Team Optimal Solution of Finite Number of Mean-Field Coupled LQG Subsystems

Jalal Arabneydi and Aditya Mahajan

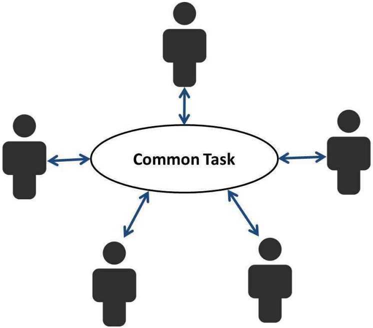
Electrical and Computer Engineering Department, McGill University

54th IEEE Conference on Decision and Control

Date: December 17th, 2015

What is Team?

- **What is team?** Any collection of decision makers (agents) that are interested to collaborate (team up) to accomplish a common task.



- In general, these problems belong to **NEXP complexity** class.

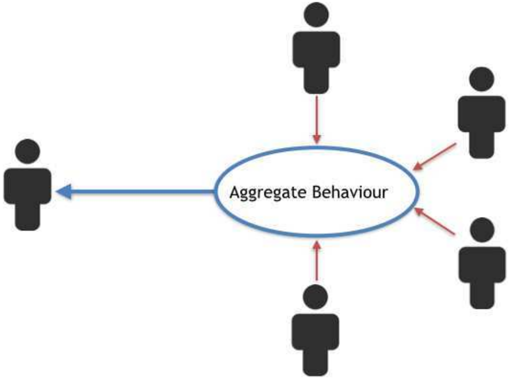
Brief literature review on Team Theory

- **Classical** information structure: All agents have identical information.
- **Non-classical** information structure: Agents have different information sets.

Examples of **non-classical** information structure:

- Static team (Radner 1962, Marschack and Radner 1972)
- Dynamic team (Witsenhausen 1971, Witsenhausen 1973)
- Specific information structure
  - Partially nested (Ho and Chu 1972)
  - One-step delayed sharing (Witsenhausen 1971, Yoshikawa 1978)
  - n-step delayed sharing (Witsenhausen 1971, Varaiya 1978, Nayyar 2011)
  - Common past sharing (Aicardi 1978)
  - Periodic sharing (Ooi 1997)
  - Belief sharing (Yuksel 2009)
  - Partial history sharing (Nayyar 2013)
  - This work analyses a new information structure : **Mean-field sharing**

What is Mean Field Team?

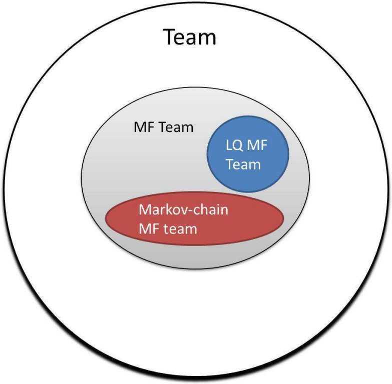


2

The number of decision makers is not necessarily large.

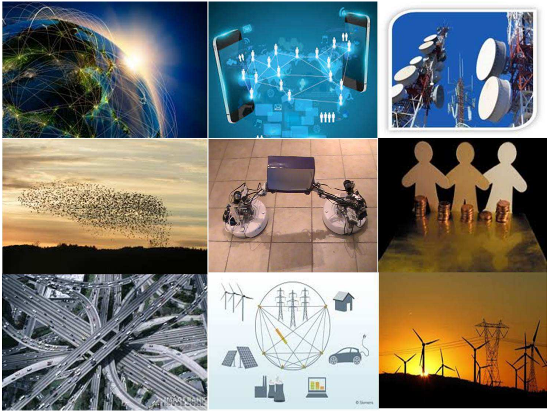
What is Mean Field Team?

- Key feature of mean-field teams is that the solution is **tractable**. In particular,
  - Markov chain mean-field team (J. Arabneydi and A. Mahajan, CDC 2014):
    - mean-field: **empirical distribution**
  - Linear quadratic mean-field team (J. Arabneydi and A. Mahajan, CDC 2015):
    - mean-field: **empirical mean**



Mean Field Team in Various Applications

- Team problems arise in various applications.
  - Networked control systems
  - Robotics
  - Communication networks
  - Transportation networks
  - Sensor networks
  - Smart grids
  - Economics
  - Etc.



