Equivalence between Controlling a Large Population and a Representative Agent: Optimality, Learning, and Games (among Teams)

Serdar Yüksel

Queen's University, Department of Mathematics and Statistics

Collaborators: Sina Sanjari (Royal Military College, Canada) and Naci Saldi (Bilkent Univ.)

December 15, 2024

Outline

Outline

Decision Making and Information Structures

- Information structures in a stochastic system determine who knows what information and the impact of actions.
- Information structures have a significant impact on optimal systems design and the execution of computational/learning algorithms.
- We investigate the interaction between information and decision-making in large stochastic teams (i.e., with many interacting decision-makers (DMs)).
- **Examples.**
  - Traffic networks where decentralized traffic controllers aim to regulate the traffic to minimize the expected delay, e.g., [Chiri, Gong, and Piccoli'23] and [Festa and Göttlich'18].
  - Networked control where decentralized controllers, sensors, actuators, and encoders-decoders act over a control system toward optimizing a common goal, e.g., [Tatikonda and Mitter'04].
  - Swarm motion with an aversion to crowded regions, e.g., [Almulla, Ferreira, and Gomes'17].
  - Distribution matching, where the agents select actions to match a target distribution, e.g., [Carmona, Lauriére, and Tan'23].
  - Interacting particle models such as Markov chain Monte Carlo simulations, e.g., [Bou-Rabee and Schuh'23].

Decision Making and Information Structures

- Information structures in a stochastic system determine who knows what information and the impact of actions.
- Information structures have a significant impact on optimal systems design and the execution of computational/learning algorithms.
- We investigate the interaction between information and decision-making in large stochastic teams (i.e., with many interacting decision-makers (DMs)).
- **Examples.**
  - Traffic networks where decentralized traffic controllers aim to regulate the traffic to minimize the expected delay, e.g., [Chiri, Gong, and Piccoli'23] and [Festa and Göttlich'18].
  - Networked control where decentralized controllers, sensors, actuators, and encoders-decoders act over a control system toward optimizing a common goal, e.g., [Tatikonda and Mitter'04].
  - Swarm motion with an aversion to crowded regions, e.g., [Almulla, Ferreira, and Gomes'17].
  - Distribution matching, where the agents select actions to match a target distribution, e.g., [Carmona, Lauriére, and Tan'23].
  - Interacting particle models such as Markov chain Monte Carlo simulations, e.g., [Bou-Rabee and Schuh'23].

Decentralized Stochastic Teams

- A *team*, consists of a collection of DMs acting together to optimize a common cost function but not necessarily sharing all the available information.
- Teams whose initial states, observations, cost function, or the evolution dynamics are disturbed by some external noise processes are called *stochastic teams*.
- A team is *sequential* if the DMs act in a pre-defined order.
- 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 *dynamic*.

1. A collection of spaces $\{\Omega, \mathcal{F}, (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{Y}^i, \mathcal{Y}^i)\}$, specifying the system's distinguishable events, control, and observation spaces which are assumed to be standard Borel.
2. The $\mathbb{Y}^i$-valued observations and $\mathbb{U}^i$-valued actions are given by

$$y^i = h^i(\omega_0, \omega_i, y^{1:i-1}, u^{1:i-1})$$
$$u^i = \gamma^i(y^i),$$

where $1 : i = \{1, \ldots, i\}$ and $\gamma^i \in \Gamma^i$ is the set of all measurable functions from $\mathbb{Y}^i$ to $\mathbb{U}^i$.

3. Information structure $I^i \subseteq \{y^{1:i}, u^{1:i-1}\}$ determines who knows what.
4. There is a *probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$. There is a rich theory on stochastic teams and information structures.

Decentralized Stochastic Teams

- A *team*, consists of a collection of DMs acting together to optimize a common cost function but not necessarily sharing all the available information.
- Teams whose initial states, observations, cost function, or the evolution dynamics are disturbed by some external noise processes are called *stochastic teams*.
- A team is *sequential* if the DMs act in a pre-defined order.
- 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 *dynamic*.

1. A collection of spaces $\{\Omega, \mathcal{F}, (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{Y}^i, \mathcal{Y}^i)\}$, specifying the system's distinguishable events, control, and observation spaces which are assumed to be standard Borel.
2. The $\mathbb{Y}^i$-valued observations and $\mathbb{U}^i$-valued actions are given by

$$\begin{aligned} y^i &= h^i(\omega_0, \omega_i, y^{1:i-1}, u^{1:i-1}) \\ u^i &= \gamma^i(y^i), \end{aligned}$$

   where $1 : i = \{1, \ldots, i\}$ and $\gamma^i \in \Gamma^i$ is the set of all measurable functions from $\mathbb{Y}^i$ to $\mathbb{U}^i$.
3. Information structure $I^i \subseteq \{y^{1:i}, u^{1:i-1}\}$ determines who knows what.
4. There is a *probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$. There is a rich theory on stochastic teams and information structures.

Decentralized Stochastic Teams

- A *team*, consists of a collection of DMs acting together to optimize a common cost function but not necessarily sharing all the available information.
- Teams whose initial states, observations, cost function, or the evolution dynamics are disturbed by some external noise processes are called *stochastic teams*.
- A team is *sequential* if the DMs act in a pre-defined order.
- 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 *dynamic*.

1. A collection of spaces $\{\Omega, \mathcal{F}, (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{Y}^i, \mathcal{Y}^i)\}$, specifying the system's distinguishable events, control, and observation spaces which are assumed to be standard Borel.
2. The $\mathbb{Y}^i$-valued observations and $\mathbb{U}^i$-valued actions are given by

$$y^i = h^i(\omega_0, \omega_i, y^{1:i-1}, u^{1:i-1})$$
$$u^i = \gamma^i(y^i),$$

   where $1 : i = \{1, \ldots, i\}$ and $\gamma^i \in \Gamma^i$ is the set of all measurable functions from $\mathbb{Y}^i$ to $\mathbb{U}^i$.
3. Information structure $I^i \subseteq \{y^{1:i}, u^{1:i-1}\}$ determines who knows what.
4. There is a *probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$. There is a rich theory on stochastic teams and information structures.

Large/Mean-Field Stochastic Static Teams

Let action and observation spaces be $\mathbb{U} \subseteq \mathbb{R}^n$, and $\mathbb{Y} \subseteq \mathbb{R}^m$ for each DM.

- **Problem $\mathcal{P}^N$:** The expected cost under policy $\underline{\gamma}_N := \{\gamma^1, \cdots, \gamma^N\}$ is

$$J^N(\underline{\gamma}_N) = E^{\underline{\gamma}_N}\left[\frac{1}{N}\sum_{i=1}^{N} c\left(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^{N}\delta_{u^p}\right)\right].$$

- **Problem $\mathcal{P}^\infty$:** The expected cost under policy $\underline{\gamma} := \{\gamma^i\}_{i\in\mathbb{N}}$ is

$$J(\underline{\gamma}) = \limsup_{N\to\infty} E^{\underline{\gamma}}\left[\frac{1}{N}\sum_{i=1}^{N} c\left(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^{N}\delta_{u^p}\right)\right].$$

Globally optimal solution: A policy $\underline{\gamma}_N^\star$ such that

$$J^N(\underline{\gamma}_N^\star) = \inf_{\underline{\gamma}_N\in\Gamma_N} J^N(\underline{\gamma}_N).$$

Large Stochastic Dynamic Teams

State and observation dynamics are given by

$$x_{t+1}^i = f_t\left(x_t^i, u_t^i, \frac{1}{N}\sum_{p=1}^{N}\delta_{x_t^p}, \frac{1}{N}\sum_{p=1}^{N}\delta_{u_t^p}, w_t^i\right),$$

$$y_t^i = h_t\left(x_{0:t}^i, u_{0:t-1}^i, v_t^i\right).$$

- **Problem $\mathcal{P}_T^N$:** The expected cost under policy $\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}} := \{\boldsymbol{\gamma}^{\boldsymbol{i}}\}_{i\in\mathbb{N}}$ with $\boldsymbol{\gamma}^{\boldsymbol{i}} = \{\gamma_t^i\}_{t=0}^T$ is

$$J_T^N(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}) = \sum_{t=0}^{T-1} E^{\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}}\left[\frac{1}{N}\sum_{i=1}^{N} c\left(\omega_0, x_t^i, u_t^i, \frac{1}{N}\sum_{p=1}^{N}\delta_{u_t^p}, \frac{1}{N}\sum_{p=1}^{N}\delta_{x_t^p}\right)\right].$$

- **Problem $\mathcal{P}_T^\infty$:** The expected cost under policy $\underline{\boldsymbol{\gamma}} := \{\boldsymbol{\gamma}^{\boldsymbol{i}}\}_{i\in\mathbb{N}}$ is

$$J_T^\infty(\underline{\boldsymbol{\gamma}}) = \limsup_{N\to\infty} J_T^N(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}).$$

- To efficiently compute an optimal solution, we need to first establish the existence and structural results for an optimal solution.

Information Structures

We study the problems under three classes of information structures:

1. Decentralized information structure:

$$I_t^{i,\mathsf{DEC}} = \{x_{0:t}^i, u_{0:t-1}^i\}.$$

2. Centralized information structure:

$$I_t^{\mathsf{CEN}} = \{x_{0:t}^{1:N}, u_{0:t-1}^{1:N}\}.$$

3. Decentralized mean-field sharing information structure:

$$I_t^{i,\mathsf{MF}} = \{x_{0:t}^i, u_{0:t-1}^i, \mu_{0:t}^N\}$$

with $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{x_t^i}$.

Outline

Fully Decentralized Information Structures

Our first set of results is for teams under a fully decentralized information structure:

$$I_t^{i,\mathsf{DEC}} = \{x_{0:t}^i, u_{0:t-1}^i\}.$$

When we have observations:

$$I_t^{i,\mathsf{DEC}} = \{y_{0:t}^i, u_{0:t-1}^i\}.$$

For clarity in presentation, we will first discuss the static case:

$$I^{i,\mathsf{DEC}} = \{y^i\}.$$

Main Results under Fully Decentralized Information Structure

Theorem: *Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$ (or $\mathcal{P}_T^N$ and $\mathcal{P}_T^\infty$). Under sufficient conditions, we show that*

Convex Teams (with Decoupled Dynamics)

Exchangeable Teams with N DMs

An optimal solution exists that is symmetric and deterministic.

$N \downarrow \infty$

Mean-field limit

An optimal solution exists that is symmetric and independently randomized.

Non-Convex Teams

Exchangeable Teams with N DMs

approx. optimal

An optimal solution exists that is exchangeable and randomized.

$N \downarrow \infty$

Mean-field limit

An optimal solution exists that is symmetric and independently randomized.

