^\gamma(\gamma^{1:2^}\infty) \leq \inf_{\gamma, \in \Gamma_N} J_N^\gamma(\gamma^{-\ast}, \gamma^{i}) + \epsilon,
\end{equation}

where \(-i := \{1, 2\}\setminus \{i\}\). If \(\epsilon = 0\), the policy profile \(\gamma^{1:2^}\infty\) 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 \(i_0\) and \(i \in \{1, 2\}\), i.e., for any permutations \(\sigma\) and \(\tau\) of \(\{1, \ldots, N_1\}\) and \(\{1, \ldots, N_2\}\), respectively,

\[c^i(\omega_0, u_1^{1:N_1}, u_2^{1:N_2}) = c^i(\omega_0, u_\sigma^{1}(1):\sigma(N_1), u_\tau^{1}(1):\tau(N_2)) \quad \forall \omega_0,
\]

where \(u_\sigma^{1}(1):\sigma(N_1) := u_1^1, \ldots, u_\sigma^{1}(N_1)\) and \(u_\tau^{1}(1):\tau(N_2) := u_2^1, \ldots, u_\tau^{2}(N_2)\).

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

\begin{equation}
c^i(\omega_0, u_1^{1:N_1}, u_2^{1:N_2}) = \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_k^1} \right), \Xi^2 \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_k^2} \right) \right)
\end{equation}

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

\begin{equation}
\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_k^1} \right), \Xi^2 \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_k^2} \right) \right)
\end{equation}

for every permutation \(\beta = \sigma\) or \(\beta = \tau\). Clearly, if \(c_k^i = \bar{c}^i\), then the above exchangeability condition holds. When the number of DMs goes to infinity, we consider the special case of \(c_k^i = \bar{c}^i\). We note that weak continuity of \(\Xi^i\) 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 \(\text{PL}^i\) involves a countably infinite number of DMs indexed by \(N_i = \mathbb{N}\). The set of admissible policies \(\gamma^i := (\gamma_1^i, \gamma_2^i, \ldots)\) for each \(\text{PL}^i\) is denoted by \(\Gamma^i := \prod_{k \in \mathbb{N}} \Gamma^i_k\), where each \(\gamma_k^i\) is a measurable function so that \(u_k^i = \gamma_k^i(y_k^i)\) for \(k \in N_i\). Let the expected cost of \(\text{PL}^i\) under a policy profile \(\gamma^{1:2^}\infty\) be given by
NASH EQUILIBRIA FOR TEAM AGAINST TEAM GAMES

\begin{equation}
J^i_\infty(\gamma^{1:2}) = \limsup_{N_1, N_2 \to \infty} \mathbb{E}^{1:2} \left[ \frac{1}{N_1} \sum_{k=1}^{N_1} c^i \left( \omega_0, u^i_k, \Xi^1 \left( \frac{1}{N_1} \sum_{p=1}^{N_1} \delta u^p_k \right), \Xi^2 \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta u^p \right) \right) \right]
\end{equation}

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

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

\begin{equation}
J^i_\infty(\gamma^{1:2:2*}) \leq J^i_\infty(\gamma^{-i, \ast}) \quad \forall \gamma^{-i} \in \Gamma^i.
\end{equation}

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 \( \bar{\mu}^N \) be the conditional distribution of observations \( y^i_N \) given \( \omega_0 \). There exists a probability measure \( Q^i_k \) on \( \mathbb{Y}^i \) and a bounded function \( f^i_k : \mathbb{Y}^i \times \Omega_0 \times \prod_{j=1, j \neq k}^{N_j} \mathbb{Y}^j \times \prod_{j=1}^{N_i} \mathbb{Y}^{-i} \to \mathbb{R}_+ \) 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 \( B := \prod_{i=1}^2 B^i_1 \times \cdots \times B^i_{N_i} \))

\begin{equation}
\bar{\mu}^N(B | \omega_0) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{B^i_k} f^i_k(y^i_k, \omega_0, y^{-i}_k, y^{-i}) Q^i_k(dy^i_k),
\end{equation}

where \( y^{-i}_k = \{y^i_1, \ldots, \hat{y}^i_k, \ldots, y^i_{N_i}\} \setminus \{y^i_k\} \).

Under the change of measure in \( (2.7) \), probability measures on observations \( y^i_k \) are \( Q^i_k \) that are independent of observations \( y^{-i}_k, y^{-i}, \) and also \( \omega_0 \). 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 \( \omega_0 \). 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 above assumptions on the observations, action spaces, and cost functions.

**Assumption 2.3.**

(i) \( \{y^i_k\}_{k \in \mathcal{N}_1} \) and \( \{y^i_k\}_{k \in \mathcal{N}_2} \) are independent, conditioned on \( \omega_0 \);

(ii) for all \( i \in \{1, 2\} \), \( \{y^i_k\}_{k \in \mathcal{N}_i} \) have an identical distribution, conditioned on \( \omega_0 \).

**Assumption 2.4.** For \( i \in \{1, 2\} \),

(i) \( U^i \) is compact;

(ii) \( c^i(\omega_0, \cdot, \cdot) \) in \( (2.3) \) is continuous for all \( \omega_0 \).

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

\begin{equation}
\bar{\mu}^N(B | \omega_0) = \prod_{i=1}^2 \prod_{k=1}^{N_i} \int_{B^i_k} \hat{f}^i(y^i_k, \omega_0) Q^i(dy^i_k),
\end{equation}

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 $t \in \{0, \ldots, T-1\}$ for simplicity, and subsets of appropriate dimensional Euclidean spaces $\mathbb{U}_k,t = \mathbb{U}^i \subseteq \mathbb{R}^{n_i}$, $\mathbb{Y}_k,t = \mathbb{Y}^i \subseteq \mathbb{R}^{n'_i}$, $\mathbb{X}_{k,t} = \mathbb{X}^i \subseteq \mathbb{R}^{n''_i}$, where $n_i$, $n'_i$, and $n''_i$ are positive integers.

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

\begin{align}
  x_{k,t+1}^i &= f_t^i \left( x_{k,t}^i, u_{k,t}^i, \Xi^i_x \left( \frac{1}{N_i} \sum_{p=1}^{N_i} \delta_{x_{p,t}^i} \right), \Xi^i_x \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,t}} \right) \right), \\
  \Xi^i_x &\left( \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,t}^i} \right), \Xi^i_x \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,t}} \right)
\end{align}

\begin{align}
  y_{k,t}^i &= h_t^i \left( x_{k,0:t}^i, u_{k,0:t-1}^i, v_{k,0:t}^i \right)
\end{align}

for some weakly continuous functions $\Xi^i_x : \mathcal{P}(\mathbb{X}^i) \rightarrow \tilde{\mathbb{X}}^i$ and $\Xi^i_u : \mathcal{P}(\mathbb{U}^i) \rightarrow \tilde{\mathbb{U}}^i$, where $\tilde{\mathbb{X}}^i$ and $\tilde{\mathbb{U}}^i$ are subsets of appropriate dimensional Euclidean spaces. In the above, functions $f_t^i$ and $h_t^i$ are measurable functions, and $v_{k,t}^i$ and $w_{k,t}^i$ 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 $\mathcal{P}_N^T$. Consider a stochastic dynamic game with a finite number of DMs. Each PL’s admissible policies $\gamma_{N,t}^i : \mathcal{X}^i_{N,t} \rightarrow \mathcal{S}(\mathcal{L}_{N,t})$ with $\gamma_{N,t}^i \equiv \gamma_{N,t}^i$ are measurable functions so that $u_{k,t}^i = \gamma_{k,t}^i(y_{k,t}^i)$ for all $k \in \mathcal{N}_i$ and $t \in \{0, \ldots, T-1\}$. We use bold letters for variables over time $t \in \{0, \ldots, T-1\}$, and underline variables to denote variables of DM $i$’s for $k \in \mathcal{N}_i$ within PL $i$. Denote the space of policies for each player with $\Gamma_N^i = \prod_{t=0}^{T-1} \prod_{k \in \mathcal{N}_i} \Gamma_{k,t}$. Let PL’s expected cost function under $\gamma_{N,1:T}^i \in \prod_{t=1}^{T} \Gamma_{N,t}$ be given by