Key Feature of LQ Mean Field Team

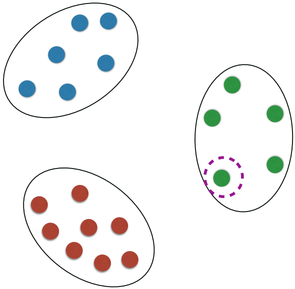
- LQ mean-field team is a team problem with non-classical information structure whose **computational complexity does not increase with number of agents**.

Outline

- Problem formulation and Main results
- Generalizations
- Numerical Example
- Summary and Conclusion

Notation

- N : number of heterogeneous agents (entire population)
- K : number of types (sub-populations)

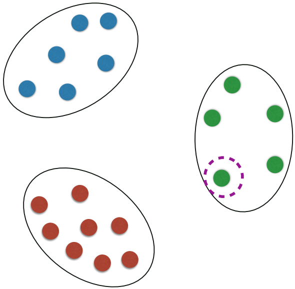


Notation

- N : number of heterogeneous agents (entire population)
- K : number of types (sub-populations)

For agent $i \in \mathcal{N}^k$ **of type** $k \in \mathcal{K}$

- $x_t^i \in \mathbb{R}^{d_x^k}$: state of agent i
- $u_t^i \in \mathbb{R}^{d_u^k}$: action of agent i



Notation

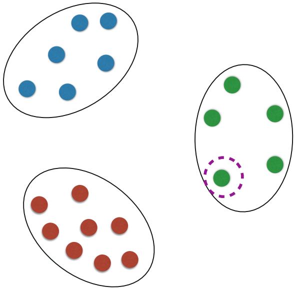
- N : number of heterogeneous agents (entire population)
- K : number of types (sub-populations)

For agent $i \in \mathcal{N}^k$ **of type** $k \in \mathcal{K}$

- $x_t^i \in \mathbb{R}^{d_x^k}$: state of agent i
- $u_t^i \in \mathbb{R}^{d_u^k}$: action of agent i

For sup-population of type $k \in \mathcal{K} = \{1, \dots, K\}$

- $\mathcal{N}^k$: entire sub-population of type k
- $\bar{x}_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} x_t^i$: mean-field of states
- $\bar{u}_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} u_t^i$: mean-field of actions



Notation

- N : number of heterogeneous agents (entire population)
- K : number of types (sub-populations)

For agent $i \in \mathcal{N}^k$ of type $k \in \mathcal{K}$

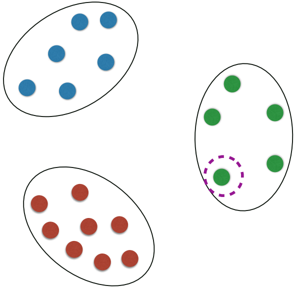
- $x_t^i \in \mathbb{R}^{d_x^k}$: state of agent i
- $u_t^i \in \mathbb{R}^{d_u^k}$: action of agent i

For sup-population of type $k \in \mathcal{K} = \{1, \ldots, K\}$

- $\mathcal{N}^k$: entire sub-population of type k
- $\bar{x}_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} x_t^i$: mean-field of states
- $\bar{u}_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} u_t^i$: mean-field of actions

For entire population

- $\mathcal{N} = \mathcal{N}^1 \bigcup \ldots \bigcup \mathcal{N}^K$: entire population
- $\mathbf{x}_t = (x_t^i)_{i \in \mathcal{N}}$: joint state of entire population at time t
- $\mathbf{u}_t = (u_t^i)_{i \in \mathcal{N}}$: joint action of entire population at time t
- $\bar{\mathbf{x}}_t = \text{vec}(\bar{x}_t^1, \ldots, \bar{x}_t^K)$: mean-field of states of entire population
- $\bar{\mathbf{u}}_t = \text{vec}(\bar{u}_t^1, \ldots, \bar{u}_t^K)$: mean-field of actions of entire population



Problem Formulation

- Dynamics of agent $i \in \mathcal{N}^k$ with type $k \in \{1, \ldots, K\}$:

$$x^i_{t+1} = A^k_t x^i_t + B^k_t u^i_t + D^k_t \bar{\mathbf{x}}_t + E^k_t \bar{\mathbf{u}}_t + w^i_t, \tag{1}$$