Connections with Mean-Field Games, Teams, and Controls

- The mean-field approach designs policies for both games with infinitely many DMs and those with a large number of DMs [Huang-Caines-Malhamé'06'07, Lasry-Lions'06'07].
- Often, an infinite model is studied, and then its implications for large games are presented. The other direction of going from finite to infinite (connections between Nash equilibria of N-DM games and solutions of MFG) has also been studied, e.g., see [Bardi-Priuli'13, Arapostathis-Biswas-Carrol'17, Fischer'17'22, Lacker'17'20, cardaliaguet-Rainer'19].
- For infinite DM controls and teams, many studies focused on the LQG setup; see, e.g., [Mahajan-Martins-Y.'13, Huang-Nguyen'11'12, Wang-Zhang'17, Arabneydi-Mahajan'15, Ouyang-Asghari-Nayyar'18].

▶ The results presented for **MFG focus on Nash equilibrium, which may be inconclusive regarding global optimality**. We aim to find globally optimal solutions for teams with a finite and infinite number of DMs under a decentralized information structure.

▶ In the first part, we study **decentralized optimal policies for finite and infinite population teams**–we do not approach the problem from a centralized agent's perspective (and thus cannot consider MP or dynamic programming).

Connections with Mean-Field Games, Teams, and Controls

- The mean-field approach designs policies for both games with infinitely many DMs and those with a large number of DMs [Huang-Caines-Malhamé'06'07, Lasry-Lions'06'07].
- Often, an infinite model is studied, and then its implications for large games are presented. The other direction of going from finite to infinite (connections between Nash equilibria of N-DM games and solutions of MFG) has also been studied, e.g., see [Bardi-Priuli'13, Arapostathis-Biswas-Carrol'17, Fischer'17'22, Lacker'17'20, cardaliaguet-Rainer'19].
- For infinite DM controls and teams, many studies focused on the LQG setup; see, e.g., [Mahajan-Martins-Y.'13, Huang-Nguyen'11'12, Wang-Zhang'17, Arabneydi-Mahajan'15, Ouyang-Asghari-Nayyar'18].

► The results presented for **MFG focus on Nash equilibrium, which may be inconclusive regarding global optimality**. We aim to find globally optimal solutions for teams with a finite and infinite number of DMs under a decentralized information structure.

► In the first part, we study **decentralized optimal policies for finite and infinite population teams**–we do not approach the problem from a centralized agent's perspective (and thus cannot consider MP or dynamic programming).

Convex Teams: Existence of a Symmetric Optimal Solution

Theorem (SIAM J Cont and Opt'21, Trans. Auto Cont'21)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $c(\omega_0, \cdot, \cdot)$ *is continuous and convex;*
- $\mathbb{U}$ *is convex and compact;*
- *Observations are i.i.d. conditioned on* ω_0 *and i.i.d. under a measure transformation.*

Then,

1. *There exists a symmetric (identical) optimal policy for* $\mathcal{P}^N$.
2. *There exists a subsequence of (independently randomized) optimal policies for* $\mathcal{P}^N$, *converging to an optimal policy for* $\mathcal{P}^\infty$ *(and every converging sequence does so).*
3. *There exists a symmetric (independently randomized) optimal policy for* $\mathcal{P}^\infty$.

Proof method: Show that for every N agent, optimal policies exist and are symmetric. Take $N \to \infty$, via lower semi-continuity, show that any subsequential limit is optimal for the infinite problem; compactness of policies ensures the existence of such a converging subsequence.

Convex Teams: Existence of a Symmetric Optimal Solution

Theorem (SIAM J Cont and Opt'21, Trans. Auto Cont'21)

Consider $\mathcal{P}^N$ *and* $\mathcal{P}^\infty$. *Assume that*

- $c(\omega_0, \cdot, \cdot)$ *is continuous and convex;*
- $\mathbb{U}$ *is convex and compact;*
- *Observations are i.i.d. conditioned on* ω_0 *and i.i.d. under a measure transformation.*

Then,

1. *There exists a symmetric (identical) optimal policy for* $\mathcal{P}^N$.
2. *There exists a subsequence of (independently randomized) optimal policies for* $\mathcal{P}^N$, *converging to an optimal policy for* $\mathcal{P}^\infty$ *(and every converging sequence does so).*
3. *There exists a symmetric (independently randomized) optimal policy for* $\mathcal{P}^\infty$.

Proof method: Show that for every N agent, optimal policies exist and are symmetric. Take $N \to \infty$, via lower semi-continuity, show that any subsequential limit is optimal for the infinite problem; compactness of policies ensures the existence of such a converging subsequence.

Convex Teams: Existence of a Symmetric Optimal Solution

Theorem (SIAM J Cont and Opt'21, Trans. Auto Cont'21)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $c(\omega_0, \cdot, \cdot)$ *is continuous and convex;*
- $\mathbb{U}$ *is convex and compact;*
- *Observations are i.i.d. conditioned on* ω_0 *and i.i.d. under a measure transformation.*

Then,

1. *There exists a symmetric (identical) optimal policy for* $\mathcal{P}^N$.
2. *There exists a subsequence of (independently randomized) optimal policies for* $\mathcal{P}^N$*, converging to an optimal policy for* $\mathcal{P}^\infty$ *(and every converging sequence does so).*
3. *There exists a symmetric (independently randomized) optimal policy for* $\mathcal{P}^\infty$.

Proof method: Show that for every N agent, optimal policies exist and are symmetric. Take $N \to \infty$, via lower semi-continuity, show that any subsequential limit is optimal for the infinite problem; compactness of policies ensures the existence of such a converging subsequence.

Convexity and Dynamic Information Structure

- Many problems, however, are not convex, especially when the information is shaped via actions, as in dynamic team problems [SICON'17].
- Example [Witsenhausen'68]: Consider a dynamic team problem with two DMs with observations y^1 and $y^2 = u^1 + w^1$, where y^1 and w^1 standard normal random variables (with density η). The cost is

$$c(\omega, u^{1:2}) = k^2(y^1 - u^1)^2 + (u^1 - u^2)^2$$

for some $k > 0$. We can probabilistically reduce the problem to a static one by incorporating the dynamic dependence into the cost via measure change:

$$\int c(\omega, u^{1:2}) Q(dy^1)\gamma^1(du^1|y^1)\gamma^2(du^2|y^2)P(dy^2|u^1)$$

$$= \int \underbrace{c(\omega, u^{1:2})\frac{\eta(y^2 - u^1)}{\eta(y^2)}}_{\text{new cost is not convex in } u^1} \gamma^1(du^1|y^1)\gamma^2(du^2|y^2)Q(dy^1)\underbrace{\eta(y^2)dy^2}_{Q(dy^2)}.$$

- Relaxing convexity allows us to study teams with coupled dynamics:

$$x^i_{t+1} = f_t\left(x^i_t, u^i_t, \frac{1}{N}\sum_{p=1}^{N}\delta_{x^p_t}, \frac{1}{N}\sum_{p=1}^{N}\delta_{u^p_t}, w^i_t\right).$$

Convexity and Dynamic Information Structure

- Many problems, however, are not convex, especially when the information is shaped via actions, as in dynamic team problems [SICON'17].
- Example [Witsenhausen'68]: Consider a dynamic team problem with two DMs with observations y^1 and $y^2 = u^1 + w^1$, where y^1 and w^1 standard normal random variables (with density η). The cost is

$$c(\omega, u^{1:2}) = k^2(y^1 - u^1)^2 + (u^1 - u^2)^2$$

for some $k > 0$. We can probabilistically reduce the problem to a static one by incorporating the dynamic dependence into the cost via measure change:

$$\int c(\omega, u^{1:2}) Q(dy^1)\gamma^1(du^1|y^1)\gamma^2(du^2|y^2)P(dy^2|u^1)$$

$$= \int \underbrace{c(\omega, u^{1:2})\frac{\eta(y^2 - u^1)}{\eta(y^2)}}_{\text{new cost is not convex in } u^1} \gamma^1(du^1|y^1)\gamma^2(du^2|y^2)Q(dy^1)\underbrace{\eta(y^2)dy^2}_{Q(dy^2)}.$$

- Relaxing convexity allows us to study teams with coupled dynamics:

$$x^i_{t+1} = f_t\left(x^i_t, u^i_t, \frac{1}{N}\sum_{p=1}^{N}\delta_{x^p_t}, \frac{1}{N}\sum_{p=1}^{N}\delta_{u^p_t}, w^i_t\right).$$

Convexity and Dynamic Information Structure

- Many problems, however, are not convex, especially when the information is shaped via actions, as in dynamic team problems [SICON'17].
- Example [Witsenhausen'68]: Consider a dynamic team problem with two DMs with observations y^1 and $y^2 = u^1 + w^1$, where y^1 and w^1 standard normal random variables (with density η). The cost is

$$c(\omega, u^{1:2}) = k^2(y^1 - u^1)^2 + (u^1 - u^2)^2$$

for some $k > 0$. We can probabilistically reduce the problem to a static one by incorporating the dynamic dependence into the cost via measure change:

$$\int c(\omega, u^{1:2}) Q(dy^1)\gamma^1(du^1|y^1)\gamma^2(du^2|y^2) P(dy^2|u^1)$$

$$= \int \underbrace{c(\omega, u^{1:2}) \frac{\eta(y^2 - u^1)}{\eta(y^2)}}_{\text{new cost is not convex in } u^1} \gamma^1(du^1|y^1)\gamma^2(du^2|y^2) Q(dy^1) \underbrace{\eta(y^2)dy^2}_{Q(dy^2)}.$$

- Relaxing convexity allows us to study teams with coupled dynamics:

$$x^i_{t+1} = f_t\left(x^i_t, u^i_t, \frac{1}{N}\sum_{p=1}^{N} \delta_{x^p_t}, \frac{1}{N}\sum_{p=1}^{N} \delta_{u^p_t}, w^i_t\right).$$

Relaxing and Metrizing Policies

Assume the team is static.

Assumption

There exist functions f^i and probability measures Q^i such that

$$P((y^1,\dots,y^N)\in\mathbb{A}\mid\omega_0)=\prod_{i=1}^{N}\int_{\mathbb{A}^i} f^i(y^i,\omega_0,y^1,\dots,y^{i-1})Q^i(dy^i),\quad \forall\mathbb{A}^i\in\mathcal{B}(\mathbb{Y}^i).$$

- We define convergence on policies as

$$\gamma^i_n \xrightarrow{n\to\infty} \gamma^i \text{ iff } \delta_{\{g^i_n(y^i)\}}(du^i)Q^i(dy^i) \xrightarrow[\text{weakly}]{n\to\infty} \delta_{\{g^i_n(y^i)\}}(du^i)Q^i(dy^i).$$

- To relax convexity, we allow independent and correlated randomization in policies.
- We identify γ^i as an element in Borel probability measures on $\mathbb{U}^i\times\mathbb{Y}^i$ with fixed marginals on $\mathbb{Y}^i$ (under the weak convergence topology):

$$\Gamma^i=\left\{\gamma^i\in\mathcal{P}(\mathbb{U}^i\times\mathbb{Y}^i)\mid\gamma^i(\mathbb{A})=\int_{\mathbb{A}}\pi^i(du^i|y^i)Q^i(dy^i)\right\}.$$

Relaxing and Metrizing Policies

Assume the team is static.