\begin{align}
  J_{N}^{i,T} (\gamma_{N,1:T}^i) = \mathbb{E} \Xi^2 \left( \frac{1}{N_i} \sum_{k=1}^{N_i} \sum_{t=0}^{T-1} c_t \left( \omega_0, x_{k,t}^i, u_{k,t}^i, \Xi_x^i \left( \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{x_{p,t}^i} \right) \right) \\
  \Xi_x^i \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,t}} \right), \Xi_u \left( \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,t}} \right), \Xi_u \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,t}} \right) \right)
\end{align}

for some weakly continuous functions $\Xi^i_x : \mathcal{P}(\mathbb{X}^i) \rightarrow \tilde{\mathbb{X}}^i$ and $\Xi^i_u : \mathcal{P}(\mathbb{U}^i) \rightarrow \tilde{\mathbb{U}}^i$, where $\tilde{\mathbb{X}}^i$ and $\tilde{\mathbb{U}}^i$ are subsets of Euclidean spaces. In the above, $c_t^i : \Omega_0 \times \mathbb{X}^i \times \mathbb{U}^i \rightarrow \mathbb{R}_{\geq 0}$ is a Borel measurable cost function that satisfies
for any permutations $\beta = \sigma$ or $\beta = \tau$. Random variable $\omega_0$ is again a cost-related random variable. If $c^i_k = \bar{c}^i$, then (2.12) holds since
\begin{equation}
\sum_{k=1}^{N_i} c^i_k(\cdot, x^i_{k,t}, u^i_{k,t}, \ldots, \cdot) = \sum_{k=1}^{N_i} c^i_{\beta(k)}(\cdot, x^i_{\beta(k),t}, u^i_{\beta(k),t}, \ldots, \cdot)
\end{equation}
for any permutations $\beta = \sigma$ or $\beta = \tau$. Random variable $\omega_0$ is again a cost-related random variable. If $c^i_k = \bar{c}^i$, then (2.12) holds since
\begin{align}
\sum_{k=1}^{N_i} c^i_k(\omega_0, x^i_{k,t}, u^i_{k,t}, \xi^1_x \sum_{p=1}^{N_1} \delta_{x_{p,t}}, \ldots, \xi^2_u \sum_{p=1}^{N_2} \delta_{u_{p,t}}) \\
= \sum_{k=1}^{N_i} \bar{c}^i(\omega_0, x^i_{\beta(k),t}, u^i_{\beta(k),t}, \xi^1_x \sum_{p=1}^{N_1} \delta_{x_{\beta(p),t}}, \ldots, \xi^2_u \sum_{p=1}^{N_2} \delta_{u_{\beta(p),t}})
\end{align}

When the number of DMs goes to infinity, we consider the special case $c^i_k = \bar{c}^i$. Denote $\bar{c}^i := \sum_{k=1}^{N_i} c^i_k$. Consider a stochastic dy-
amic 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 $N_i = \mathbb{N}$. The set of admissible policies $\gamma^i_N := (\gamma^i_{0:T-1})$ for each $PL^i$ is denoted by $\Gamma^i := \prod_{t=0}^{T-1} \prod_{k=1}^{N_i} \Gamma^i_{k,t}$, where each $\gamma^i_{k,t}$ is a measurable function so that $u^i_{k,t} = \gamma^i_{k,t}(y^i_{k,t})$ for $k \in N_i$ and $t \in \{0, \ldots, T-1\}$. Let the expected cost of $PL^i$ under a policy profile $\gamma^i_N$ be given by
\begin{equation}
J^i_T(\gamma^i_N) = \sum_{k=1}^{N_i} c^i_k(\omega_0, x^i_{k,t}, u^i_{k,t}, \xi^1_x \sum_{p=1}^{N_1} \delta_{x_{p,t}}, \ldots, \xi^2_u \sum_{p=1}^{N_2} \delta_{u_{p,t}})
\end{equation}

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) $f^i(\cdot, \ldots, w^i_{k,t})$ and $h^i(\cdot, \ldots, v^i_{k,0:t})$ are continuous and uniformly bounded for all $(w^i_{k,t}, v^i_{k,0:t})$ and for $t \in \{0, \ldots, T-1\}$;
(ii) $c^i(\omega_0, \cdot, \cdot)$ is continuous and uniformly bounded for all $\omega_0$;
(iii) $U^i$ is compact.

Assumption 2.6. For $i \in \{1, 2\}$,
(i) $(x^i_{k,0})_{k \in N_i}$ are i.i.d., conditioned on $\omega_0$;
(ii) $(w^i_{k,t})_{k \in N_i}$ are i.i.d. for $t \in \{0, \ldots, T-1\}$, and $(w^i_{k,0:T-1})$ are mutually independent and independent of $\omega_0$ and $(x^i_{k,0:T-1})$; random variables $(v^i_{k,t})_{k \in N_i}$ are i.i.d. for $t \in \{0, \ldots, T-1\}$, and $(v^i_{k,0:T-1})$ are mutually independent, and independent of $\omega_0$, $(x^i_{k,0:T-1})$, and $(w^i_{k,t})_{k \in N_i}$ for $t \in \{0, \ldots, T-1\}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
In section 4, we establish the existence of a randomized NE for $\mathcal{P}_N^T$, 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 $\mathcal{P}_N^T$. 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 $\omega_0$. 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

\begin{equation}
\Theta_k^i := \left\{ P \in \mathcal{P}(U^i \times Y^i) \middle| P(B) = \int_B \delta_{(g^i_k(y^i_k))}(du^i_k)Q^i_k(dy^i_k), \quad g^i_k : Y^i \rightarrow U^i, \ B \in \mathcal{B}(U^i \times Y^i) \right\}.
\end{equation}

Also, let

\begin{equation}
\mathcal{R}_k^i := \left\{ P \in \mathcal{P}(U^i \times Y^i) \middle| P(B) = \int_B \Pi_k^i(du^i_k|y^i_k)Q^i_k(dy^i_k), \quad B \in \mathcal{B}(U^i \times Y^i) \right\},
\end{equation}

where $\Pi_k^i$ is a stochastic kernel from $Y^i$ to $U^i$. The set $\Theta_k^i$ is the set of extreme points of $\mathcal{R}_k^i$. Hence, $\Theta_k^i$ 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 $\mathcal{D}_k^i$, $\Gamma_k^i$, by $\mathcal{R}_k^i$, with the convergence in policies defined by

\begin{equation}
\gamma_{n,k}^i \xrightarrow{n \rightarrow \infty} \gamma_k^i(du^i_k|y^i_k)Q^i_k(dy^i_k) \quad \text{weakly} \quad \gamma_{n,k}^i(du^i_k|y^i_k)Q^i_k(dy^i_k).
\end{equation}

We equip $\mathcal{P}(Y^i \times U^i)$ with the $w$-$s$ topology, which is the coarsest topology under which the map $\int \kappa(y^i_k, u^i_k)P(dy^i_k, du^i_k) : \mathcal{P}(Y^i \times U^i) \rightarrow \mathbb{R}$ is continuous for every measurable and bounded function $\kappa$, which is continuous in $u^i_k$ for every $y^i_k$. Unlike the weak convergence topology, $\kappa$ need not be continuous in $y^i_k$ (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 $\mathcal{P}L^i$ as a collection of Borel probability measures on $\Gamma_N^i$, where Borel $\sigma$-field $\mathcal{B}(\Gamma_N^i)$ is induced by the introduced topology in (3.3) and $\Gamma_N^i$ is equipped with the

\(^1\)This set corresponds to the set of Young measures introduced in [41].
Let \( z \) be exogenous system random variables. In the above, for every fixed \( z \), \( P_{\pi,k} \in \mathcal{P}(\Gamma_k) \) corresponds to an independent randomized policy for each DM \( k \in \mathcal{N} \) (affecting only finitely many elements) of \( \mathbb{N} \). Since \( L_{\text{CO},N} \) and \( L_{\text{N}} \) are equal (see [32, Theorem A.1]), \( L_{\text{N}} \) 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 \( \sigma \) of \( \{1, \ldots, m\} \), \( \mathcal{L}(x^{\sigma(1)},x^{\sigma(m)}) = \mathcal{L}(x^{1:m}) \), where \( \mathcal{L}(\cdot) \) denotes the (joint) law of random variables. Random variables \( (x^1,x^2,\ldots) \) are infinitely exchangeable if finite distributions of \( (x^1,x^2,\ldots) \) and \( (x^{\sigma(1)},x^{\sigma(2)},\ldots) \) are identical for any finite permutation (affecting only finitely many elements) of \( \mathbb{N} \).