Problem Formulation

- Dynamics of agent $i \in \mathcal{N}^k$ with type $k \in \{1, \dots, K\}$:

$$x_{t+1}^i = A_t^k x_t^i + B_t^k u_t^i + D_t^k \bar{\mathbf{x}}_t + E_t^k \bar{\mathbf{u}}_t + w_t^i, \tag{1}$$

- Per-step cost: for $t = 1, \dots, T-1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \bar{\mathbf{x}}_t{}^\intercal P_t^x \bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t{}^\intercal P_t^u \bar{\mathbf{u}}_t + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \left[x_t^{i\intercal} Q_t^k x_t^i + u_t^{i\intercal} R_t^k u_t^i \right] \right] \tag{2}$$

and $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) = \bar{\mathbf{x}}_T{}^\intercal P_T^x \bar{\mathbf{x}}_T + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} x_T^{i\,\intercal} Q_T^k x_T^i \right], \tag{3}$$

where P_t^x, P_t^u, Q_t^k, and R_t^k are matrices of appropriate dimension; above matrices are symmetric and satisfy the following condition:

$$\begin{aligned} Q_t^k \geq 0 \ \ \forall k \in \mathcal{K}, \quad \operatorname{diag}\{Q_t^1, \dots, Q_t^K\} + P_t^x \geq 0, \\ R_t^k > 0 \ \ \forall k \in \mathcal{K}, \quad \operatorname{diag}\{R_t^1, \dots, R_t^K\} + P_t^u > 0. \end{aligned} \tag{4}$$

- **Mean-field sharing** Information structure: $U_t^i = g_t^i(x_{1:t}^i, \bar{\mathbf{x}}_{1:t})$, where g_t^i is called control law of agent i at time t.

Problem Formulation

- **Mean-field sharing** Information structure: $U_t^i = g_t^i(x_{1:t}^i, \bar{\mathbf{x}}_{1:t})$, where g_t^i is called control law of agent i at time t.
- Optimization problem: We are interested in finding a strategy $\mathbf{g}$ that minimizes

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)], \tag{5}$$

  where the expectation is with respect to the measure induced on all system variables by the choice of strategy $\mathbf{g}$.

Theorem (1)

Let

$$\bar{A}_t := \text{diag}\{A_t^1, \ldots, A_t^K\} + \text{vec}(D_t^1, \ldots, D_t^K), \quad \bar{Q}_t := \text{diag}\{Q_t^1, \ldots, Q_t^K\}$$
$$\bar{B}_t := \text{diag}\{B_t^1, \ldots, B_t^K\} + \text{vec}(E_t^1, \ldots, E_t^K), \quad \bar{R}_t := \text{diag}\{R_t^1, \ldots, R^K\}.$$

1 *Structure of optimal strategy: The optimal strategy is unique and is linear in local state and the mean-field of the system. In particular,*

$$u_t^i = \breve{L}_t^k(x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{6}$$

Our main Results

Theorem (1)

Let

$$\bar{A}_t := \operatorname{diag}\{A_t^1, \dots, A_t^K\} + \operatorname{vec}(D_t^1, \dots, D_t^K), \quad \bar{Q}_t := \operatorname{diag}\{Q_t^1, \dots, Q_t^K\}$$

$$\bar{B}_t := \operatorname{diag}\{B_t^1, \dots, B_t^K\} + \operatorname{vec}(E_t^1, \dots, E_t^K), \quad \bar{R}_t := \operatorname{diag}\{R_t^1, \dots, R^K\}.$$

1. *Structure of optimal strategy: The optimal strategy is unique and is linear in local state and the mean-field of the system. In particular,*

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{6}$$

where the above gains are obtained by the solution of $K+1$ Riccati equations: one for computing each $\breve{L}_t^k$, $k \in \mathcal{K}$, and one for $\bar{L}_t := \operatorname{vec}(\bar{L}_t^1, \dots, \bar{L}_t^K)$.