Assumption

There exist functions f^i and probability measures Q^i such that

$$P((y^1,\ldots,y^N)\in\mathbb{A}\mid\omega_0)=\prod_{i=1}^{N}\int_{\mathbb{A}^i}f^i(y^i,\omega_0,y^1,\ldots,y^{i-1})Q^i(dy^i),\quad\forall\mathbb{A}^i\in\mathcal{B}(\mathbb{Y}^i).$$

- We define convergence on policies as

$$\gamma_n^i\xrightarrow{n\to\infty}\gamma^i\ \text{ iff }\ \delta_{\{g_n^i(y^i)\}}(du^i)Q^i(dy^i)\xrightarrow[\text{weakly}]{n\to\infty}\delta_{\{g_n^i(y^i)\}}(du^i)Q^i(dy^i).$$

- To relax convexity, we allow independent and correlated randomization in policies.
- We identify γ^i as an element in Borel probability measures on $\mathbb{U}^i\times\mathbb{Y}^i$ with fixed marginals on $\mathbb{Y}^i$ (under the weak convergence topology):

$$\Gamma^i=\left\{\gamma^i\in\mathcal{P}(\mathbb{U}^i\times\mathbb{Y}^i)\mid\gamma^i(\mathbb{A})=\int_{\mathbb{A}}\pi^i(du^i|y^i)Q^i(dy^i)\right\}.$$

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\text{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\text{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\mathsf{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\mathsf{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Sets of Randomized Policies and Classes of Correlations

- We introduce the following sets of randomized policies:
  - L^N: set of all joint distributions of $\gamma^{1:N}$;
  - L^N_{EX}: set of all joint distributions of γ^i such that

$$N\text{-exchangeability} \leftarrow \qquad \mathcal{L}(\gamma^1, \ldots, \gamma^N) = \mathcal{L}(\gamma^{\sigma(1)}, \ldots, \gamma^{\sigma(N)}) \quad \forall \text{ permutations } \sigma;$$

  - L^N_{CO}: set of all joint distributions of γ^is that are independent, conditioned on common randomness of z;
  - $L^i_{\text{CO,SYM}}$: set of all γ^is that are identical and independent, conditioned on common randomness of z;
  - L^i_{PR}: set of all γ^is that are independent.
  - $L^N_{\text{PR,SYM}}$: set of all joint distributions of γ^is that are identical (symmetric) and independent.
- We denote those for $\mathcal{P}^\infty$ by dropping super-index N, e.g., L_{CO}.
- A detailed study of correlations under decentralized information is [Saldi-Y.-(Prob. Surveys'22)].

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $\mathbb{U}$ *is compact;*
- $c(\omega_0, \cdot, \cdot)$ *is continuous;*
- *Observations of DMs are i.i.d. conditioned on* ω_0 *and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $\mathbb{U}$ *is compact;*
- $c(\omega_0, \cdot, \cdot)$ *is continuous;*
- *Observations of DMs are i.i.d. conditioned on ω_0 and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- *$\mathbb{U}$ is compact;*
- *$c(\omega_0, \cdot, \cdot)$ is continuous;*
- *Observations of DMs are i.i.d. conditioned on ω_0 and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $\mathbb{U}$ *is compact;*
- $c(\omega_0, \cdot, \cdot)$ *is continuous;*
- *Observations of DMs are i.i.d. conditioned on ω_0 and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- $\mathbb{U}$ *is compact;*
- $c(\omega_0, \cdot, \cdot)$ *is continuous;*
- *Observations of DMs are i.i.d. conditioned on* ω_0 *and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- *$\mathbb{U}$ is compact;*
- *$c(\omega_0, \cdot, \cdot)$ is continuous;*
- *Observations of DMs are i.i.d. conditioned on ω_0 and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Relaxing Convexity: Exchangeability and Symmetry for Optimal Policies

Theorem (Math of OR'23)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Assume that

- *$\mathbb{U}$ is compact;*
- *$c(\omega_0, \cdot, \cdot)$ is continuous;*
- *Observations of DMs are i.i.d. conditioned on ω_0 and i.i.d. under change of measure.*

Then:

1. *There exists an optimal policy for $\mathcal{P}^N$, that is exchangeable (i.e., it belongs to L^N_{EX}).*
2. *There exists an optimal policy for $\mathcal{P}^\infty$, that is symmetric and privately randomized (i.e., it belongs to $L_{\mathsf{PR,SYM}}$).*
3. *A symmetric optimal solution for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Note that N-exchangeability $\not\Longrightarrow$ symmetry.

Example: $P(X^1 = 1, X^2 = 0) = P(X^1 = 0, X^2 = 1) = 0.5, P(X^1 = 0, X^2 = 0) = P(X^1 = 1, X^2 = 1) = 0$. However, if they are conditionally i.i.d., we must have

$$P(X^1 = 0, X^2 = 0) = 0 = \int_0^1 p_z^2 \eta(dz), P(X^1 = 1, X^2 = 1) = 0 = \int_0^1 (1 - p_z)^2 \eta(dz).$$

Proof of Part (i)

1. Optimality of an exchangeable policy:
   - For any permutation σ of $\{1,\ldots,N\}$, show that the expected cost is invariant under any permutation of policies, i.e.,

$$J^N(P_\pi^{N,\sigma}) = J^N(P_\pi^N).$$

   - For any arbitrary policy P_π^N, construct an exchangeable policy $\widehat{P}_\pi^N$ by averaging policy over all possible permutations $\sigma \in S_N$.
   - By linearity of the expected cost in randomized policies yields that

$$J^N(\widehat{P}_\pi^N) = \frac{1}{|S_N|}\sum_{\sigma\in S_N} J^N(P_\pi^{N,\sigma}) = J^N(P_\pi^N).$$

2. Existence of an optimal solution within exchangeable ones:

$$\inf_{P_\pi^N \in \underbrace{L_{\mathsf{EX}}^N}_{\text{Compact}}} \underbrace{J^N(P_\pi^N)}_{\text{Continuous}} = \min_{P_\pi^N \in \underbrace{L_{\mathsf{EX}}^N}} \underbrace{J^N(P_\pi^N)}$$

Proof (Cont.) – Key Lemma

Lemma

Consider $\mathcal{P}^N$. Suppose that

- *The cost function $c : \Omega_0 \times \mathbb{U} \times \mathbb{U} \to \mathbb{R}_+$ is bounded and continuous in its second and third arguments;*
- *$\mathbb{U}$ is compact;*
- *Observations are i.i.d. under measure transformation.*

Then:

$$\limsup_{N\to\infty} \inf_{P_\pi^N \in L_{EX}^N} J^N(P_\pi^N) = \limsup_{N\to\infty} \inf_{P_\pi \in L_{EX}} J^N(P_{\pi,N}).$$

Sketch of a Proof of Lemma

1. *N*-exchangeable optimal actions and policies converge in distribution to infinitely exchangeable actions and policies:
   - By [Diaconis and Friedman'80], we construct infinitely exchangeable policies and actions close in total variation to N-exchangeable optimal actions and policies for any finite marginals:
$$\left\| \mathcal{L}(\gamma_N^{1\star}, \ldots, \gamma_N^{m\star}) - \mathcal{L}(\gamma_{N,\infty}^{1}, \ldots, \gamma_{N,\infty}^{m}) \right\|_{\mathsf{TV}} \xrightarrow[N\to\infty]{} 0.$$
   - Using the fact that L_{EX} is compact, show that a subsequence of optimal policy and actions converge in distribution to infinitely exchangeable policy and actions: for every $m \geq 1$
$$\mathcal{L}(\gamma_n^{1\star}, \ldots, \gamma_n^{m\star}) \xrightarrow[n\to\infty]{} \mathcal{L}(\gamma_\infty^{1}, \ldots, \gamma_\infty^{m})$$
$$\mathcal{L}(u_n^{1\star}, \ldots, u_n^{m\star}) \xrightarrow[n\to\infty]{} \mathcal{L}(u_\infty^{1}, \ldots, u_\infty^{m}).$$
   - By [Aldous '85], the empirical measures of exchangeable optimal actions converge weakly since they converge in distribution.
2. Show that the limsup of expected cost converges as policies converge.
   - This follows from a generalized dominated convergence theorem, since marginals and empirical measures of actions converge weakly.

Proof (Cont.)

Theorem (de Finetti Representation Theorem for Policies)

Any infinitely exchangeable policy is conditionally independent and identical, i.e., for any $P_\pi \in L_{\mathsf{EX}}$

$$P_\pi(\underline{\gamma} \in \mathbb{A}) = \prod_{i=1}^{\infty} \int_0^1 \widehat{P}_\pi(\gamma^i \in \mathbb{A}^i | z)\eta(dz) \in L_{\mathsf{CO,SYM}}.$$

This yields that (second inequality to be justified in the next slide)

$$\begin{aligned}
\inf_{P_\pi \in L} \limsup_{N\to\infty} J_N(P_{\pi,N}) &\geq \limsup_{N\to\infty} \inf_{P_\pi^N \in L_{\mathsf{EX}}^N} J_N(P_\pi^N) \\
&= \limsup_{N\to\infty} \inf_{P_\pi \in L_{\mathsf{EX}}} J_N(P_{\pi,N}) \\
&= \limsup_{N\to\infty} \inf_{P_\pi \in L_{\mathsf{CO,SYM}}} J_N(P_\pi^N) \\
&= \limsup_{N\to\infty} \inf_{P_\pi^N \in L_{\mathsf{PR,SYM}}^N} J_N(P_{\pi,N}) \\
&\geq \inf_{P_\pi \in L_{\mathsf{PR,SYM}}} \limsup_{N\to\infty} J_N(P_{\pi,N}) \\
&\geq \inf_{P_\pi \in L} \limsup_{N\to\infty} J_N(P_{\pi,N}).
\end{aligned}$$

We have

$$\begin{aligned}
&\limsup_{N\to\infty} \inf_{P_\pi^N \in L^N_{\text{PR,SYM}}} \int P_\pi^N(d\underline{\gamma})\mu^N(d\omega_0, d\underline{y})c^N(\underline{\gamma}, \underline{y}, \omega_0) \\
&\geq \lim_{n\to\infty} \int \left(\int c\left(\omega_0, u, \int_{\mathbb{U}} u\Lambda_n(du \times \mathbb{Y})\right) \Lambda_n(du, dy)\right) \prod_{i=1}^{\infty} \gamma_n^\star(du^i, dy^i)\mathbb{P}_0(d\omega_0) \\
&\geq \int \left(\int c\left(\omega_0, u, \int_{\mathbb{U}} u\Lambda(du \times \mathbb{Y})\right) \Lambda(du, dy)\right) \prod_{i=1}^{\infty} \gamma^\star(du^i, dy^i)\mathbb{P}_0(d\omega_0) \qquad (*) \\
&\geq \inf_{P_\pi \in L_{\text{PR,SYM}}} \limsup_{N\to\infty} \int P_{\pi,N}(d\underline{\gamma})\mu^N(d\omega_0, d\underline{y})c^N(\underline{\gamma}, \underline{y}, \omega_0)
\end{aligned}$$

where $\Lambda_n(\cdot) = \frac{1}{N}\sum_{i=1}^n \delta_{(u_n^{i\star}, y^i)}(\cdot)$.
Equality $(*)$ follows from generalized DCT since $\gamma_n^\star$ and Λ_n converge to $\gamma^\star$ and Λ weakly, respectively as $n \to \infty$.