Let \( L_{\text{EX},N} \subseteq L_{\text{N}} \) for each PL \( k \) be given by

\[
L_{\text{EX},N} := \left\{ P_{\pi} \in L_{\text{N}} \left| \forall A_k \in \mathcal{B}(\Gamma^k_i) : P_{\pi}(\gamma^1_k \in A_1, \ldots, \gamma^i_k \in A_i) \right. \right\},
\]

(3.4)

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

\[
L_{\text{CO},\text{SYM},N} := \left\{ P_{\pi} \in L_{\text{N}} \left| \forall A_k \in \mathcal{B}(\Gamma^k_i) : P_{\pi}(\gamma^1_k \in A_1, \ldots, \gamma^i_k \in A_i) \right. \right\},
\]

where for all \( k \in N_i \), conditioned on \( z^i \), \( \tilde{P}_{\pi} \in \mathcal{P}(\Gamma^k_i) \) corresponds to an independent randomized policy for each DM \( k \)'s DM, which is symmetric among PL \( k \)'s DMs. Note that \( L_{\text{CO},\text{SYM},N} \subseteq L_{\text{EX},N} \).

**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 \( L^i, L_{\text{CO}}, L_{\text{EX}}, \) and \( L_{\text{CO},\text{SYM}} \) 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 \( \Gamma^i := \prod_{k \in \mathbb{N}} \Gamma^k_i \) is defined as \( L := \mathcal{P}(\Gamma^i) \). Let the set of all randomized policies with common and independent randomness be given by

\[
L_{\text{CO}} := \left\{ P_{\pi} \in L^i \left| \forall A_k \in \mathcal{B}(\Gamma^k_i) : P_{\pi}(\gamma^1_k \in A_1, \gamma^2_k \in A_2, \ldots) \right. \right\},
\]

where

\[
= \int_{z^i \in [0,1]} \prod_{k \in \mathbb{N}} P_{\pi,k}(\gamma^k_k \in A_k|z^i)\eta^i(dz^i), \quad \eta^i \in \mathcal{P}([0,1]).
\]
Let the set of all infinitely exchangeable randomized policies \( L_{\text{EX}} \) be given by
\[
L_{\text{EX}} := \left\{ P_\pi \in L^i \mid \forall A_k \in \mathcal{B}(\Gamma_{k}) , \forall N_i \in \mathbb{N} , \text{ and } \forall \sigma \in S_{N_i} : 
\pi(\gamma_{1}^{i} \in A_{1} , \ldots , \gamma_{N_i}^{i} \in A_{N_i}) = P_\pi(\gamma_{\sigma(1)}^{i} \in A_{1} , \ldots , \gamma_{\sigma(N_i)}^{i} \in A_{N_i}) \right\}.
\]
Let the symmetric policies with common and independent randomness be given by
\[
L_{\text{CO}} := \left\{ P_\pi \in L^i \mid \forall A_k \in \mathcal{B}(\Gamma_{k}) : \forall \sigma \in S_{N_i} : 
P_\pi(\gamma_{1}^{i} \in A_{1} , \ldots , \gamma_{N_i}^{i} \in A_{N_i}) = \prod_{k \in \mathbb{N}} P_\pi(\gamma_k^{i} \in A_k) \right\}.
\]
For the infinite population setting, we also define the following set of randomized policies with only private independent randomness:
\[
L_{\text{PR}} := \left\{ P_\pi \in L^i \mid \forall A_k \in \mathcal{B}(\Gamma_{k}) : 
P_\pi(\gamma_{1}^{i} \in A_{1} , \gamma_{2}^{i} \in A_{2} , \ldots) = \prod_{k \in \mathbb{N}} P_{\pi,k}(\gamma_k^{i} \in A_k) \right\}.
\]

Define the set of symmetric randomized policies with private independent randomness as
\[
L_{\text{PR,SYM}} := \left\{ P_\pi \in L^i \mid \forall A_k \in \mathcal{B}(\Gamma_{k}) : 
P_\pi(\gamma_{1}^{i} \in A_{1} , \gamma_{2}^{i} \in A_{2} , \ldots) = \prod_{k \in \mathbb{N}} \tilde{P}_{\pi}(\gamma_k^{i} \in A_k) \right\}.
\]

We note that \( L_{\text{PR}} \) and \( L_{\text{PR,SYM}} \) are not convex sets; however, \( L_{\text{PR}} \) (\( L_{\text{PR,SYM}} \)) contains the set of extreme points of the convex set \( L_{\text{CO}} \) (\( L_{\text{CO,SYM}} \)). 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 \( L_{\text{EX}} \), \( L_{\text{CO,SYM}} \), \( L_{\text{CO}} \), and \( L_{\text{CO,SYM}} \). Then
(i) any \( P_\pi \in L_{\text{EX}} \) belongs also to \( L_{\text{CO,SYM}} \), i.e., \( L_{\text{CO,SYM}} = L_{\text{EX}} \);
(ii) \( L_{\text{CO,SYM}} \subseteq L_{\text{EX}} \); however, in general \( L_{\text{CO,SYM}} \neq L_{\text{EX}} \).

**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 \( \mathcal{U} \) be compact. Then \( L_{\text{EX}} \) and \( L_{\text{EX}} \) 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 \((P^1_\pi, P^2_\pi)\) \in \(L^1_N \times L^2_N\) are given by

\[
J_{\pi,N}(P^1_\pi, P^2_\pi) := \int \int P^1_\pi(dy_1) P^2_\pi(dy_2) \mu_N^N(d\omega_0, dy) c^N(\gamma, y, \omega_0)
\]

\[
:= \int \int c^i(\omega_0, u^1_{1:N}, u^2_{1:N}) \prod_{i=1}^2 \prod_{k=1}^{N_i} \gamma_k^i(du^i_k | y^i_k) \times \prod_{i=1}^2 P^i_\pi(dy^i_{1:N}) \mu^N(d\omega_0, dy^1_{1:N}, dy^2_{1:N})
\]

where \(\mu^N\) is the joint probability measure of \(\omega_0, dy^1_{1:N}, dy^2_{1:N}\) and

\[
c^N(\gamma, y, \omega_0) := \int c^i(\omega_0, u^1_{1:N}, u^2_{1:N}) \prod_{i=1}^2 \prod_{k=1}^{N_i} \gamma_k^i(du^i_k | y^i_k).
\]

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 \( \mathcal{U} \) is compact and \( c^\omega(\omega, \ldots, \cdot) \) is continuous and (uniformly) bounded for all \( \omega_0 \). Then there exists a randomized policy profile \((P^1_\pi, P^2_\pi)\) \in \(L^1_{\mathcal{C},N} \times L^2_{\mathcal{C},N}\) that constitutes an NE for \( \mathcal{P}_N \).

Note that allowing common and independent randomness in randomization of policies leads to the convexity of \(L^1_{\mathcal{C},N}\). Since \(L^1_{\mathcal{C},N}\) and \(L^2_{\mathcal{C},N}\) are identical, \(L^1_{\mathcal{C},N}\) 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 \( \mathcal{P}_N \).** 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 \(L^1_{\mathcal{E},N} \times L^2_{\mathcal{E},N}\) (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) $U^i$ is compact;
(ii) $c^i(\omega_0, \ldots, \cdot)$ is continuous and (uniformly) bounded for all $\omega_0$;
(iii) $(y_k^i)_{k \in \mathcal{N}_i}$ are $N_i$-exchangeable, conditioned on $\omega_0$;
(iv) $(y_k^2)_{k \in \mathcal{N}_2}$ and $(y_k^1)_{k \in \mathcal{N}_1}$ are mutually independent, conditioned on $\omega_0$.

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 $(P^1, P^2) \in L^1_{\mathcal{E}, N} \times L^2_{\mathcal{E}, N}$ that constitutes an NE for $\mathcal{P}_N$ among all policy profiles in $L^1_N \times L^2_N$.

Since the costs are continuous by Assumption 3.1, the convexity and compactness of $L^1_{\mathcal{E}, N}$ 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 $L^1_N \times L^2_N$. 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 $L^1_{\mathcal{E}, N} \times L^2_{\mathcal{E}, N}$).

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 $(P^1, P^2) \in L^1_N \times L^2_N$ are given by

$$J^1_{\pi, N}(P^1, P^2) := \int P^1_N(d\gamma^1)P^2_N(d\gamma^2)\mu^N(d\omega_0, dy)c^i_N(\gamma, y, \omega_0)$$

with

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

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 $c^i_N$ which satisfies Assumption 2.1.