Theorem (1)

2 *Riccati equations: For* $t \in \{1, \ldots, T-1\}$,

$$\breve{L}_t^k = -\left(B_t^{k\intercal} \breve{M}_{t+1}^k B_t^k + R_t^k\right)^{-1} B_t^{k\intercal} \breve{M}_{t+1}^k A_t^k \tag{7}$$

and

$$\bar{L}_t = -\left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t, \tag{8}$$

where $\{\breve{M}_t^k\}_{t=1}^T$ *and* $\{\bar{M}_t\}_{t=1}^T$ *are the solutions of following Riccati equations:*

$$\breve{M}_T^k = Q_T^k, \quad \bar{M}_T = \bar{Q}_T + P_T^x, \tag{9}$$

and for $t = T-1, \ldots, 1$,

$$\begin{aligned}\breve{M}_t^k = &-A_t^{k\intercal} \breve{M}_{t+1}^k B_t^k \left(B_t^{k\intercal} \breve{M}_{t+1}^k B_t^k + R_t^k\right)^{-1} B_t^{k\intercal} \breve{M}_{t+1}^k A_t^k \\ &+ A_t^{k\intercal} \breve{M}_{t+1}^k A_t^k + Q_t^k,\end{aligned} \tag{10}$$

and

$$\begin{aligned}\bar{M}_t = &-\bar{A}_t^\intercal \bar{M}_{t+1} \bar{B}_t \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t \\ &+ \bar{A}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{Q}_t + P_t^x.\end{aligned} \tag{11}$$

Salient feature of the results

1. The optimal decentralized control laws perform as well as the optimal centralized control laws.

Salient feature of the results

1 The optimal decentralized control laws perform as well as the optimal centralized control laws.

2 The results generalize to **Infinite horizon**, **Tracking problem**, **Infinite Population**, and **Noisy observation**.

Generalization 1:Infinite horizon

The solution methodology naturally extend to infinite horizon under standard assumptions.

- Time average
- Discounted cost

Theorem (2)

1 *Structure of optimal strategy:*

$$u_t^i = \breve{L}^k (x_t^i - \bar{x}_t^k) + \bar{L}^k \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{12}$$

where above gains are obtained by the solution of $K+1$ algebraic Riccati equations.

2 *Algebraic Riccati equations: algebraic version of Riccati equations defined in Theorem 1, for time average and discounted cost problems.*

Generalization 2:Tracking problem

- Per-step cost:for $t = 1, \dots, T-1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = (\bar{\mathbf{x}}_t - \mathbf{s}_t)^\intercal P_t^x (\bar{\mathbf{x}}_t - \mathbf{s}_t) + \bar{\mathbf{u}}_t^\intercal P_t^u \bar{\mathbf{u}}_t$$
$$+ \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \left[(x_t^i - r_t^i)^\intercal Q_t^k (x_t^i - r_t^i) + u_t^{i\intercal} R_t^k u_t^i \right] \right],$$

and for $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) = (\bar{\mathbf{x}}_T - \mathbf{s}_T)^\intercal P_T^x (\bar{\mathbf{x}}_T - \mathbf{s}_T)$$
$$+ \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} (x_T^i - r_T^i)^\intercal Q_T^k (x_T^i - r_T^i) \right].$$

- Everything else remains the same as in the basic model.

Generalization 2:Tracking problem

Theorem (3)

1 *Structure of optimal strategy:*

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t + \breve{F}_t^k v_t^i + \bar{F}_t^k \bar{v}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{13}$$

where gains $\{\breve{L}_t^k, \bar{L}_t^k\}_{t=1}^{T-1}$ *are the same as in Theorem 1.*

Theorem (3)

1 *Structure of optimal strategy:*

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t + \breve{F}_t^k v_t^i + \bar{F}_t^k \bar{v}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{13}$$

where gains $\{\breve{L}_t^k, \bar{L}_t^k\}_{t=1}^{T-1}$ *are the same as in Theorem 1.*

2 *Riccati equations: Let* $\{\breve{M}_t^k\}_{t=1}^{T}$ *and* $\{\bar{M}_t\}_{t=1}^{T}$ *be the solution of (K+1) Riccati equations defined in Theorem 1. For* $t = 1, \dots, T-1$*:*

$$\breve{F}_t^k = \left(B_t^{k\intercal} \breve{M}_{t+1}^k B_t^k + R_t^k \right)^{-1} B_t^{k\intercal} \quad \text{and} \quad \bar{F}_t = \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u \right)^{-1} \bar{B}_t^\intercal, \tag{14}$$

where $\bar{F}_t =: \mathrm{vec}(\bar{F}_t^1, \dots, \bar{F}_t^K)$*. For* $t = T$*,*

$$v_T^i = Q_T^k r_T^i, \quad \bar{v}_T = \bar{Q}_T \bar{\mathbf{r}}_T + P_T^x \mathbf{s}_T \tag{15}$$