Summary: Main Results under Decentralized Information Structure

Theorem: *Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$ (or $\mathcal{P}_T^N$ and $\mathcal{P}_T^\infty$). Under sufficient conditions, we show that*

Convex Teams (with Decoupled Dynamics)

Exchangeable Teams with N DMs

An optimal solution exists that is symmetric and deterministic.

$N \downarrow \infty$

Mean-field limit

An optimal solution exists that is symmetric and independently randomized.

Non-Convex Teams

Exchangeable Teams with N DMs

approx. optimal

An optimal solution exists that is exchangeable and randomized.

$N \downarrow \infty$

Mean-field limit

An optimal solution exists that is symmetric and independently randomized.

Outline

Recall: Information Structures

We now focus on teams with centralized and decentralized mean-field sharing information structures:

1. Centralized information structure:

$$I_t^{\mathsf{CEN}} = \{x_{0:t}^{1:N}, u_{0:t-1}^{1:N}\}.$$

2. Decentralized mean-field sharing information structure:

$$I_t^{i,\mathsf{MF}} = \{x_{0:t}^i, u_{0:t-1}^i, \mu_{0:t}^N\}$$

with $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{x_t^i}$.

Centralized Finite and Infinite Horizon Discounted Problems

State dynamics is given by

$$x^i_{t+1} = f_t\left(x^i_t, u^i_t, \mu^N_t, w^i_t\right),$$

- For $\beta \in (0,1)$, the expected cost under policy $\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}} := \{\boldsymbol{\gamma}^{\boldsymbol{i}}\}_{i\in\mathbb{N}}$ with $\boldsymbol{\gamma}^{\boldsymbol{i}} = \{\gamma^i_t\}_{t=0}^T$ is

$$\mathcal{P}^N_T: \qquad J^N_T(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}) = E^{\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T-1}\beta^t c\left(x^i_t, u^i_t, \mu^N_t\right)\right],$$

$$\mathcal{P}^N: \qquad J^N(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}) = E^{\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{\infty}\beta^t c\left(x^i_t, u^i_t, \mu^N_t\right)\right].$$

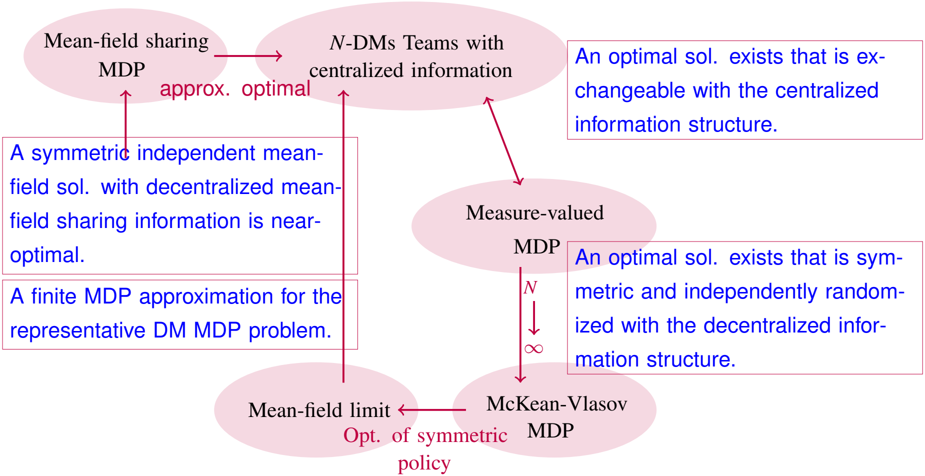
- The expected cost under policy $\underline{\boldsymbol{\gamma}} := \{\boldsymbol{\gamma}^{\boldsymbol{i}}\}_{i\in\mathbb{N}}$ is

$$\mathcal{P}^\infty_T: \qquad J^\infty_T(\underline{\boldsymbol{\gamma}}) = \limsup_{N\to\infty} J^N_T(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}),$$

$$\mathcal{P}^\infty: \qquad J^\infty(\underline{\boldsymbol{\gamma}}) = \limsup_{N\to\infty} J^N(\underline{\boldsymbol{\gamma}}_{\boldsymbol{N}}).$$

Main Results for Problems under Centralized and Mean-Field Sharing Information Structures

Theorem: *Consider $\mathcal{P}_T^N$ and $\mathcal{P}_T^\infty$. Under sufficient conditions, we show that*



Connections with McKean-Vlasov and Mean-Field Control Problems

- In [Bäuerle'23], an equivalent MDP formulation is characterized for finite population problems with centralized information. [Bäuerle'23] also showed that the limsup of value functions for the finite population MDP is equal to the value function of a limiting MDP.
- [Bäuerle'23] did not establish that the optimal value function can be attained by independently randomized symmetric policies, except when dynamics are decoupled.
- In [Carmona, Lauriére, and Tan'23], dynamic programming equations have been established for McKean-Vlasov control problems with a representative agent using a lifted measure-valued MDP.
- In [Motte and Pham'22], mean-field MDP, the connection between the finite population problem, and the limiting MDP for the McKean-Vlasov MDP has been established, assuming the policies are open-loop, symmetric, and decentralized.
- In [Arabneydi and Mahajan'14], mean-field sharing information structure studied for finite population mean-field teams under symmetric policies, where dynamic programming equations have been obtained using the common information approach.

► We demonstrate that the solution of the measure-valued MDP can be realized as an exchangeable policy for finite population and a symmetric and decentralized policy with mean-field sharing information for the infinite population problems.

Randomized Policies

- Centralized (correlated) policies: $\pi_t(u_t^{1:N} \in \cdot \mid x_{0:t}^{1:N}, u_{0:t-1}^{1:N})$;
- Centralized joint Markov policies: $\pi_t(u_t^{1:N} \in \cdot \mid x_t^{1:N}, \mu_t^N)$;
- Centralized exchangeable policies:

$$\pi_t(u_t^{\sigma(1):\sigma(N)} \in \cdot \mid x_{0:t}^{\sigma(1):\sigma(N)}, u_{0:t-1}^{\sigma(1):\sigma(N)}) = \pi_t(u_t^{1:N} \in \cdot \mid x_{0:t}^{1:N}, u_{0:t-1}^{1:N});$$

- Conditionally symmetric and independent with decentralized mean-field sharing information

$$\pi_t(u_t^{1:N} \in \cdot | x_{0:t}^{1:N}, u_{0:t-1}^{1:N}) = \int_0^1 \prod_{i=1}^N \overline{\pi}_t(u_t^i \in \cdot | x_{0:t}^i, \mu_{0:t}^N, z)\nu_t(dz);$$

- Symmetric and independent with decentralized mean-field sharing information

$$\pi_t(u_t^{1:N} \in \cdot | x_{0:t}^{1:N}, u_{0:t-1}^{1:N}) = \prod_{i=1}^N \overline{\pi}_t(u_t^i \in \cdot | x_{0:t}^i, \mu_{0:t}^N).$$

Measure-Valued Mean-Field MDP

Theorem (Bäuerle'23)

For centralized $\mathcal{P}^N$, an optimal policy can be expressed as a map from μ_t^N to θ_t^N for every $t \geq 0$.

Reformulate the centralized N-agent problem as a measure-valued MDP:

1. State and action (μ_t^N and θ_t^N):
   - Let $\mathcal{P}_{\mathsf{E}}^N(\mathbb{X} \times \mathbb{U})$ be the set of all empirical measures on $(x^{1:N}, u^{1:N})$.
   - For any $\mu^N \in \mathcal{P}_{\mathsf{E}}^N(\mathbb{X})$, a new measure-valued action set is

$$U(\mu^N) := \left\{\theta^N \in \mathcal{P}_{\mathsf{E}}^N(\mathbb{X} \times \mathbb{U}) \mid \theta^N(\cdot \times \mathbb{U}) = \mu^N(\cdot)\right\}.$$

2. Dynamics and costs (η and $\widetilde{c}$):
   - Show that

$$\mathbb{P}(\mu_{t+1}^N \in \cdot | \mu_{0:t}^N, \theta_{0:t}^N) = \mathbb{P}(\mu_{t+1}^N \in \cdot | \mu_t^N, \theta_t^N) := \eta(\cdot | \mu_t^N, \theta_t^N)$$

$$\frac{1}{N}\sum_{i=1}^{N} c(x^i, u^i, \mu^N) = \int c(x, u, \mu^N)\theta^N(dx, du) := \widetilde{c}(\mu^N, \theta^N).$$

3. Markov policies (g_t):
   - Deterministic (Markov) policies are measurable functions $g_t : \mu_t^N \mapsto \theta_t^N$.