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

$$J^1_{\pi, \infty}(P^1, P^2) := \limsup_{N_1, N_2 \to \infty} \int P^1_{\pi, N_1}(d\gamma^1)P^2_{\pi, N_2}(d\gamma^2)\mu^N(d\omega_0, dy)c^i_N(\gamma, y, \omega_0).$$
where $P^{i}_{N}$ is the marginal of $P^{i}_{N} \in L^{1}$ to the first $N_{i}$ components, and $\mu^{N}$ is the marginal of the fixed probability measure on $(\omega_{0}, y_{1:N_{1}}, y_{1:N_{2}})$.

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 $(P^{1s}_{\pi}, P^{2s}_{\pi}) \in L^{1}_{PR,SYM} \times L^{2}_{PR,SYM}$ 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 $L^{1}_{PR,SYM} \times L^{2}_{PR,SYM}$, 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 $L^{1}_{PR,SYM} \times L^{2}_{PR,SYM}$. 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 $L^{1}_{N}$, 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 $\epsilon_{N_{1},2} \rightarrow 0$ as $N_{1}, N_{2} \rightarrow \infty$.

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 $\epsilon_{N_1,T}$-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 $\epsilon_{N_1,T}$-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 $\mathcal{P}_T^\dagger$ that is exchangeable, and an NE for $\mathcal{P}_\infty$ 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 $\zeta_k := (w_{k,t},\psi_{k,t})$ with $\zeta_{k,-1} := x_{k,0}, \zeta^0 := (\xi^0,\ldots,\xi^T)$, $\zeta^{1:2}_{1:T-1} = (\xi^1_{T-1},\ldots,\xi^2_{T-1},\ldots,\xi^2_{1:T-1})$, and $\zeta^i := \xi^i_{1:T-1}$. Also, define the information history at time $t$ for any $t \in \{1,\ldots,T-1\}$ by

\begin{equation}
H_t := \left\{ \omega_0, \xi^1_{1:t-1}, \xi^2_{0:t-1}, \eta^1_{1:t-1} \right\}.
\end{equation}

Assumption 4.1. For every $N \in \mathbb{N} \cup \{\infty\}$, let $\nu^N_i$ be the distribution of $y^1_{t,i}$ 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 $t \in \{0,\ldots,T-1\}$:

(i) (Independent reduction) There exist a probability measure $\tau^N_{i,k,t}$ and a function $\psi^N_{k,t}(\cdot)$ such that for all Borel sets $A^i_{k,t}$ on $\mathbb{Y}^i$ (with $A = \prod_{i=1}^T A^i_1 \times \cdots \times A^i_N$)

\[ \nu^N_i(A|H_t) = \sum_{k=1}^{2N} \prod_{i=1}^T \psi^N_{k,t}(y^i_{k,t}, H_t) \tau^N_{i,k,t}(dy^i_{k,t}). \]

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

\[ \nu^N_i(A|H_t) = \sum_{k=1}^{2N} \prod_{i=1}^T \tilde{\omega}^N_{k,t}(y^i_{k,t}, H_t) \tilde{\eta}^N_{k,t}(dy^i_{k,t}|H^i_{k,t}) \]

with $H^i_{k,t} := \xi^i_{k,-t-1}, y^i_{k,0:t-1}, y^i_{k,0:t-1}$ and for each DM $i_k$ over time, there exists a static reduction under which the observations $y^i_{k,t}$ are expanding and distributed by fixed probability measures $\eta^i_{k,t}$, i.e., $\sigma(y^i_{k,t}) \subseteq \sigma(y^i_{k,t+1})$.

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

\[ \gamma_{k}^{i,n} \xrightarrow{n\to\infty} \gamma_{k}^{i} \iff \gamma_{k}^{i,n}(du^i_{k,t},|y^i_{k,t}) \gamma_{k}^{i}(dy^i_{k,t}) \text{ weakly } \gamma_{k}^{i,n}(du^i_{k,t},|y^i_{k,t}) \gamma_{k}^{i}(dy^i_{k,t}) \quad \forall t \in \mathbb{T}. \]

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

\[ \gamma_{k}^{i,n} \xrightarrow{n\to\infty} \gamma_{k}^{i} \iff \gamma_{k}^{i,n}(du^i_{k,t},|y^i_{k,0:t}) \gamma_{k}^{i}(dy^i_{k,0:t}) \text{ weakly } \gamma_{k}^{i,n}(du^i_{k,t},|y^i_{k,0:t}) \gamma_{k}^{i}(dy^i_{k,0:t}) \]

That is, under Assumption 4.1, by considering $\gamma_{k}^{i}$ we can define sets of randomized policies $L_{\mathcal{Y}}, L_{\mathcal{X}}\mathcal{N}, L_{\mathcal{X}}E\mathcal{Y}, L_{\mathcal{X}}\mathcal{S}Y, L_{\mathcal{X}}S\mathcal{M}, L_{\mathcal{X}}P_{\mathcal{Y}}, L_{\mathcal{X}}P_{\mathcal{Y}}$, and $L_{\mathcal{X}}P_{\mathcal{Y}}$ 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 $I_{k,t} := \{y_{k,t}\}$, where observations are symmetric and given by

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

where the functions $h_t^i$ are identical for all $k \in N_i$. In the above, we used $M_{k,0:t-1}^{1:2} := (c^i_{k,-1:t}, u^i_{k,0:t-1}, \zeta^i_{k,-1:t}, \omega^i_{k,0:t-1})$ to denote the information up to time $t$ excluding $(c^i_{k,-1:t}, u^i_{k,0:t-1})$. 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

$$J_{\pi,N}^{1:T}(P_{\pi}^1, P_{\pi}^2) := \int P_{\pi}^1(dy^1)P_{\pi}^2(dy^2)\mu^N(d\omega_0, d\zeta) \int_1^2 N N_0 dy^1_N \gamma^i N(du^i|y^i, \pi) \prod_{k=1}^{N_i} \gamma_{k,0:t-1}^{i}(du^i|y^i)$$

$$\times \prod_{i=1}^{2} \int_1^2 N_0 dy^i N_0 \prod_{t=0}^{T-1} \nu_t^{N}(dy^1_t|H_t)$$

where $\mu^N$ denotes the joint distribution of $(\omega_0, \zeta_{1:2}^1)$, and

$$c^i_N(\gamma, \zeta, y, \omega_0) := \int c^i(\omega_0, \zeta_{1:2}^1, u_{1:2}^1) \prod_{k=1}^{N_i} \gamma_{k,0:t-1}^{i}(du^i|y^i)$$

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

Assumption 4.2. $c^i$ is (separately) exchangeable for any permutations $\sigma$ and $\tau$ of $\{1, \ldots, N_1\}$ and $\{1, \ldots, N_2\}$, respectively, for all $\omega_0, \zeta^1_{1:2}$

$$c^i(\omega_0, \zeta^1_{1:2}, u_{1:2}^1) = c^i(\omega_0, \zeta^1_{\sigma(1:\tau)}(\sigma(N_i)), \zeta^2_{\tau(1:\tau)}(\tau(N_i)), u^1_{\sigma(1:\tau)}(\sigma(N_i)), u^2_{\tau(1:\tau)}(\tau(N_i))).$$

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

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

(i) $U^i$ is compact;

(ii) $c^i(\omega_0, \zeta_{1:2}^1, \cdot)$ in (4.4) is continuous and (uniformly) bounded for all $\omega_0, \zeta^1_{1:2}$;

(iii) $h_t^i$ in (4.2) is continuous in actions for $t = 0, \ldots, T - 1$;

(iv) $\zeta^1_{1:2}$ are exchangeable, conditioned on $\omega_0$;

(v) for all Borel sets $A_i^k$ on $\mathbb{Y}^i$ (with $A^i = A^1_1 \times \cdots \times A^i_{N_i}$)

$$\nu^N_t(A|H_t) = \prod_{i=1}^{2} \prod_{k=1}^{N_i} \nu_t^i(A_i^k|\omega_0, \zeta_i^i, \omega_{0:t-1}^i, u_{0:t-1}^{1:2}/M_{i,0:t-1}^{1:2}) \quad \forall t \in T.$$
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 \((P_1^N, P_2^N) \in L_{EX,N}^1 \times L_{EX,N}^2\) 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 is without any loss.