and for $t = T-1, \dots, 1$*,*

$$v_t^i = (A_t^k - B_t^k \breve{L}_t^k)^\intercal v_{t+1}^i + Q_t^k r_t^i \quad \text{and} \quad \bar{v}_t = (\bar{A}_t - \bar{B}_t \bar{L}_t)^\intercal \bar{v}_{t+1} + \bar{Q}_t \bar{\mathbf{r}}_t + P_t^x \mathbf{s}_t. \tag{16}$$

$$\bar{x}^k_{t+1} = A^k_t \bar{x}^k_t + B^k_t(\bar{L}^k_t \bar{\mathbf{x}}_t) + D^k_t \bar{\mathbf{x}}_t + E^k_t \bar{L}_t \bar{\mathbf{x}}_t + \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} w^i_t. \quad (17)$$

Theorem (4)

Structure of optimal strategy:

$$u^i_t = \breve{L}^k_t(x^i_t - z^k_t) + \bar{L}^k_t \mathbf{z}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \quad (18)$$

where $\{\breve{L}^k_t, \bar{L}^k_t\}_{t=1}^{T-1}$ *are same as in Theorem 1 and*

$$z^k_t = \begin{cases} \bar{x}^k_t, & \textit{sub-population } k \textit{ is finite} \\ A^k_{t-1} z^k_{t-1} + (B^k_{t-1} \bar{L}^k_{t-1} + D^k_{t-1} + E^k_{t-1} \bar{L}_{t-1}) \mathbf{z}_{t-1}, & \textit{sub-population } k \textit{ is infinite.} \end{cases} \quad (19)$$

Generalization 4: Noisy observation

- Let $y_t^i = C_t^k x_t^i + \bar{C}_t^k \bar{\mathbf{x}}_t + v_t^i$.
- Information structure: $u_t^i = g_t^i(y_{1:t}^i, \bar{\mathbf{x}}_{1:t})$.

Generalization 4: Noisy observation

- Let $y^i_t = C^k_t x^i_t + \bar{C}^k_t \bar{\mathbf{x}}_t + v^i_t$.
- Information structure: $u^i_t = g^i_t(y^i_{1:t}, \bar{\mathbf{x}}_{1:t})$.

Theorem (5)

1 *Structure of optimal strategy:*

$$u^i_t = \breve{L}^k_t(\hat{x}^i_t - \bar{x}^k_t) + \bar{L}^k_t \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{20}$$

where gains $\{\breve{L}^k_t, \bar{L}^k_t\}_{t=1}^{T-1}$ *are the same as in Theorem 1.*

- *Kalman filters: initial estimate* $\hat{x}^i_t = 0$, *and*

$$\hat{x}^i_{t+1} = A^k_t \hat{x}^i_t + B^k_t u^i_t + F^k_t(y^i_t - C^k_t \hat{x}^i_t - \bar{C}^k_t \bar{\mathbf{x}}_t) \tag{21}$$

where the Kalman filter gain is given by

$$F^k_t = A^k_t S^k_t {C^k_t}^\intercal (C^k_t S^k_t {C^k_t}^\intercal + \Sigma^k_v)^{-1}, \tag{22}$$

where the state estimation error covariances satisfy the (filter) Riccati equation: $S^k_1 = \Sigma^k_x$ *and for* $t > 1$,

$$S^k_{t+1} = A^k_t S^k_t {A^k_t}^\intercal - A^k_t S^k_t {C^k_t}^\intercal (C^k_t S^k_t {C^k_t}^\intercal + \Sigma^k_v)^{-1} C^k_t S^k_t {A^k_t}^\intercal + \Sigma^k_w. \tag{23}$$

Proof Approach

- **Step 1**: Construct an auxiliary system with complete centralized information i.e. $\mathbf{x}_t$.
- **Step 2**: Exploit the structure of the problem to solve the auxiliary system.
- **Step 3**: The obtained optimal strategy of auxiliary system is implementable under mean-field sharing.

Step 1: Auxiliary System

- Define $\breve{x}_t^i := x_t^i - \bar{x}_t^k$ and $\breve{u}_t^i := u_t^i - \bar{u}_t^k$.