Exchangeability of an Optimal Policy for N-DM Problems

Theorem (CDC'24)

Suppose that

(i) $\mathbb{U}$ *and* $\mathbb{X}$ *are compact;*

(ii) $c : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{R}$ *and* $f(\cdot,\cdot,\cdot,w) : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{X}$ *are jointly continuous.*

Then:

1. *There exists an optimal policy for* $\mathcal{P}_N$*, that is exchangeable.*
2. *The value iterations*

$$J^N_{T-1,T}(\mu^N) = \inf_{\theta^N \in U(\mu^N)} \widetilde{c}(\mu^N, \theta^N)$$

$$J^N_{t,T}(\mu^N) = \inf_{\theta^N \in U(\mu^N)} \left\{ \widetilde{c}(\mu^N, \theta^N) + \beta \int J^N_{t+1,T}(\overline{\mu}^N)\eta(d\overline{\mu}^N | \mu^N, \theta^N) \right\}$$

admit a Markov solution for $\mathcal{P}^N_T$ *that can be realized by an exchangeable Markov policy.*

3. *The value iterations*

$$J^N_{\infty}(\mu^N) = \inf_{\theta^N \in U(\mu^N)} \left\{ \widetilde{c}(\mu^N, \theta^N) + \beta \int J^N_{\infty}(\overline{\mu}^N)\eta(d\overline{\mu}^N | \mu^N, \theta^N) \right\}$$

admit a Markov stationary solution for $\mathcal{P}^N$ *that can be realized by an exchangeable stationary policy.*

Sketch of a Proof

1. Exchangeability of a realized policy at $t = T-1$:
   - Suppose that $g^\star_{T-1}$ is optimal, inducing $\theta^{N\star}_{T-1}$ and realized by $P^\star_{T-1}$ induced by a Markovian randomized policy $\pi^\star_{T-1}$. Denote the permutation of $P^\star_{T-1}$ by $P^{\sigma\star}_{T-1}$. This permutation does not change μ^N and $\theta^{N\star}_{T-1}$.
   - Construct an exchangeable distribution $\widehat{P}^\star_{T-1}$ on $x^{1:N}_{T-1}, u^{1:N}_{T-1}$ by averaging over all permutation of $P^\star_{T-1}$.
   - Show that

$$\int \widetilde{c}(\mu, \theta^N)\widehat{P}^\star_{T-1}(dx^{1:N}_{T-1}, du^{1:N}_{T-1}) = \int \tilde{c}(\mu, \theta^N)P^\star_{T-1}(dx^{1:N}_{T-1}, du^{1:N}_{T-1})$$

2. Exchangeability of a realized policy for $t = 0, \ldots, T-1$:
   - Recursively, show that

$$\int\int J^N_{t,T}(\mu')\eta\left(d\mu'|\mu^N, \theta^N\right)\widehat{P}^\star_{t-1}(dx^{1:N}_{T-2}, du^{1:N}_{T-2}) = \int J^N_{t,T}(\mu')\eta(d\mu'|\mu^N, g^\star_{t-1}(\mu)).$$

3. Existence of an optimal Markov policy $\boldsymbol{g}^\star$:
   - Use a measurable selection theorem.

Example: Symmetric Policies Might Not be Optimal for N-DM Problems

It is known that the map from μ^N to θ^N might not be attainable by symmetric policies (e.g., [Arabneydi and Mahajan'14]). A simple example is as follows:

Consider the dynamics

$$x_{t+1}^i = u_t^i.$$

Let $\mathbb{X} = \mathbb{U} = \{0, 1\}$ and $N = 2$. Take $x_0^1 = x_0^2 = 1$.

Let the cost be

$$\sum_{t=0}^{t=1} \sum_{x \in \mathbb{X}} \left| \mu_t^2(x) - \frac{1}{2} \right|^2, \qquad \mu_t^2(x) := \frac{1}{2} \left(\delta_{x_t^1}(x) + \delta_{x_t^2}(x) \right)$$

Under a symmetric policy, $u_t^i = \gamma_t^i(x_t^i, \mu_t^2) = \gamma_t(x_t^i, \mu_t^2)$, with $\mu_0^2 = [0 \quad 1]$, μ_1^2 will always stay away from the uniform measure $[\frac{1}{2} \quad \frac{1}{2}]$ with arbitrarily high probability;

- For deterministic policies: μ_1^2 will be a full mass on either 0 or 1;
- For randomized policies, μ_1^2 will be a positive measure on $(0, 0), (0, 1), (1, 0), (1, 1)$.

The optimal symmetric policy selects the actions with probabilities 0.5, attaining $0.5 + 0.25 = 0.75$.

However, if we choose $\gamma_0^1(1, \mu_0^2) = 1$ and $\gamma_0^2(1, \mu_0^2) = 0$, the total cost will be 0.5.

Symmetric Policies are Optimal in the Limit and Near-Optimal for Large N

Theorem (CDC'24)

Suppose that

(i) $\mathbb{U}$ *and* $\mathbb{X}$ *are compact;*

(ii) $c : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{R}$ *is jointly continuous;*

(iii) $f(\cdot, \cdot, \cdot, w) : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{X}$ *is jointly continuous for every* w.

Then:

1. *The sequence of exchangeable optimal policies for* $\mathcal{P}_T^N$ *obtained via value iterations converges through a subsequence to a symmetric and independent policy that is optimal for* $\mathcal{P}_T^\infty$.
2. *An optimal policy exists for* $\mathcal{P}_T^\infty$ *and* $\mathcal{P}^\infty$ *that is decentralized, symmetric, and independent with Markovian policies for* $\mathcal{P}_T^N$ *and stationary policies for* $\mathcal{P}^\infty$.
3. *A symmetric (conditionally) independent optimal solution for* $\mathcal{P}_T^\infty$ ($\mathcal{P}^\infty$) *under the decentralized mean-field-sharing information structure is near optimal for* $\mathcal{P}_T^N$ ($\mathcal{P}^N$) *within any policy with the centralized information structure for large* N.

Implications of Optimality of Symmetric Policies for Infinite Population Teams

Theorem (CDC'24)

Under the same assumptions,

1. *Any solution of the single-agent representative DM (Mckean-Vlasov) problem with*

$$x_{t+1}^R = f(x_t^R, u_t^R, \mu_t^R, w_t^R), \quad \mu_t^R = \mathcal{L}(x_t^R)$$

$$J_T^R(\pi^R) = \mathbb{E}^{\pi^R}\left[\sum_{t=0}^{T-1} \beta^t c(x_t^R, u_t^R, \mu_t^R)\right],$$

is optimal for $\mathcal{P}_T^\infty$.

2. *An optimal solution of the representative agent Mckean-Vlasov problem can be expressed as a map from* μ_t *to* $\theta_t = \mathcal{L}(x_t^R, u_t^R)$ *for every* $t \geq 0$.

Proof method. Reformulate by MDP with $\mu_{t+1}^R \sim \eta^R(\cdot|\mu_t^R, \theta_t^R)$ and the running cost

$$\widetilde{c}(\theta_t^R, \mu_t^R) = \int c(x_t^R, u_t^R, \mu_t^R)\theta_t^R(dx_t^R, du_t^R).$$

Endow the policy with Young topology (which makes it compact) and view it as the action space

$$\Gamma(\mu) = \left\{\theta_\pi \in \mathcal{P}(\mathbb{X} \times \mathbb{U}) \mid \theta_\pi(\mathbb{A}) = \int_{\mathbb{A}} \pi(du|x)\mu(dx)\right\}.$$

Implications of Near Optimality of Symmetric Policies for Large Teams

Lemma (CDC'24)

Under symmetric randomized policies,

$$P^{\boldsymbol{\pi}}\left((x_{t+1}^i, \mu_{t+1}^N) \in \cdot | x_{0:t}^{1:N}, u_{0:t}^{1:N}\right) = P^{\boldsymbol{\pi}}\left((x_{t+1}^i, \mu_{t+1}^N) \in \cdot | x_t^i, \mu_t^N\right).$$

Proposition (CDC'24)

Suppose that

(i) $\mathbb{U}$ *and* $\mathbb{X}$ *are compact;*

(ii) $c : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{R}$ *is jointly continuous;*

(iii) $f(\cdot, \cdot, \cdot, w) : \mathbb{X} \times \mathbb{U} \times \mathcal{P}(\mathbb{X}) \to \mathbb{X}$ *is jointly continuous for every* w.

1. *A symmetric independent optimal solution under mean-field sharing is nearly optimal for large teams under the centralized information structure.*
2. *Under symmetry, each agent with a mean-field sharing information structure faces an MDP.*

Verification Theorem for the Representative Agent

Theorem (CDC'24)

Consider the MDP for the representative DM under the same assumption.

(i) *For the finite horizon, an optimal Markov policy* $\boldsymbol{g}_T^\star$ *exists and satisfies the value iterations:*

$$J_{t,T}(\mu) = \mathbb{T}(J_{t+1,T})(\mu), \qquad J_{T-1,T}(\mu) := \inf_{\theta \in U(\mu)} \widetilde{c}(\mu,\theta) = \widetilde{c}(\mu, g_{T-1}^\star(\mu))$$

$$\mathbb{T}(\nu)(z) := \inf_{\theta \in U(z)} \left\{ \widetilde{c}(z,\theta) + \beta \int \nu(\bar{z}) \eta(d\bar{z}|z,\theta) \right\}.$$

Furthermore, an optimal policy $\boldsymbol{g}_T^\star$ *can be realized by a sequence of Markov policies* $\{\pi_t^\star\}_{t\geq 0}$ *with* $\pi_t^\star \in \mathcal{P}(\mathbb{U} \mid \mathbb{X})$.

(ii) *For the infinite horizon, an optimal stationary Markov policy* $\boldsymbol{g}^\star$ *exists and satisfies the value iteration:*

$$J_\infty(\mu) := \mathbb{T}(J_\infty)(\mu), \qquad J_\infty(\mu) := \left\{ \widetilde{c}(\mu, g^\star(\mu)) + \beta \int \nu(\overline{\mu}) \eta(d\overline{\mu}|\mu, g^\star(\mu)) \right\}.$$

Furthermore, an optimal policy $\boldsymbol{g}^\star$ *can be realized by sequence of stationary Markov policies* $(\pi^\star, \pi^\star, \ldots)$ *with* $\pi^\star \in \mathcal{P}(\mathbb{U} \mid \mathbb{X})$ *for the representative DM.*

Proof method. We show that transition kernel η is weakly continuous, and $\widetilde{c}$ is continuous, and we use the compactness of new action space under the Young topology.

Finite Approximation for the Representative DM Problem

Under our assumptions, the representative DM's MDP has a weak Feller kernel with the state space $\mathcal{P}(\mathbb{X})$ and the action space

$$\Gamma(\mu) = \left\{ \theta_\pi \in \mathcal{P}(\mathbb{X} \times \mathbb{U}) \mid \theta_\pi(\mathbb{A}) = \int_{\mathbb{A}} \pi(du|x)\mu(dx) \right\},$$

which is compact under the Young topology.

Denote the quantization of $\mathbb{X}$ under the nearest neighborhood quantizer by $\widehat{\mathbb{X}}_n$. Similarly, we quantize the action space $\widehat{\mathbb{U}}_n$.