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

Let \( PL^i \)'s expected cost for \( \mathcal{P}_N^T \) under a randomized policy profile \((P_1^N, P_2^N) \in L_{N}^1 \times L_{N}^2\) be given by

\[
J_{i,N}^T(P_1^N, P_2^N) = \int P_1^N(d\gamma^1_P)P_2^N(d\gamma^2_P)\mu^N(d\omega_0, d\xi)^c_i^N(\gamma_\ast, \xi_\ast, \omega_0)\nu^N(dy\mid \xi_\ast, \gamma_\ast, \omega_0),
\]

where

\[
c_i^N(\gamma_\ast, \xi_\ast, \omega_0) := \int \frac{1}{N_1} \sum_{k=1}^{N_1} \sum_{t=0}^{T-1} \epsilon^i(\omega_0, x_{k,t}, u_{k,t}, \xi_t^1 \left( \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{x_{p,t}}^1 \right), \xi_t^2 \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{x_{p,t}}^2 \right)), \sum \delta_{u_{i,t}} \left( \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,t}} \right) \delta_{u_{i,t}} \ast \prod_{i=1}^2 \delta_{y_{i,t}}(du\mid y_i).
\]

where \( \mu^N \) is the joint distribution of \((\omega_0, \xi^1:2)\) and \( \nu^N = \prod_{t=0}^{T-1} \nu_i^t \), with \( \nu_i^t \) the joint distribution of \( y_{i}^t \| H_t \) in (4.1).

Let \( PL^i \)'s expected cost for \( \mathcal{P}_N^\infty \) under a randomized policy profile \((P_1^N, P_2^N) \in L_{N}^1 \times L_{N}^2\) be given by

\[
J_{i,\infty}^T(P_1^N, P_2^N) = \limsup_{N \to \infty} \int P_1^N(d\gamma^1_P)P_2^N(d\gamma^2_P)\mu^N(d\omega_0, d\xi)^c_i^N(\gamma_\ast, \xi_\ast, \omega_0)\nu^N(dy\mid \xi_\ast, \gamma_\ast, \omega_0),
\]

where \( P_i^N \) is the marginal of \( P_i^\infty \in L^i \) to its first \( N_i \) components.

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

\[
\psi_i^t \left( y_{i,k,t}, \omega_0, c_{k,-t-1}^i, c_{k,-t-1}^i, y_{i,k,0:t-1}, u_{i,k,0:t-1}, \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}}^1, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}}^2 \right), \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}}^1, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}}^2 \right), \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}}^1, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}}^2 \right), \frac{1}{N_1} \sum_{p=1}^{N_1} \delta_{u_{p,0:t-1}}^1, \frac{1}{N_2} \sum_{p=1}^{N_2} \delta_{u_{p,0:t-1}}^2 \right).
\]

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
where \( \psi^i_k(y^i_k, t, \omega, \sigma^i, \zeta^i, \gamma^i, \omega_0, \ldots) \) and \( \phi^i_k(y^i_k, \omega_0, \ldots) \) are continuous (weakly continuous in their last arguments) and uniformly bounded for all \( y^i_k, \omega_0, \gamma^i, \sigma^i, \zeta^i \).

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

\[
P^i_{\sigma,n} (\gamma^i_1 \in A^i_1, \ldots, \gamma^i_{N_i} \in A^i_{N_i}) = P^i_{\sigma} (\gamma^i_{\sigma(1)} \in A^i_1, \ldots, \gamma^i_{\sigma(N_i)} \in A^i_{N_i}).
\]
Let $\mathcal{T}$ be a countable measure-determining subset of the set of all real-valued continuous functions on $\prod_{k=1}^{N_i} \Gamma_k$. For a function $f \in \mathcal{T}$, we have

\begin{equation}
(A.1) \quad \int f(\gamma_{1:N_i})dP_{\pi}^{\gamma_{1:N_i},\infty} = \lim_{n \to \infty} \int f(\gamma_{1:N_i})dP_{\pi}^{\gamma_n,n} = \int f(\gamma_{1:N_i})dP_{\pi}^{\gamma_{\infty},\infty},
\end{equation}

where (A.1) follows from the fact that $P_{\pi}^{\gamma_{1:N_i},n}$ is $N_i$-exchangeable for every $n$, and from convergence of $\{P_{\pi}^{\gamma_{1:N_i},n}\}_n$ and $\{P_{\pi}^{\gamma_{1:N_i},n}\}_n$ 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 $P_{\pi}^{\gamma_{1:N_i},\infty}$ and $P_{\pi}^{\gamma_{\infty},\infty}$ are the same, and hence $L_{EX,N}$ is compact since it is tight. An argument similar to the above establishes that $L_{EX}$ 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 $L_{CO,N}$ and $L_N$ are identical. This is because $L_{CO,N}^i$ is a convex set and the set of extreme points of $L_{N}^i$ is a subset of $L_{CO,N}^i$. This implies that the set $L_{CO,N}^i$ 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 $L_{CO,N}^i$ is compact since it is tight.

Define the best response correspondence $\Phi : \prod_{i=1}^{2} L_{CO,N}^i \to 2\prod_{i=1}^{2} L_{CO,N}^i$ as

$$
\Phi(P_{\pi}^{1}, P_{\pi}^{2}) := \bigcap_{i=1}^{2} \text{BR}(P_{\pi}^{i}), \text{ where for } i = 1, 2,
$$

$$
\text{BR}(P_{\pi}^{i}) := \left\{ P_{\pi}^{i} \in L_{CO,N}^i \left| J_{1}^{i}(P_{\pi}^{1}, P_{\pi}^{i}) \leq J_{1}^{i}(\hat{P}_{\pi}^{i}, P_{\pi}^{i}) \right. \forall P_{\pi}^{i} \in L_{CO,N}^i \right\}.
$$

In the following, we show that $\Phi$ admits a fixed point, which will be denoted by $(P_{\pi}^{1*}, P_{\pi}^{2*})$.

Next, we show that the graph

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

is closed. Consider sequences of policies $\{P_{\pi}^{1-i,n}\}_n \subseteq L_{CO,N}^i$ and $\{P_{\pi}^{i,n}\}_n \subseteq L_{CO,N}^i$ that converge weakly to $P_{\pi}^{1-i}$ and $P_{\pi}^{i}$, respectively. If $P_{\pi}^{i} \in \text{BR}(P_{\pi}^{1-i})$, we get $P_{\pi}^{i} \in \text{BR}(P_{\pi}^{1-i})$ since

$$
\inf_{\hat{P}_{\pi}^{i} \in L_{CO,N}^i} J_{1}^{i}(\hat{P}_{\pi}^{i}, P_{\pi}^{1-i}) = \lim_{n \to \infty} \inf_{P_{\pi}^{i} \in L_{CO,N}^i} J_{1}^{i}(P_{\pi}^{i}, P_{\pi}^{1-i,n}) \geq J_{1}^{i}(P_{\pi}^{i}, P_{\pi}^{1-i}) \geq \inf_{\hat{P}_{\pi}^{i} \in L_{CO,N}^i} J_{1}^{i}(\hat{P}_{\pi}^{i}, P_{\pi}^{1-i}).
$$

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, and $\{P_{\pi}^{1-i,n}\}_n$ converges weakly to $P_{\pi}^{1-i}$. Inequality (A.4) follows from exchanging limit and infimum, $P_{\pi}^{i,n} \in \text{BR}(P_{\pi}^{1-i,n})$, and an argument similar to that used for (A.3), and (A.5) follows from the fact that $P_{\pi}^{i} \in L_{CO,N}^i$ (since $L_{CO,N}^i$ is closed). This implies that the graph $G$ is closed. Moreover, since $L_{CO,N}^i$ is compact, and $c$ are continuous in actions, using [18, Proposition D.5(b)], the map

$$
F_{\pi}^{i}: P_{\pi}^{1-i} \to \inf_{\hat{P}_{\pi}^{i} \in L_{CO,N}^i} J_{1}^{i}(\hat{P}_{\pi}^{i}, P_{\pi}^{1-i})
$$

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 $n \to \infty$. 

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
is continuous. This implies that BR($P^{i-1}_\pi$) is nonempty. Thus, by the Kakutani–Fan–Glicksberg fixed point theorem, $\Phi$ admits a fixed point, and this completes the proof since the sets $L^{i}_{\text{CO},N}$ and $L^{i}_{N}$ are identical.