Step 1: Auxiliary System

- Define $\breve{x}_t^i := x_t^i - \bar{x}_t^k$ and $\breve{u}_t^i := u_t^i - \bar{u}_t^k$.
- State and action : $\mathring{\mathbf{x}}_t = \text{vec}((\breve{x}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$ and $\mathring{\mathbf{u}}_t = \text{vec}((\breve{u}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$.
- Dynamics:

$$\breve{x}_{t+1}^i = A_t^k \breve{x}_t^i + B_t^k \breve{u}_t^i + w_t^i - \bar{w}_t^k \tag{24}$$

where $\bar{w}_t^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} w_t^i$ and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{25}$$

where $\bar{\mathbf{w}}_t := \text{vec}(\bar{w}_t^1, \dots, \bar{w}_t^K)$.

Step 1: Auxiliary System

- Define $\breve{x}_t^i := x_t^i - \bar{x}_t^k$ and $\breve{u}_t^i := u_t^i - \bar{u}_t^k$.
- State and action : $\mathring{\mathbf{x}}_t = \text{vec}((\breve{x}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$ and $\mathring{\mathbf{u}}_t = \text{vec}((\breve{u}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$.
- Dynamics:

$$\breve{x}_{t+1}^i = A_t^k \breve{x}_t^i + B_t^k \breve{u}_t^i + w_t^i - \bar{w}_t^k \tag{24}$$

where $\bar{w}_t^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} w_t^i$ and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{25}$$

where $\bar{\mathbf{w}}_t := \text{vec}(\bar{w}_t^1, \dots, \bar{w}_t^K)$.

- Per-step cost $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ at $t \leq T-1$ and terminal cost $c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)$ at $t = T$.

Step 1: Auxiliary System

- Define $\breve{x}_t^i := x_t^i - \bar{x}_t^k$ and $\breve{u}_t^i := u_t^i - \bar{u}_t^k$.
- State and action : $\mathring{\mathbf{x}}_t = \operatorname{vec}((\breve{x}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$ and $\mathring{\mathbf{u}}_t = \operatorname{vec}((\breve{u}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$.
- Dynamics:

$$\breve{x}_{t+1}^i = A_t^k \breve{x}_t^i + B_t^k \breve{u}_t^i + w_t^i - \bar{w}_t^k \tag{24}$$

where $\bar{w}_t^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} w_t^i$ and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{25}$$

where $\bar{\mathbf{w}}_t := \operatorname{vec}(\bar{w}_t^1, \dots, \bar{w}_t^K)$.

- Per-step cost $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ at $t \leq T - 1$ and terminal cost $c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)$ at $t = T$.
- Information structure:

$$\mathring{\mathbf{u}}_t = \mathring{g}_t(\mathring{\mathbf{x}}_{1:t}, \mathring{\mathbf{u}}_{1:t-1}).$$

Step 1: Auxiliary System

- Define $\breve{x}_t^i := x_t^i - \bar{x}_t^k$ and $\breve{u}_t^i := u_t^i - \bar{u}_t^k$.
- State and action : $\mathring{\mathbf{x}}_t = \text{vec}((\breve{x}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$ and $\mathring{\mathbf{u}}_t = \text{vec}((\breve{u}_t^i)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$.
- Dynamics:

$$\breve{x}_{t+1}^i = A_t^k \breve{x}_t^i + B_t^k \breve{u}_t^i + w_t^i - \bar{w}_t^k \tag{24}$$

where $\bar{w}_t^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} w_t^i$ and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{25}$$

where $\bar{\mathbf{w}}_t := \text{vec}(\bar{w}_t^1, \dots, \bar{w}_t^K)$.

- Per-step cost $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ at $t \leq T-1$ and terminal cost $c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)$ at $t = T$.
- Information structure:

$$\mathring{\mathbf{u}}_t = \mathring{g}_t(\mathring{\mathbf{x}}_{1:t}, \mathring{\mathbf{u}}_{1:t-1}).$$

- Optimization problem: we are interested in finding strategy $\mathring{\mathbf{g}} := (\mathring{g}_1, \dots, \mathring{g}_T)$ that minimizes

$$\mathring{J}(\mathring{\mathbf{g}}) = \mathbb{E}^{\mathring{\mathbf{g}}}[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)], \tag{26}$$

Step 2: Solution of Auxiliary System

Lemma (1)