Define

$$\widehat{\Gamma}_n(\widehat{\mu}) = \left\{ \widehat{\theta}_\pi \in \mathcal{P}(\widehat{\mathbb{X}}_n \times \widehat{\mathbb{U}}_n) \mid \widehat{\theta}_\pi(\mathbb{A}) = \int_{\mathbb{A}} \widehat{\pi}(du|x)\widehat{\mu}(dx) \right\}.$$

Define the transition kernel $\widehat{\eta}_n \in \mathcal{P}(\mathcal{P}(\widehat{\mathbb{X}}_n)|\mathcal{P}(\widehat{\mathbb{X}}_n) \times \widehat{\Gamma}_n)$ as a quantized version of η. Also, the cost $\widehat{c}(\widehat{\theta}_\pi, \widehat{\mu}) = \int c(x, u, \widehat{\mu})\widehat{\theta}_\pi(dx, du)$. Hence,

$$\left(\mathcal{P}(\widehat{\mathbb{X}}_n), \mathcal{P}(\widehat{\mathbb{X}}_n \times \widehat{\mathbb{U}}_n), \{\widehat{\Gamma}_n(\widehat{\mu}) | \widehat{\mu} \in \mathcal{P}(\widehat{\mathbb{X}}_n)\}, \widehat{\eta}_n, \widehat{c} \right)$$

constitutes a finite MDP denoted by MDP_n. Builds on [Saldi, Linder, and Y.'18].

Finite Approximation for the Representative DM Problem (Cont.)

Let

$$(\mathbb{T}_n(\widehat{\nu}))(\widehat{\mu}) := \inf_{\theta \in \widehat{\Gamma}_n(\widehat{\mu})} \left\{ \widehat{c}(\widehat{\mu}, \theta) + \beta \int \widehat{\nu}(\overline{\mu}) \widehat{\eta}_n(d\overline{\mu} \mid \widehat{\mu}, \theta) \right\}.$$

We write the value iterations for MDP_n:

$$\widehat{J}^n_{T-1,T}(\widehat{\mu}) = \inf_{\theta \in \widehat{\Gamma}_n(\widehat{\mu})} \widehat{c}(\widehat{\mu}, \theta)$$

$$\widehat{J}^n_{t,T}(\widehat{\mu}) = \mathbb{T}_n(\widehat{J}^n_{t,T})(\widehat{\mu}).$$

For the infinite horizon cost, we have

$$\widehat{J}^n_{\infty}(\widehat{\mu}) = \mathbb{T}_n(\widehat{J}^n_{\infty})(\widehat{\mu}).$$

Finite Approximation for the Representative DM Problem (Cont.)

Let $J_{0,T}$ and J_∞ be the optimal performance for the finite horizon and infinite horizon under MDP, respectively.

Let $\widehat{J}^n_{0,T}$ and $\widehat{J}^n_\infty$ be the optimal performances for the finite horizon and infinite horizon under $\mathsf{MDP_n}$, respectively.

Theorem (CDC'24)

Consider the MDP for the representative DM. Suppose further that

(i) μ_0 *is non-atomic,*

(ii) $\mathcal{T}(\cdot|x_0^R, u_0^R, \mu_0)$ *is non-atomic for every* x_0^R, u_0^R *and* μ_0.

Then,

$$\lim_{n\to\infty} \left|\widehat{J}^n_{0,T}(\mu_0) - J_{0,T}(\mu_0)\right| = 0,$$

$$\lim_{n\to\infty} \left|\widehat{J}^n_{\infty}(\mu_0) - J_{\infty}(\mu_0)\right| = 0.$$

Sketch of a Proof

1. Convergence at $t = T - 1$:
   - Under the assumptions, $\{\mu_t\}_{t\geq 0}$ is non-atomic under any policy. Show that $\widehat{\Gamma}_n(\widehat{\mu})$ is dense in $\Gamma(\mu)$.
   - For every $\widehat{\mu} \in \mathcal{P}(\widehat{\mathbb{X}})$, show that

$$\left|\widehat{J}^n_{T-1,T}(\widehat{\mu}) - J_{T-1,T}(\widehat{\mu})\right| = \left|\int c(x,u,\widehat{\mu})\pi^\star_n(du|x)\widehat{\mu}(dx) - \int c(x,u,\widehat{\mu})\pi^\star(du|x)\widehat{\mu}(dx)\right| \to 0.$$

   - Suppose that $\widehat{\mu}_n$ converges weakly to μ as $n \to \infty$. Show that the following converges to 0:

$$\left|\widehat{J}^n_{T-1,T}(\widehat{\mu}_n) - J_{T-1,T}(\mu)\right| = \left|\int c(x,u,\widehat{\mu}_n)\widetilde{\pi}^\star_n(du|x)\widehat{\mu}_n(dx) - \int c(x,u,\mu)\widetilde{\pi}^\star(du|x)\mu(dx)\right|.$$

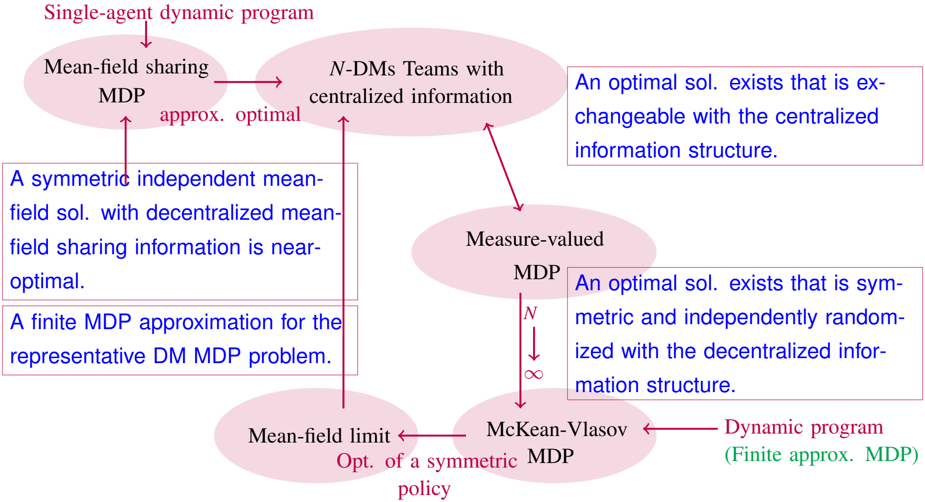
     This implies that $\widehat{J}^n_{T-1,T}$ converges continuously to $J_{T-1,T}$.
2. Convergence at $t = T - 2$:
   - Show that

$$\left|\int c(x,u,\widehat{\mu})\pi^\star_n(du|x,\widehat{\mu})\widehat{\mu}(dx) - \int c(x,u,\widehat{\mu})\pi^\star(du|x,\widehat{\mu})\widehat{\mu}(dx)\right|$$
$$+ \beta\left|\int \widehat{J}^n_{T-1,T}(\widehat{\mu}_{T-1})\widehat{\eta}_n(d\widehat{\mu}_{T-1}|\widehat{\mu}_{T-2},\widehat{\theta}^\star_n) - \int J_{T-1,T}(\mu_{T-1})\eta(d\mu_{T-1}|\widehat{\mu}_{T-2},\theta^\star)\right| \to 0.$$

   - Show $\widehat{J}^n_{T-2,T}$ converges continuously to $J_{T-2,T}$ and recursively $\widehat{J}^n_{0,T}$ converges to $J_{0,T}$.
3. Recursively, establish convergence at $t = 0$.

Review: Main Results for Problems under Centralized and Mean-Field Sharing Information Structures

Theorem: *Consider $\mathcal{P}_T^N$ and $\mathcal{P}_T^\infty$. Under sufficient conditions, we show that*



Outline

Games Among Teams

Several real-life competitive game scenarios occur on a team basis with multiple (large number of) DMs having access to local information.

Examples.

1. Multiple distributed renewable energy stations in the energy market:
   Multiple distributed renewable energy stations (each can be viewed as a large team distributed in various locations) compete for optimal energy production and pricing policies (where the abundance of energy resources might lead to price reduction, creating an incentive to interact strategically).
2. Large sensor networks vs large decentralized jammers:
   In sensor networks, a large collection of decentralized sensors (act as a large team) shares their information (by their actions) to a fusion center in the presence of jamming in the system, entailing a large team of decentralized jammers.

Games Among Teams with Fully Decentralized Information Structure

Consider a game between two teams.

Let action and observation spaces be $\mathbb{U}^i \subseteq \mathbb{R}^n$, and $\mathbb{Y}^i \subseteq \mathbb{R}^m$ for each DM of team i.

- **Problem $\mathcal{P}_N$:** The expected cost under policy $\underline{\gamma}_N^{1:2} := \{\underline{\gamma}_N^1, \underline{\gamma}_N^2\}$ with $\underline{\gamma}_N^i := \{\gamma_k^i\}_{k=1}^{N_i}$ is

$$J_N^i(\underline{\gamma}_N^{1:2}) = E^{\underline{\gamma}_N^{1:2}}\left[\frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u_k^i, \frac{1}{N_1}\sum_{p=1}^{N_1} u_p^1, \frac{1}{N_2}\sum_{p=1}^{N_2} u_p^2\right)\right].$$

- **Problem $\mathcal{P}_\infty$:** The expected cost under policy $\underline{\gamma}_\infty^{1:2} := \{\underline{\gamma}_\infty^1, \underline{\gamma}_\infty^2\}$ with $\underline{\gamma}_\infty^i := \{\gamma_k^i\}_{k=1}^{\infty}$ is

$$J_\infty^i(\underline{\gamma}_\infty^{1:2}) = \limsup_{N_1, N_2 \to \infty} E^{\underline{\gamma}_\infty^{1:2}}\left[\frac{1}{N_i}\sum_{k=1}^{N_i} c^i\left(\omega_0, u_k^i, \frac{1}{N_1}\sum_{p=1}^{N_1} u_p^1, \frac{1}{N_2}\sum_{p=1}^{N_2} u_p^2\right)\right].$$

Information is decentralized:

$$I_k^i = \{y_k^i\}.$$

We can also define the dynamic formulation with the decentralized information structure.

(Team-Wise) Nash Equilibrium

Definition

A policy $\underline{\gamma}_L^{1\star:2\star}$ is an ϵ-Nash equilibrium if

$$J_L^i(\underline{\gamma}_L^{1\star:2\star}) \leq \inf_{\underline{\gamma}_L^i \in \Gamma_L^i} J_L^i(\underline{\gamma}_L^{-i\star}, \underline{\gamma}_L^i) + \epsilon \quad i = 1, 2,$$

where $-i := \{1,2\} \backslash \{i\}$ and $L = N$ or ∞. If $\epsilon = 0$, then $\underline{\gamma}_L^{1\star:2\star}$ constitutes a Nash equilibrium.

Our notion of (team-wise) Nash equilibrium is a suitable equilibrium notion for such games since (global) optimality is desirable among DMs within teams.

Existence of Exchangeable Nash equilibrium for $\mathcal{P}_N$

Theorem (SIAM J Cont and Opt'24)

Consider $\mathcal{P}_N$. Let

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

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

(iii) $(y^i_k)_{k=1}^{N_i}$ *are* N_i*-exchangeable, conditioned on* ω_0.

(iv) $(y^1_k)_{k=1}^{N_1}$ *and* $(y^2_k)_{k=1}^{N_2}$ *are mutually independent, conditioned on* ω_0.