**A.3. Proof of Theorem 3.5.** By Lemma 3.3, the set $L^i_{\text{EX},N}$ 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 $(P^{1*}_\pi, P^{2*}_\pi) \in L^1_{\text{EX},N} \times L^2_{\text{EX},N}$ among all policies in $L^1_{\text{EX},N} \times L^2_{\text{EX},N}$. We only need to show that $(P^{1*}_\pi, P^{2*}_\pi)$ constitutes an NE among all randomized policies in $L^1_N \times L^2_N$. It is sufficient to show that for any $P^{i*}_\pi \in L^i_{\text{EX},N}$, we have

(A.6) \[ \inf_{\hat{P}^i_{\pi} \in L^i_N} J^i_{\pi}(\hat{P}^i_{\pi}, P^{i*}_\pi) = \inf_{\hat{P}^i_{\pi} \in L^i_{\text{EX},N}} J^i_{\pi}(\hat{P}^i_{\pi}, P^{i*}_\pi). \]

This follows from [32, Lemma 1] using Assumptions 2.1 and 3.1 (iii) and (iv) by first showing that for any policy $P^i_{\pi} \in L^i_N$, we have that $J^i_{\pi}(P^i_{\pi}, P^{i*}_\pi) = J^i_{\pi}(P^{i*}_\pi, P^{i*}_\pi)$, where $P^{i*}_\pi$ is a $\sigma$-permutation of $P^i_{\pi}$, and then by showing that an exchangeable policy constructed as an average of all possible permutations of $P^i_{\pi}$ does not perform worst for PL. This implies (A.6), and hence, under Assumption 2.2, there exists an NE profile $(P^{1*}_\pi, P^{2*}_\pi)$ belonging to $L^1_{\text{EX},N} \times L^2_{\text{EX},N}$. 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, $(P^{1*}_\pi, P^{2*}_\pi)$ 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 $(P^{1*}_\pi, P^{2*}_\pi)$ that constitutes an NE for $\mathcal{P}_\infty$, among all randomized policies in $L^1_{\text{PR},\text{SYM}} \times L^2_{\text{PR},\text{SYM}}$. In Step 2, we show that the policy profile $(P^{1*}_\pi, P^{2*}_\pi)$ 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 $L^1_{\text{PR},\text{SYM}} \times L^2_{\text{PR},\text{SYM}}$. 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 $\omega_0$ using a change of measure argument in (2.8). Since our policy space is restricted to $L^1_{\text{PR},\text{SYM}} \times L^2_{\text{PR},\text{SYM}}$, 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 $(P^{1*}_\pi, P^{2*}_\pi)$ in $L^1_{\text{PR},\text{SYM}} \times L^2_{\text{PR},\text{SYM}}$

\[
\limsup_{N_1, N_2 \to \infty} \int \frac{1}{N_1} \sum_{k=1}^{N_1} c^i (\omega_0, u^i_k, \Xi^1 \left( \frac{1}{N_1} \sum_{k=1}^{N_1} u^1_k \right), \Xi^2 \left( \frac{1}{N_2} \sum_{k=1}^{N_2} u^2_k \right)) \times P^i_\pi (d\gamma^*_k) P^{i-1}(d\gamma^*_k) \mathbb{P}_0 (d\omega_0) \prod_{i=1}^{N_1} \int f^i (dy^i_k, \omega_0) Q^i(dy^i_k) \gamma^i_k (du^i_k|y^i_k) \\
= \int c^i (\omega_0, u^i_R, \Xi^1 (\Lambda^1 (du^i_R)), \Xi^2 (\Lambda^2 (du^2_R))) \Lambda^i (du^i_R) \\
\times \mathbb{P}_0 (d\omega_0) \prod_{i=1}^{N_1} \int f^i (y^i_R, \omega_0) Q^i (dy^i_R) P^{*i}_{\pi,R} (dy^i_R) \gamma^i_R (du^i_R|y^i_k) \\
:= J^i_{\pi,\infty} (P^{1*}_\pi, P^{2*}_\pi, \Lambda^i, \Lambda^{-1}),
\]

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

(A.7) \[ P^{\pi}_{i} (\gamma^i_k \in \cdot) = \prod_{k \in N} P^{\pi}_{i,R} (\gamma^i_k \in \cdot) \]
for some $P^i_{\pi,R}(\gamma^i_k \in \cdot)$ belonging to the set $\mathcal{P}(\Gamma^i_R)$, and
\begin{equation}
\Lambda^i(\cdot) = \mathcal{L}(u^i_R).
\end{equation}
We note that $\Lambda^i(\cdot)$ depends on the randomized policy $P^i_{\pi,R}$ 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 $P^i_{\pi,R}$. In view of the above costs, a policy profile $(P^1_{\pi,R}, P^2_{\pi,R})$ of the forms (A.7) constitutes an NE for $\mathcal{P}_\infty$ if $P^i_{\pi,R} \in \text{BR}(P^i_{\pi,R}, \Lambda^1, \Lambda^2)$ for $i = 1, 2$ with
\begin{equation}
\text{BR}(P^i_{\pi,R}, \Lambda^1, \Lambda^2) := \left\{P^i_{\pi,R} \in \mathcal{P}(\Gamma^i_R) \right\}
\end{equation}
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 $\Phi : \prod_{i=1}^2 \mathcal{P}(\Gamma^i_R) \to 2\prod_{i=1}^2 \mathcal{P}(\Gamma^i_R)$ as $\Phi = \Psi \circ \Theta$ with $\Theta(P^1_{\pi,R}, P^2_{\pi,R}) := (\Lambda^1, \Lambda^2)$ satisfying the consistency condition (A.8), and $\Psi(\Lambda^1, \Lambda^2) = \prod_{i=1}^2 \text{BR}(P^i_{\pi,R}, \Lambda^1, \Lambda^2)$. We have that $\mathcal{P}(\Gamma^i_R)$ is nonempty, compact, and convex. Next, we show that the graph
\begin{equation*}
G = \left\{((P^1_{\pi,R}, P^2_{\pi,R}), \Phi(P^1_{\pi,R}, P^2_{\pi,R})) \left| (P^1_{\pi,R}, P^2_{\pi,R}) \in \prod_{i=1}^2 \mathcal{P}(\Gamma^i_R) \right. \right\}
\end{equation*}
is closed.

Suppose that sequences $\{P^1_{\pi,R}\}_n$ and $\{P^2_{\pi,R}\}_n$ converge to $P^1_{\pi,R}$ and $P^2_{\pi,R}$, respectively. Let $P^i_{\pi,R} \in \text{BR}(P^i_{\pi,R}, \Lambda^1, \Lambda^2)$ for $i = 1, 2$ and $\Theta(P^1_{\pi,R}, P^2_{\pi,R}) := (\Lambda^1, \Lambda^2)$. Since the distribution of $u^i_R$, induced by the policy $P^i_{\pi,R}$, converges to the distribution of $u^i_R$, induced by the policy $P^i_{\pi,R}$, using the fact that $\Theta(P^1_{\pi,R}, P^2_{\pi,R})$ is singleton, we get that $\Theta$ is continuous. Next, we show that $\Psi$ is upper-hemicontinuous. We have, for $i = 1, 2$,
\begin{equation}
\inf_{\tilde{P}^i_{\pi,R} \in \mathcal{P}(\Gamma^i_R)} \int c^i(\omega_0, \tilde{u}^i_R, \Xi^i(\Lambda^i_\infty(\tilde{u}^i_R)), \Xi^{-i}(\Lambda^{-i}_\infty(\tilde{u}^{-i}_R)))) \Lambda^i_\infty(\tilde{u}^i_R) \nonumber
\end{equation}
\begin{equation}
\times \tilde{P}^i_{\pi,R}(d\gamma^i_R) \mathbb{P}(d\gamma^i_R) \mathbb{P}(d\omega_0) \prod_{s=1}^2 \int f^s(y^s_R, \omega_0) Q^s(dy^s_R) \gamma^s_R(dy^s_R | y^s_R) \nonumber
\end{equation}
\begin{equation}
(A.10) = \lim_{\tilde{P}^i_{\pi,R} \in \mathcal{P}(\Gamma^i_R), n \to \infty} \inf \int c^i(\omega_0, \tilde{u}^i_R, \Xi^i(\Lambda^i_n(\tilde{u}^i_R)), \Xi^{-i}(\Lambda^{-i}_n(\tilde{u}^{-i}_R)))) \Lambda^i_n(\tilde{u}^i_R) \nonumber
\end{equation}
\begin{equation}
\times \tilde{P}^i_{\pi,R}(d\gamma^i_R) \mathbb{P}(d\gamma^i_R) \mathbb{P}(d\omega_0) \prod_{s=1}^2 \int f^s(y^s_R, \omega_0) Q^s(dy^s_R) \gamma^s_R(dy^s_R | y^s_R) \nonumber
\end{equation}
\begin{equation}
(A.11) \geq \inf \int c^i(\omega_0, \tilde{u}^i_R, \Xi^i(\Lambda^i_\infty(\tilde{u}^i_R)), \Xi^{-i}(\Lambda^{-i}_\infty(\tilde{u}^{-i}_R)))) \Lambda^i_\infty(\tilde{u}^i_R) \nonumber
\end{equation}
\begin{equation}
\times \tilde{P}^i_{\pi,R}(d\gamma^i_R) \mathbb{P}(d\gamma^i_R) \mathbb{P}(d\omega_0) \prod_{s=1}^2 \int f^s(y^s_R, \omega_0) Q^s(dy^s_R) \gamma^s_R(dy^s_R | y^s_R), \nonumber
\end{equation}
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 $P^{i,n}_\pi \in \mathrm{BR}(P^{i,n}_\pi, \Lambda^1_n, \Lambda^2_n)$, and along the same lines as (A.10). Hence, the above chain of inequalities becomes the chain of equalities, and this implies that $Q$ is closed. Moreover, since $P(\Gamma_k)$ is compact and $J^i_{\pi, \infty}$ is upper-hemicontinuous, we have that $\mathrm{BR}(P^{i,n}_\pi, \Lambda^1_n, \Lambda^2_n)$ is nonempty. The Kakutani–Fan–Glicksberg fixed point theorem completes the proof of this step.