For any time t, $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t)$ *such that for* $t = 1, \dots, T-1$,

$$\mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t) = \bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}_t^k(\breve{x}_t^i, \breve{u}_t^i),$$

and $t = T$,

$$\mathring{c}_T(\mathring{\mathbf{x}}_T) = \bar{c}_T(\bar{\mathbf{x}}_T) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}_T^k(\breve{x}_T^i),$$

where for $t = 1, \dots, T-1$,

$$\bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \bar{\mathbf{x}}_t^\intercal(\bar{Q}_t + P_t^x)\bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal(\bar{R}_t + P_t^u)\bar{\mathbf{u}}_t,$$

$$\breve{c}_t^k(\breve{x}_t^i, \breve{u}_t^i) = \frac{1}{|\mathcal{N}^k|}\left[\breve{x}_t^{i\,\intercal} Q_t^k \breve{x}_t^i + \breve{u}_t^{i\,\intercal} R_t^k \breve{u}_t^i\right],$$

and $t = T$,

$$\bar{c}_T(\bar{\mathbf{x}}_T) = \bar{\mathbf{x}}_T^\intercal(\bar{Q}_T + P_T^x)\bar{\mathbf{x}}_T, \quad \breve{c}_T^k(\breve{x}_T^i) = \frac{1}{|\mathcal{N}^k|}\left[\breve{x}_T^{i\,\intercal} Q_T^k \breve{x}_T^i\right].$$

Step 2: Solution of Auxiliary System

- Consider

$$\breve{x}_{t+1}^{i} = A_t^k \breve{x}_t^i + B_t^k \breve{u}_t^i + w_t^i - \bar{w}_t^k \tag{27}$$

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{28}$$

Step 2: Solution of Auxiliary System

- Consider

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t + w^i_t - \bar{w}^k_t \tag{27}$$

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{28}$$

- By virtue of certainty equivalence Theorem,

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t \tag{29}$$

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t \tag{30}$$

- Consider

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t + w^i_t - \bar{w}^k_t \tag{27}$$

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{28}$$

- By virtue of certainty equivalence Theorem,

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t \tag{29}$$

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t \tag{30}$$

- Deterministic auxiliary system consists on $(N+1)$ decoupled components:
  - N components with state $\breve{x}^i_t$ and action $\breve{u}^i_t$, $i \in \mathcal{N}^k$, $k \in \mathcal{K}$, with quadratic cost $\breve{c}^k_t(\breve{x}^i_t, \breve{u}^i_t)$.
  - one component with state $\bar{\mathbf{x}}_t$ and action $\bar{\mathbf{u}}_t$ with quadratic cost $\bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$.

Theorem (5)

The optimal control strategy of auxiliary model is unique and given by

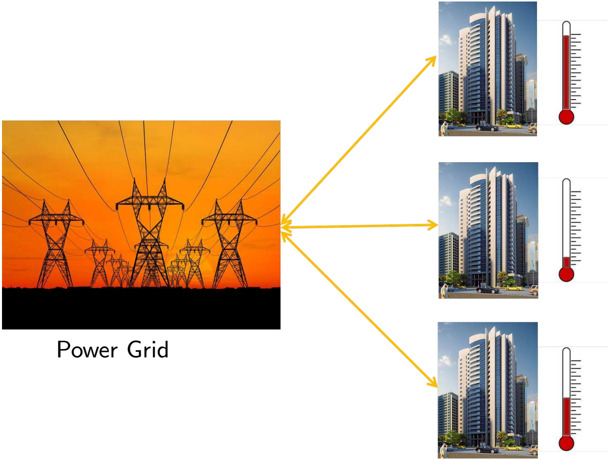
$$\breve{u}_t^i = \breve{L}_t^k \breve{x}_t^i, \quad \bar{\mathbf{u}}_t = \bar{L}_t \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K}, \tag{31}$$

where the gains $\{\breve{L}_t^k\}_{t=1}^{T-1}$ and $\{\bar{L}_t\}_{t=1}^{T-1}$ are given as in Theorem 1.

Note that

$$u_t^i = \breve{u}_t^i + \bar{u}_t^k = \breve{L}_t^k(x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t.$$

Example: Demand Response in Smart Grids



Keep mean-field close to desired reference trajectory with minimum discomfort for heaters.