Then, there exists a Nash equilibrium for $\mathcal{P}_N$ that is N-exchangeable, i.e., it belongs to $L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$.

Proof method. First, under (i), show that $L^i_{\mathsf{EX},N}$ is convex and compact. Then, use (ii) and the Kakutani-Fan-Glicksberg fixed point theorem to show the existence of a Nash equilibrium among all exchangeable policies. Finally, use (iii) and (iv) to show that we can restrict the search over exchangeable policies in finding the best response policies to fixed exchangeable ones of the other team.

- For more general non-exchangeable games, we showed the existence of a Nash equilibrium with common randomness among each team.

Existence of Exchangeable Nash equilibrium for $\mathcal{P}_N$

Theorem (SIAM J Cont and Opt'24)

Consider $\mathcal{P}_N$. Let

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

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

(iii) $(y_k^i)_{k=1}^{N_i}$ *are* N_i*-exchangeable, conditioned on* ω_0.

(iv) $(y_k^1)_{k=1}^{N_1}$ *and* $(y_k^2)_{k=1}^{N_2}$ *are mutually independent, conditioned on* ω_0.

Then, there exists a Nash equilibrium for $\mathcal{P}_N$ *that is N-exchangeable, i.e., it belongs to* $L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$.

Proof method. First, under (i), show that $L^i_{\mathsf{EX},N}$ is convex and compact. Then, use (ii) and the Kakutani-Fan-Glicksberg fixed point theorem to show the existence of a Nash equilibrium among all exchangeable policies. Finally, use (iii) and (iv) to show that we can restrict the search over exchangeable policies in finding the best response policies to fixed exchangeable ones of the other team.

- For more general non-exchangeable games, we showed the existence of a Nash equilibrium with common randomness among each team.

Existence of Exchangeable Nash equilibrium for $\mathcal{P}_N$

Theorem (SIAM J Cont and Opt'24)

Consider $\mathcal{P}_N$. Let

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

(ii) $c^i(\omega_0, \cdot, \dots, \cdot)$ *is continuous and (uniformly) bounded for all* ω_0.

(iii) $(y^i_k)_{k=1}^{N_i}$ *are* N_i*-exchangeable, conditioned on* ω_0.

(iv) $(y^1_k)_{k=1}^{N_1}$ *and* $(y^2_k)_{k=1}^{N_2}$ *are mutually independent, conditioned on* ω_0.

Then, there exists a Nash equilibrium for $\mathcal{P}_N$ *that is* N*-exchangeable, i.e., it belongs to* $L^1_{\mathsf{EX},N} \times L^2_{\mathsf{EX},N}$.

Proof method. First, under (i), show that $L^i_{\mathsf{EX},N}$ is convex and compact. Then, use (ii) and the Kakutani-Fan-Glicksberg fixed point theorem to show the existence of a Nash equilibrium among all exchangeable policies. Finally, use (iii) and (iv) to show that we can restrict the search over exchangeable policies in finding the best response policies to fixed exchangeable ones of the other team.

- For more general non-exchangeable games, we showed the existence of a Nash equilibrium with common randomness among each team.

Existence of Exchangeable Nash equilibrium for $\mathcal{P}_N$

The zero-sum case: Conditional independence can be relaxed [Hogeboom-Burr-Y. (SCL'23)], as the proof builds on a minimax theorem (and not a fixed point argument).

Existence of Symmetric Nash equilibrium for $\mathcal{P}_\infty$ and Approximate Nash equilibrium for $\mathcal{P}_N$

Theorem (SIAM J Cont and Opt'24)

Consider $\mathcal{P}_\infty$. Let

(i) $\mathbb{U}^i$ *is compact and* $c^i(\omega_0, \cdot, \cdot, \cdot)$ *is continuous for all* ω_0.

(ii) $(y_k^1)_{k=1}^{N_1}$ *and* $(y_k^2)_{k=1}^{N_2}$ *are independent, conditioned on* ω_0;

(iii) $(y_k^i)_{k=1}^{N_i}$ *have an identical distribution, conditioned on* ω_0.

Then:

1. *There exists a Nash equilibrium that is independently randomized and symmetric, i.e., it belongs to* $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$.
2. *An independently randomized symmetric Nash equilibrium for* $\mathcal{P}_\infty$ *constitutes an approximate Nash equilibrium for* $\mathcal{P}_N$ *among all randomized policies.*

Proof method. First, use the Kakutani-Fan-Glicksberg fixed point theorem for a generic DM representing the team population and then the de Finetti Theorem.

- This indicates that games among teams are asymptotically equivalent to those among representative agents. We are not claiming uniqueness.

Existence of Symmetric Nash equilibrium for $\mathcal{P}_\infty$ and Approximate Nash equilibrium for $\mathcal{P}_N$

Theorem (SIAM J Cont and Opt'24)

Consider $\mathcal{P}_\infty$. Let

(i) $\mathbb{U}^i$ *is compact and* $c^i(\omega_0, \cdot, \cdot, \cdot)$ *is continuous for all* ω_0.

(ii) $(y_k^1)_{k=1}^{N_1}$ *and* $(y_k^2)_{k=1}^{N_2}$ *are independent, conditioned on* ω_0;

(iii) $(y_k^i)_{k=1}^{N_i}$ *have an identical distribution, conditioned on* ω_0.

Then:

1. *There exists a Nash equilibrium that is independently randomized and symmetric, i.e., it belongs to* $L^1_{\mathsf{PR,SYM}} \times L^2_{\mathsf{PR,SYM}}$.
2. *An independently randomized symmetric Nash equilibrium for* $\mathcal{P}_\infty$ *constitutes an approximate Nash equilibrium for* $\mathcal{P}_N$ *among all randomized policies.*

Proof method. First, use the Kakutani-Fan-Glicksberg fixed point theorem for a generic DM representing the team population and then the de Finetti Theorem.

- This indicates that games among teams are asymptotically equivalent to those among representative agents. We are not claiming uniqueness.

Outline

Large Continuous-Time Stochastic Teams: Dynamics

In our model, each DM has access to only a private state process that evolves as follows

$$dX_t^i = b_t(X_t^i, U_t^i)dt + \sigma_t(X_t^i)dW_t^i, \quad t \in \mathbb{T} := [0, T], \tag{*}$$

where $b_t(\cdot,\cdot)$ and $\sigma_t(\cdot)$ are continuous.

Suppose that (*) admits a unique, strong solution under non-anticipative open-loop policies and satisfies a nondegeneracy condition.

Also, we study the coupled dynamics using Girsanov's change of measures:

$$dX_t^i = b_t\left(X_t^i, U_t^i, \frac{1}{N}\sum_{p=1}^{N}\delta_{X_t^p}, \frac{1}{N}\sum_{p=1}^{N}\delta_{U_t^p}\right)dt + \sigma_t(X_t^i)dW_t^i, \quad t \in \mathbb{T},$$

For this talk, we focus on (*).

Large Continuous-Time Stochastic Teams: Expected Costs

Admissible policies: A policy $\boldsymbol{\gamma}^{\boldsymbol{i}}$ maps available information to actions such that the action process $\{U_t^i\}_t$ is adapted to the natural filtration generated by $X_{[0,t]}^i$ and t.
DMs collectively minimize the following expected cost function:

$$J_N(\boldsymbol{\gamma}^{\mathbf{1:N}}) := \mathbb{E}^{\boldsymbol{\gamma}^{1:N}} \left[\int_0^T \frac{1}{N} \sum_{i=1}^{N} c \left(X_t^i, U_t^i, \frac{1}{N} \sum_{p=1}^{N} \delta_{X_t^p}, \frac{1}{N} \sum_{p=1}^{N} \delta_{U_t^p} \right) dt \right].$$

The corresponding mean-field limit $\mathcal{P}^{\infty}$:

$$J_\infty(\underline{\boldsymbol{\gamma}}) = \limsup_{N\to\infty} J_N(\boldsymbol{\gamma}^{\mathbf{1:N}}).$$

Globally optimal solution: policy $\boldsymbol{\gamma}^{\mathbf{1\star:N\star}}$ such that

$$J_N(\boldsymbol{\gamma}^{\mathbf{1\star:N\star}}) = \inf_{\boldsymbol{\gamma}^{\mathbf{1:N}}} J_N(\boldsymbol{\gamma}^{\mathbf{1:N}}).$$

Convex Teams: Existence of a Symmetric Markovian Optimal Solution

Markovian Policies: $\boldsymbol{\gamma}^i$ is Markovian if for any $t \in [0,T]$, $\gamma_t^i : \mathbb{X} \to \mathbb{U}$ such that U_t^i is measurable with respect to the σ-field generated by only (t, X_t^i).

Theorem (ACC'24)

Consider $\mathcal{P}^N$. Then:

(i) *Without any loss, we can restrict the search for an optimal solution to Markovian policies.*

(ii) *If $\mathbb{U}$ is convex and $J(\boldsymbol{\gamma}^{\mathbf{1:N}})$ is convex in $\boldsymbol{\gamma}^{\mathbf{1:N}}$, then without any loss, we can restrict the search for an optimal solution to symmetric Markovian policies.*

Example: Linear quadratic teams.

- For convex teams with a finite number of DMs, we can restrict the search for an optimal solution to symmetric Markov policies.

Randomized Markov Policies

Definition

For each DM^i, the set of randomized Markov policies is the set of all Borel measurable functions $\boldsymbol{\nu^i} : [0,T] \times \mathbb{X} \to \mathcal{P}(\mathbb{U})$ such that the $\mathcal{P}(\mathbb{U})$-valued process $U_t^i = \nu_t^i(X_t^i)$ becomes adapted to $\sigma(X_t^i, t)$.

We endow the set of randomized Markov policies with Borkar's topology in [Borkar'89] under which $\boldsymbol{\nu_n^i}$ converges to $\boldsymbol{\nu^i}$ as $n \to \infty$ if

$$\lim_{n\to\infty} \int_0^T \int_{\mathbb{X}} f(t,x) \int_{\mathbb{U}} g(t,x,u)\nu_{n,t}^i(x)(du)\, dxdt = \int_0^T \int_{\mathbb{X}} f(t,x) \int_{\mathbb{U}} g(t,x,u)\nu_t^i(x)(du)\, dxdt$$

for all $f \in L_1([0,T] \times \mathbb{X}) \cap L_2([0,T] \times \mathbb{X})$ and $g \in \mathcal{C}_b([0,T] \times \mathbb{X} \times \mathbb{U})$. Denote this set of probability measures by $\mathcal{M}^i$. Let $\boldsymbol{\mu^i}(\cdot;\boldsymbol{\nu^i}) := \mathcal{L}(\boldsymbol{X^i})(\cdot)$, be a probability measure on path space $\mathcal{C}([0,T];\mathbb{X})$ induced by $\boldsymbol{\nu^i}$.