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

$$
(A.12) \quad \inf_{P^i_\pi \in L_{\pi, \infty}^{L^1}} J^i_{\pi, \infty}(P^i_\pi, P^{i, \infty}_\pi) = \inf_{P^i_\pi \in L^1} J^i_{\pi, \infty}(P^i_\pi, P^{i, \infty}_\pi).
$$

By fixing $P^{i, \infty}_\pi$, $P^i_\pi$ 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 $L^k_{\mathbb{E}}|_N$ be the set of randomized policies in $L^k_{\mathbb{E}}$ restricted to their first $N_i$ components. By Lemma 3.2, we have $L^k_{\mathbb{E}}|_N = L^k_{\mathbb{C}, \mathbb{O}, \mathbb{S}, \mathbb{Y}, \mathbb{M}}|_N$, but $L^k_{\mathbb{E}}|_N \neq L^k_{\mathbb{E}}|_N$. We have

$$
(A.13) \quad \inf_{P^i_\pi \in L^1} \limsup_{N_1, N_2 \to \infty} \int \frac{1}{N_i} \sum_{k=1}^{N_i} \omega_k, u_k^i, \Xi^1 \left( \frac{1}{N_1} \sum_{k=1}^{N_1} u_k \right), \Xi^2 \left( \frac{1}{N_2} \sum_{k=1}^{N_2} u_k \right) dP^i_\pi P^{i, \infty}_\pi(d\omega_k) \prod_{i=1}^{N} \hat{f}^i(dy_k, \omega_k) Q^i(dy_k) \gamma_k(dy_k) d\gamma_k
$$

$$
(A.14) \quad \inf_{P^i_\pi \in L^1} \limsup_{N_1, N_2 \to \infty} \int \frac{1}{N_i} \sum_{k=1}^{N_i} \omega_k, u_k^i, \Xi^1 \left( \frac{1}{N_1} \sum_{k=1}^{N_1} u_k \right), \Xi^2 \left( \frac{1}{N_2} \sum_{k=1}^{N_2} u_k \right) dP^i_\pi P^{i, \infty}_\pi(d\omega_k) \prod_{i=1}^{N} \hat{f}^i(dy_k, \omega_k) Q^i(dy_k) \gamma_k(dy_k) d\gamma_k
$$

$$
(A.15) \quad \inf_{P^i_\pi \in L^1} \limsup_{N_1, N_2 \to \infty} \int \frac{1}{N_i} \sum_{k=1}^{N_i} \omega_k, u_k^i, \Xi^1 \left( \frac{1}{N_1} \sum_{k=1}^{N_1} u_k \right), \Xi^2 \left( \frac{1}{N_2} \sum_{k=1}^{N_2} u_k \right) dP^i_\pi P^{i, \infty}_\pi(d\omega_k) \prod_{i=1}^{N} \hat{f}^i(dy_k, \omega_k) Q^i(dy_k) \gamma_k(dy_k) d\gamma_k
$$

$$
(A.16) \quad \inf_{P^i_\pi \in L^1} J^i_{\pi, \infty}(P^i_\pi, P^{i, \infty}_\pi),
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
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 \(L_{\text{EX}}^{i} \subset L_{\text{EX},N}^{i}\), 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 \(P_{\pi, N}\) since \(L_{\text{PR,SYM}}^{i} | N\) contains the set of extreme points of \(L_{\text{CO,SYM}} | N\). 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 \(\{\tilde{P}_{\pi, N_1, n}^{i}\}_{n \in N_1} \subseteq L_{\text{PR,SYM}}^{i} | N\). Since \(L_{\text{PR,SYM}}^{i} | N\) is compact, there exists a subsequence \(\{\tilde{P}_{\pi, n}^{i}\}_{n \in N_1} \subseteq L_{\text{PR,SYM}}^{i} | N\), with index \(n \in N_1\), that converges to \(\tilde{P}_{\pi, \infty}^{i} \in L_{\text{PR,SYM}}^{i} | N\). Hence, compactness of \(L_{\text{PR,SYM}}^{i}\) allows us to extract subsequences of policies \(\{\tilde{P}_{\pi, n}^{i}\}_{n} \) and \(\{\tilde{P}_{\pi, n}^{i}\}_{n}\) that converge to \(\tilde{P}_{\pi, \infty}^{i}\) and \(\tilde{P}_{\pi, \infty}^{2}\), respectively. Define the empirical measures of actions as

\[
(A.17) \quad \tilde{\Lambda}_{n}^{i}(\cdot) := \frac{1}{n} \sum_{k=1}^{n} \delta(\tilde{u}_{k,n}^{i})\left(\cdot\right), \quad \tilde{\Lambda}_{n, \infty}^{i}(\cdot) := \frac{1}{n} \sum_{k=1}^{n} \delta(\tilde{u}_{k, \infty}^{i})\left(\cdot\right),
\]

where actions \(\tilde{u}_{k,n}^{i}\) and \(\tilde{u}_{k, \infty}^{i}\) are induced by randomized policies \(\tilde{P}_{\pi, n}^{i}\) and \(\tilde{P}_{\pi, \infty}^{i}\), respectively. Similarly, we define \(\Lambda_{n, \infty}^{-i}\), where the actions \(u_{k, \infty}^{-i}\) are induced by randomized policies \(P_{\pi, \infty}^{-i}\). By Assumption 2.3 and under the change of measure in (2.8) (see Assumption 2.2), the observations \((u_{k}^{i})_{k \in N}\) are i.i.d. Since randomized policies \(\tilde{P}_{\pi, n}^{i}\) and \(\tilde{P}_{\pi, \infty}^{i}\) belong to \(L_{\text{PR,SYM}}^{i} | n\) and \(L_{\text{PR,SYM}}^{i}\), respectively, we can conclude that \(\tilde{u}_{k, \infty}^{i}, \tilde{u}_{k, \infty}^{-i}\), and \(u_{k, \infty}^{-i}\) 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 \(\{\tilde{\Lambda}_{n}^{i}\}_{n}\) and \(\{\tilde{\Lambda}_{n, \infty}^{i}\}_{n}\) converge weakly to the same limit \(\tilde{\Lambda}_{\infty}^{i} := \mathcal{L}(u_{R, \infty}^{i})\). Also, \(\Lambda_{n, \infty}^{-i}\) converges weakly to \(\Lambda_{\infty}^{-i} := \mathcal{L}(u_{R, \infty}^{-i})\). 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 \((P_{1}^{*i}, P_{2}^{*i}) \in L_{1}^{i} | \text{PR,SYM} \times L_{2}^{i} | \text{PR,SYM}\) for \(P_{\infty}\). Proceeding along the same lines as (A.15)–(A.16), we get

\[
(A.18) \quad J_{i, \infty}^{i}(P_{1}^{*i}, P_{2}^{*i}) = \limsup_{N_{1}, N_{2} \to \infty} \inf_{P_{\pi, N_{1}} \in L_{\text{PR,SYM}}^{i} | N_{1}} J_{i, \infty}^{i}(P_{\pi, N_{1}}^{i}, P_{\pi, \infty}^{-i}) \quad \forall i \in \{1, 2\}.
\]