- Dynamics of agent i :

$$x_{t+1}^i = A_t x_t^i + B_t u_t^i + w_t^i, \tag{32}$$

Example: Demand Response in Smart Grids

- Dynamics of agent i :

$$x^i_{t+1} = A_t x^i_t + B_t u^i_t + w^i_t, \tag{32}$$

- Per-step cost:for $t = 1, \dots, T-1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = (\bar{\mathbf{x}}_t - \mathbf{s}_t)^\intercal P^x_t (\bar{\mathbf{x}}_t - \mathbf{s}_t) + \bar{\mathbf{u}}_t^\intercal P^u_t \bar{\mathbf{u}}_t + \frac{1}{N}\left[\sum_{i=1}^{N}\left[(x^i_t - x^i_1)^\intercal Q_t (x^i_t - x^i_1) + x^{i\intercal}_t H_t x^i_t + u^{i\intercal}_t R_t u^i_t\right]\right],$$

and for $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) = (\bar{\mathbf{x}}_T - \mathbf{s}_T)^\intercal P^x_T (\bar{\mathbf{x}}_T - \mathbf{s}_T) + \frac{1}{N}\left[\sum_{i=1}^{N}(x^i_T - x^i_1)^\intercal Q_T (x^i_T - x^i_1) + x^{i\intercal}_T H_T x^i_T\right].$$

Extreme control

Summary & Conclusion

- We present a class of team problems that we call linear quadratic mean-field teams with non-classical information structure.

Summary & Conclusion

- We present a class of team problems that we call linear quadratic mean-field teams with non-classical information structure.

- We identify team-optimal solution for arbitrary number of agents. In particular, we show that optimal strategy is unique, identical across sub-populations, and linear in local state and (global) mean-field.

Summary & Conclusion

- We present a class of team problems that we call linear quadratic mean-field teams with non-classical information structure.
- We identify team-optimal solution for arbitrary number of agents. In particular, we show that optimal strategy is unique, identical across sub-populations, and linear in local state and (global) mean-field.
- To compute the optimal gains, we obtain $K+1$ standard Ricatti equations that do not depend on the size of population; hence, the solution and the solution complexity is independent of the number of agents.

Summary & Conclusion

- We present a class of team problems that we call linear quadratic mean-field teams with non-classical information structure.
- We identify team-optimal solution for arbitrary number of agents. In particular, we show that optimal strategy is unique, identical across sub-populations, and linear in local state and (global) mean-field.
- To compute the optimal gains, we obtain $K + 1$ standard Ricatti equations that do not depend on the size of population; hence, the solution and the solution complexity is independent of the number of agents.
- The results generalize to infinite horizon, tracking problem, infinite population, and noisy observation.

Thank You

- Let $y_t^i = C_t^k x_t^i + \bar{C}_t^k \bar{\mathbf{x}}_t + v_t^i$.
- Information structure: $u_t^i = g_t^i(y_{1:t}^i, \bar{\mathbf{x}}_{1:t})$.

Generalization 4: Noisy observation

- Let $y_t^i = C_t^k x_t^i + \bar{C}_t^k \bar{\mathbf{x}}_t + v_t^i$.
- Information structure: $u_t^i = g_t^i(y_{1:t}^i, \bar{\mathbf{x}}_{1:t})$.

Theorem (5)

1 *Structure of optimal strategy:*

$$u_t^i = \breve{L}_t^k(\hat{x}_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K} \tag{33}$$

where gains $\{\breve{L}_t^k, \bar{L}_t^k\}_{t=1}^{T-1}$ *are the same as in Theorem 1.*

- *Kalman filters: initial estimate* $\hat{x}_t^i = 0$*, and*

$$\hat{x}_{t+1}^i = A_t^k \hat{x}_t^i + B_t^k u_t^i + F_t^k(y_t^i - C_t^k \hat{x}_t^i - \bar{C}_t^k \bar{\mathbf{x}}_t) \tag{34}$$

where the Kalman filter gain is given by

$$F_t^k = A_t^k S_t^k C_t^{k\intercal}(C_t^k S_t^k C_t^{k\intercal} + \Sigma_v^k)^{-1}, \tag{35}$$

where the state estimation error covariances satisfy the (filter) Riccati equation: $S_1^k = \Sigma_x^k$ *and for* $t > 1$,

$$S_{t+1}^k = A_t^k S_t^k A_t^{k\intercal} - A_t^k S_t^k C_t^{k\intercal}(C_t^k S_t^k C_t^{k\intercal} + \Sigma_v^k)^{-1} C_t^k S_t^k A_t^{k\intercal} + \Sigma_w^k. \tag{36}$$