Lemma (Borkar'89)

Suppose that $\mathbb{U}$ is compact. Then:

- *$\mathcal{M}^i$ is compact.*
- *If $\boldsymbol{\nu_n^i}$ converges to $\boldsymbol{\nu^i}$, then $\boldsymbol{\mu^i}(\cdot;\boldsymbol{\nu_n^i})$ converges weakly to $\boldsymbol{\mu^i}(\cdot;\boldsymbol{\nu^i})$ as $n \to \infty$.*

Randomized Markov Policies

Definition

For each DM^i, the set of randomized Markov policies is the set of all Borel measurable functions $\boldsymbol{\nu}^{\boldsymbol{i}} : [0,T] \times \mathbb{X} \to \mathcal{P}(\mathbb{U})$ such that the $\mathcal{P}(\mathbb{U})$-valued process $U_t^i = \nu_t^i(X_t^i)$ becomes adapted to $\sigma(X_t^i, t)$.

We endow the set of randomized Markov policies with Borkar's topology in [Borkar'89] under which $\boldsymbol{\nu}_{\boldsymbol{n}}^{\boldsymbol{i}}$ converges to $\boldsymbol{\nu}^{\boldsymbol{i}}$ as $n \to \infty$ if

$$\lim_{n\to\infty} \int_0^T \int_{\mathbb{X}} f(t,x) \int_{\mathbb{U}} g(t,x,u)\nu_{n,t}^i(x)(du)\, dxdt = \int_0^T \int_{\mathbb{X}} f(t,x) \int_{\mathbb{U}} g(t,x,u)\nu_t^i(x)(du)\, dxdt$$

for all $f \in L_1([0,T] \times \mathbb{X}) \cap L_2([0,T] \times \mathbb{X})$ and $g \in \mathcal{C}_b([0,T] \times \mathbb{X} \times \mathbb{U})$. Denote this set of probability measures by $\mathcal{M}^i$. Let $\boldsymbol{\mu}^{\boldsymbol{i}}(\cdot;\boldsymbol{\nu}^{\boldsymbol{i}}) := \mathcal{L}(\boldsymbol{X}^{\boldsymbol{i}})(\cdot)$, be a probability measure on path space $\mathcal{C}([0,T];\mathbb{X})$ induced by $\boldsymbol{\nu}^{\boldsymbol{i}}$.

Lemma (Borkar'89)

Suppose that $\mathbb{U}$ is compact. Then:

- *$\mathcal{M}^i$ is compact.*
- *If $\boldsymbol{\nu}_{\boldsymbol{n}}^{\boldsymbol{i}}$ converges to $\boldsymbol{\nu}^{\boldsymbol{i}}$, then $\boldsymbol{\mu}^{\boldsymbol{i}}(\cdot;\boldsymbol{\nu}_{\boldsymbol{n}}^{\boldsymbol{i}})$ converges weakly to $\boldsymbol{\mu}^{\boldsymbol{i}}(\cdot;\boldsymbol{\nu}^{\boldsymbol{i}})$ as $n \to \infty$.*

Existence of Globally Optimal Solutions for $\mathcal{P}^N$

Theorem (ACC'24)

Consider $\mathcal{P}^N$ *. Let*

- c *be continuous in all its arguments.*
- $\mathbb{U}$ *be compact.*

Then:

1. *There exists a randomized globally optimal policy* $(\boldsymbol{\nu}^{\mathbf{1}}, \ldots, \boldsymbol{\nu}^{\boldsymbol{N}})$ *that is Markovian, i.e.,* $\boldsymbol{\nu}^{\boldsymbol{i}} \in \mathcal{M}^i$ *for all* $i \in \mathcal{N}$.
2. *If, additionally, the team is convex, then there exists a randomized globally optimal policy that is Markovian and symmetric.*

Proof method. We first show that $\boldsymbol{\nu}^{\mathbf{1:N}} \mapsto J_N(\boldsymbol{\nu}^{\mathbf{1:N}})$ is continuous. Then, since $\mathcal{M}^i$ is compact, there exists an optimal solution $\boldsymbol{\nu}^{\mathbf{1\star:N\star}}$ for $\mathcal{P}^N$ among randomized Markov policies. Since restricting the search for global optimality to randomized Markov policies is without loss, this completes the proof of part (i). Part (ii) follows a similar line of reasoning by restricting symmetric Markovian policies apriori.

Existence of Globally Optimal Solutions for $\mathcal{P}^\infty$

Theorem (ACC'24)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Let

- *c be continuous in all its arguments.*
- *$\mathbb{U}$ be compact.*
- *$\mathcal{P}^N$ be convex for every N.*

Then:

1. *Any sequence of randomized Markov globally optimal policies $\{\boldsymbol{\nu}_{\boldsymbol{N}}^\star\}_N$ for $\mathcal{P}^N$ that is symmetric admits a subsequence that converges to a globally optimal policy for $\mathcal{P}^\infty$ that is symmetric and Markovian.*
2. *There exists a randomized globally optimal policy for $\mathcal{P}^\infty$ that is Markovian and symmetric.*

Proof method. For a converging subsequence $\{\boldsymbol{\nu}_{\boldsymbol{n}}^\star\}_n \subseteq \mathcal{M}^i$, we show that empirical measures converge weakly almost surely using the Skorohod's representation theorem. Use generalized DCT to show

$$\limsup_{N\to\infty} J_N(\boldsymbol{\nu}_{\boldsymbol{N}}^\star, \ldots, \boldsymbol{\nu}_{\boldsymbol{N}}^\star) = J_\infty(\underline{\boldsymbol{\nu}^\star}).$$

Relaxing convexity: Sets of randomized policies-Part I

Let Γ^i be the set of all progressively measurable random probability measure processes $\boldsymbol{\gamma}^{\boldsymbol{i}}(\omega)$ on $[0, T] \times \mathbb{U}$ that satisfies the three following conditions:

(i) Marginals on $[0, T]$ are fixed to be a Lebesgue measure.

(ii) Random variable $\gamma_t^i : \omega \mapsto \boldsymbol{\gamma}^{\boldsymbol{i}}(\cdot|t)(\omega)$ is independent of $W_s^i - W_t^i$ for $s > t$ and for any $t \in \mathbb{T}$, and $W_{[0,T]}^j$ any $j \in \mathcal{N}$.

We equip Γ^i with the Young topology, that is, $\mathcal{P}([0, T] \times \mathbb{U}^i)$ is endowed with the weak convergence topology.

Sets of randomized policies-Part II

We denote the sets of all randomized policies P_π on $\prod_{i=1}^N \Gamma^i$ by $L^N := \mathcal{P}(\prod_{i=1}^N \Gamma^i)$ and we equip it with the product topology.

We define the following sets of randomized policies as subsets of L^N:

$$L^N_{\mathsf{EX}} := \left\{ P_\pi(\boldsymbol{\gamma}^{\mathbf{1}} \in \cdot, \ldots, \boldsymbol{\gamma}^{\boldsymbol{N}} \in \cdot) = P_\pi(\gamma^{\sigma(1)} \in \cdot, \ldots, \gamma^{\sigma(N)} \in \cdot) \quad \forall \sigma \in S_N \right\}$$

$$L^N_{\mathsf{CO,SYM}} := \left\{ P_\pi(\boldsymbol{\gamma}^{\mathbf{1}} \in \cdot, \ldots, \boldsymbol{\gamma}^{\boldsymbol{N}} \in \cdot) = \int_{z\in[0,1]} \prod_{i=1}^{N} \tilde{P}_\pi(\boldsymbol{\gamma}^{\boldsymbol{i}} \in \cdot | z)\eta(dz), \ \eta \in \mathcal{P}([0,1]) \right\}$$

$$L^N_{\mathsf{PR,SYM}} := \left\{ P_\pi(\boldsymbol{\gamma}^{\mathbf{1}} \in \cdot, \ldots, \boldsymbol{\gamma}^{\boldsymbol{N}} \in \cdot) = \prod_{i=1}^{N} \widetilde{P}_\pi(\boldsymbol{\gamma}^{\boldsymbol{i}} \in \cdot) \right\}.$$

Lemma (ACC'24)

Let $\mathbb{U}$ be compact. Suppose that $b_t(\cdot,\cdot)$ and $\sigma_t(\cdot)$ are uniformly bounded for all $t \in \mathbb{T}$. Then:

1. *L^N_{EX} is convex and compact.*
2. *The mapping $\Theta : L^N_{\mathsf{EX}} \to \mathcal{P}(\mathcal{C}([0,T],\mathbb{X}))$ is continuous.*

Existence of Randomized Globally Optimal Solutions for $\mathcal{P}^N$ and $\mathcal{P}^\infty$

Theorem (ACC'24)

Consider $\mathcal{P}^N$ and $\mathcal{P}^\infty$. Let

- *c be continuous in all its arguments.*
- *$\mathbb{U}$ be compact.*
- *$b_t(\cdot,\cdot)$ and $\sigma_t(\cdot)$ be uniformly bounded for all $t \in \mathbb{T}$.*

Then:

1. *There exists an exchangeable randomized globally optimal policy $P_\pi^{N\star}$ for $\mathcal{P}^N$ among all randomized policies L^N, i.e., $P_\pi^{N\star} \in L^N_{\mathsf{EX}}$.*
2. *There exists a symmetric privately randomized globally optimal policy $P_\pi^\star$ for $\mathcal{P}^\infty$ among all randomized policies L, i.e., $P_\pi^\star \in L_{\mathsf{PR,SYM}}$.*
3. *A symmetric privately randomized globally optimal policy $P_\pi^\star$ for $\mathcal{P}^\infty$ is approximately optimal for $\mathcal{P}^N$.*

Outline

Conclusion

Under sufficient conditions, we have shown that

1. Finite population convex teams: *an optimal solution exists that is symmetric and deterministic.*
2. Infinite population convex teams: *an optimal solution exists that is symmetric and independently randomized.* $\Longleftarrow$ *representative agent measure-valued MDP.*
3. Finite population non-convex teams: *an optimal solution exists that is exchangeable and randomized, but there exists an approximate optimal policy that is symmetric and independently randomized for sufficiently large teams.*
4. Infinite population non-convex teams: *an optimal solution exists that is symmetric and independently randomized.* $\Longleftarrow$ *representative agent measure-valued MDP.*
5. Large centralized teams: *a symmetric solution of N-DM mean-field sharing problem is nearly optimal for large centralized teams* $\Longleftarrow$ *single-agent mean-field sharing MDP*.
6. Justification in *Agent-Based-Modeling* and *Team-against-Team games*, via the representative agents defining the game played in equilibrium–Agent-based-Modeling is often used in economics, mathematical biology, and other sciences.