Following [44, Theorem 5.2], there exists a sequence of optimal policies as the best response to \(P_{\pi, \infty}^{-i}\), denoted by \(\{\tilde{P}_{\pi, N_{1}}^{i}\}_{N_{1}}\). 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\},

\[
(A.19) \quad \liminf_{N_{1}, N_{2} \to \infty} \inf_{P_{\pi, N_{1}} \in L_{\text{PR,SYM}}^{i} | N_{1}} J_{i, \infty}^{i}(P_{\pi, N_{1}}^{i}, P_{\pi, \infty}^{-i}) \geq \inf_{P_{\pi} \in L_{\text{PR,SYM}}^{i}} J_{i, \infty}^{i}(P_{\pi}^{i}, P_{\pi, \infty}^{-i}) = J_{i, \infty}^{i}(P_{1}^{*i}, P_{2}^{*i}),
\]

where (A.19) follows from an argument similar to that used in (A.15)–(A.16). This together with (A.18) implies that
By the strong law of large numbers since \((P^{1*}_{i}, P^{2*}_{i}) \in L^1_{\text{PR,SYM}} \times L^2_{\text{PR,SYM}}\), we get
\[
J^i_{\pi,\infty}(P^{1*}_{\pi}, P^{2*}_{\pi}) = \lim_{N_1, N_2 \to \infty} \inf_{P^i_{\pi,N} \in L^1_{\text{PR,SYM}}|N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{i*}_{\pi}) \quad \forall i \in \{1, 2\}.
\] (A.21)

Equality (A.20) together with (A.21) implies that there exists \(\epsilon_{N_{1,2}} > 0\), converging to zero as \(N_1, N_2 \to \infty\), such that
\[
\left| J^i_{\pi,N}(P^i_{\pi}, P^{i*}_{\pi}) - \inf_{P^i_{\pi,N} \in L^1_{\text{PR,SYM}}|N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{i*}_{\pi}) \right| \leq \epsilon_{N_{1,2}} \quad \forall i \in \{1, 2\}.
\] (A.22)

This implies that \((P^{1*}_{\pi}, P^{2*}_{\pi})\) constitutes an \(\epsilon_{N_{1,2}}\)-NE for \(P_N\) among independently randomized symmetric policy profiles \(L^1_{\text{PR,SYM}}|N_1} \times L^2_{\text{PR,SYM}}|N_2\). Proceeding along the same lines as (A.13)–(A.16), we get, for \(i \in \{1, 2\}\),
\[
\limsup_{N_1, N_2 \to \infty} \inf_{P^i_{\pi,N} \in L^1_{\text{PR,SYM}}|N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{i*}_{\pi}) = \limsup_{N_1, N_2 \to \infty} \inf_{P^i_{\pi,N} \in L^1_{\text{PR,SYM}}|N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{i*}_{\pi}).
\]

Hence, using (A.22), there exists \(\epsilon_{N_{1,2}} > 0\), converging to zero as \(N_1, N_2 \to \infty\) such that
\[
\left| J^i_{\pi,N}(P^i_{\pi}, P^{i*}_{\pi}) - \inf_{P^i_{\pi,N} \in L^1_{\text{PR,SYM}}|N_i}} J^i_{\pi,N}(P^i_{\pi,N}, P^{i*}_{\pi}) \right| \leq \epsilon_{N_{1,2}} \quad \forall i \in \{1, 2\}.
\] (A.23)

This implies that \((P^{1*}_{\pi}, P^{2*}_{\pi})\) constitutes an \(\epsilon_{N_{1,2}}\)-NE for \(P_N\) 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 \((P^{1*}_{\pi}, P^{2*}_{\pi})\) among all policies in \(L^1_{\text{EX,N}} \times L^2_{\text{EX,N}}\). 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 \(P^T_{\pi}\) 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 \(P^T_{\pi}\) to randomized policies belonging to \(L^1_{\text{PR,SYM}} \times L^2_{\text{PR,SYM}}\). 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 \(J^i_{\pi,\infty}(P^i_{\pi,R}, P^{i*}_{\pi,R})\) in (4.7) for any \((P^i_{\pi,R}, P^{i*}_{\pi,R}) \in L^1_{\text{PR,SYM}} \times L^2_{\text{PR,SYM}}\) as \(J^i_{\pi,\infty}(P^i_{\pi,R}, P^{i*}_{\pi,R}, \Lambda^i, \Lambda^{-i})\) with \(P^i_{\pi,R}(\gamma^i_k \in \cdot) = \prod_{k \in \mathbb{N}} P^i_{\pi,R}(\gamma^i_k \in \cdot)\) for some \(P^i_{\pi,R}(\gamma^i_k \in \cdot)\) belonging to the set \(\mathcal{P}(\prod_{t=0}^{T-1} \Gamma_{R,t})\), and
\[
\Lambda^i \left( \times \prod_{t=0}^{T-1} X^i \times Y^i \times Z^i \right) = \mathcal{L}(u^i_{\pi,R}), \quad \Lambda^i \left( \times \prod_{t=0}^{T-1} U^i \times Y^i \times Z^i \right) = \mathcal{L}(x^i_{\pi,R}).
\] (A.24)

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

Suppose that sequences $\{P_{\pi,R}^{1,n}\}_n$ and $\{P_{\pi,R}^{2,n}\}_n$ converge to $P_{\pi,R}^{1,\infty}$ and $P_{\pi,R}^{2,\infty}$, respectively. Let $P_{\pi,R}^{i,n} \in \text{BR}(P_{\pi,R}^{i,n}, \Lambda_i^n, \Lambda_i^2)$ for $i = 1, 2$ and $\Theta(P_{\pi,R}^{1,n}, P_{\pi,R}^{2,n}) := (\Lambda_1^n, \Lambda_2^n)$. Since the distribution of $\pi,R, \Theta_i^n$ induced by the policy $P_{\pi,R}^{i,n}$ converges to the distribution of $\pi,R, \Theta_i$, induced by the policy $P_{\pi,R}^{i,\infty}$, the marginal of $\Lambda_i^n$ on $\pi,R, \Theta_i$ converges to the marginal of $\Lambda_i^{\infty}$ on $\pi,R, \Theta_i$. By continuity of $f^{\pi,R}_i, \Xi_i, \pi,R, \Theta_i$, we get that $\pi,R, \Theta_i^n$ converges weakly to $\pi,R, \Theta_i$. Hence, marginals of $\Lambda_i^n$ on $\pi,R, \Theta_i^n$ converge to marginals of $\Lambda_i^{\infty}$ on $\pi,R, \Theta_i$. Hence, we can recursively show that $\Lambda_i^n$ converges weakly to $\Lambda_i^{\infty}$ for $i \in \{1, 2\}$ (and $\mathbb{P}^0$-a.s.). This leads to the continuity of $\Theta$. Along the same lines as (A.10)–(A.11), using the generalized dominated convergence theorem, it can be shown that $\Psi$ is upper-hemicontinuous. Hence, by the Kakutani–Fan–Glicksberg fixed point theorem, there exists a policy profile $(P_{1,\pi,R}^{1,\infty}, P_{2,\pi,R}^{2,\infty})$ that constitutes an NE for $\mathcal{P}_{T}^{\pi,R}$ among all randomized policies in $L_{R,\text{SYM}}^1 \times L_{R,\text{SYM}}^2$. By fixing $P_{\pi,R}^{1,\infty}$ and $\Lambda_{\pi,R}^{\infty}$, PL faces a team problem, and hence [32, Theorem 4] implies that (A.12) holds. Hence, $(P_{1,\pi,R}^{1,\infty}, P_{2,\pi,R}^{2,\infty})$ constitutes an NE for $\mathcal{P}_{\pi,R}^{\ast}$, among all randomized policy profiles in $L^1 \times L^2$, and completes the proof.

REFERENCES


NASH EQUILIBRIA FOR TEAM AGAINST TEAM GAMES

