Mean Field Teams

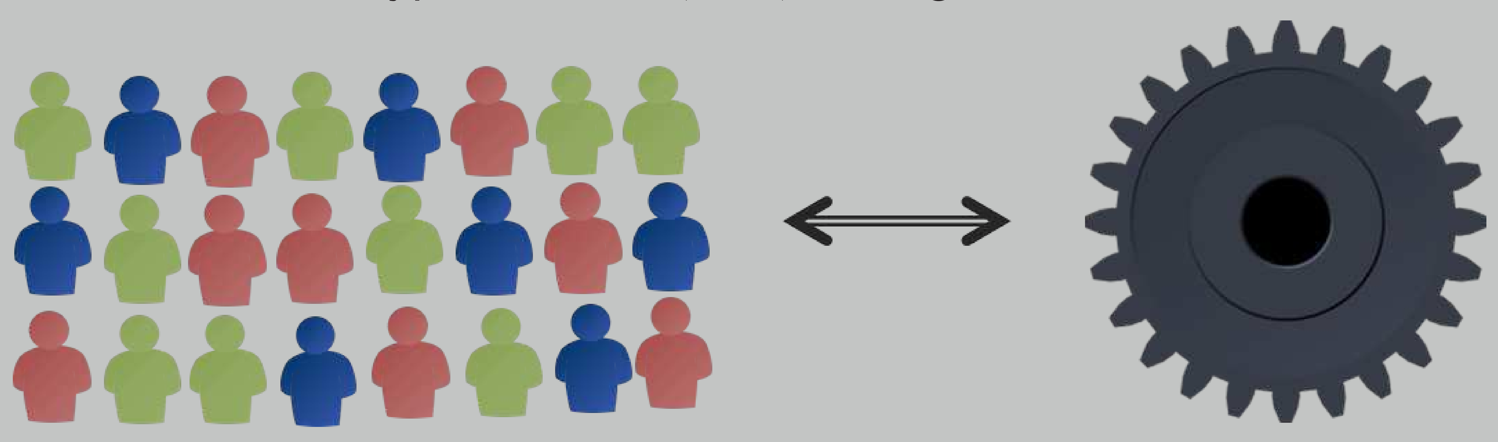
Jalal Arabneydi and Aditya Mahajan
ECE Department, McGill University

Motivation

- Mean-field teams arises in many applications ranging from **networked control systems**, **robotics**, **communication networks**, **transportation networks**, **sensor networks**, **economics**, and **smart grids**.

Notation

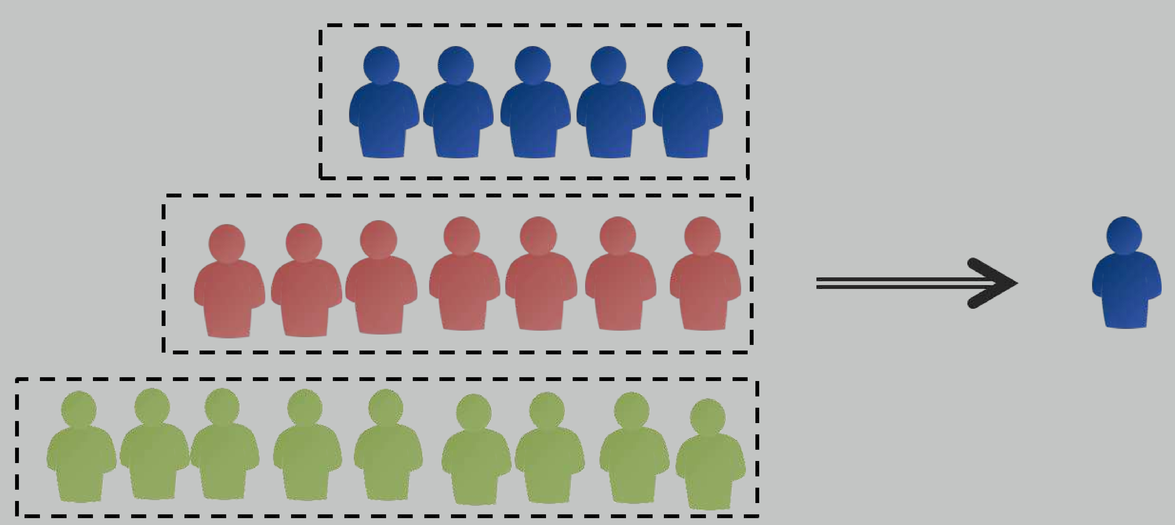
- Agents: $\mathcal{K}$ different types with $\mathcal{N}^1, \dots, \mathcal{N}^K$ agents.



- State and action of agent $i \in \mathcal{N}^k, k \in \mathcal{K}$: $X_t^i \in \mathcal{X}^k$ and $U_t^i \in \mathcal{U}^k$.
- Joint state and action: $\mathbf{X}_t = (X_t^i)_{i \in \bigcup_{k\in\mathcal{K}} \mathcal{N}^k}$ and $\mathbf{U}_t = (U_t^i)_{i \in \bigcup_{k\in\mathcal{K}} \mathcal{N}^k}$.
- Mean-field of type $k \in \mathcal{K}$ (empirical distribution): $Z_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \delta_{X_t^i}$.
- Mean-field of system: $\mathbf{Z}_t = (Z_t^1, \dots, Z_t^K)$.

Model and Main Results

- Dynamics of agent $i \in \mathcal{N}^k, k \in \mathcal{K}$: $X_{t+1}^i = f_t(X_t^i, U_t^i, W_t^i, \mathbf{Z}_t)$, where noise $\{W_t^i\}_{t=1}^T$ is an independent process.



- Information structure: $U_t^i = g_t^i(X_t^i, \mathbf{Z}_{1:t})$.
- Optimization problem: $min_{\mathbf{g}} \mathbb{E}^{\mathbf{g}} \left[\sum_{t=1}^T \ell_t(\mathbf{X}_t, \mathbf{U}_t) \right]$.
- **Assumption**: identical control laws within types i.e. $g_t^i = \tilde{g}_t^k, \forall i \in \mathcal{N}^k$. In large-scale systems, it is reasonable to treat subsystems identically, for the purposes of simplicity, robustness, and fairness.

- **Theorem 1:** Let $\psi_t^*(\mathbf{z}_t)$ denote any argmin of the right-hand side of following dynamic program. For $t = T, \dots, 1$, and for z_t,

$$V_t(\mathbf{z}_t) := \min_{\gamma_t} \mathbb{E}[\ell(\mathbf{X}_t, \mathbf{U}_t) + V_{t+1}(\mathbf{Z}_{t+1}) | \mathbf{Z}_t = \mathbf{z}_t, \Gamma_t = \gamma_t] \quad (1)$$

  where $V_{T+1}(\mathbf{z}_{T+1}) = 0$ and the minimization is over all functions $\gamma_t = (\gamma_t^1, \dots, \gamma_t^K), \gamma_t^k : \mathcal{X}^k \to \mathcal{U}^k, k \in \mathcal{K}$. Then, $\tilde{g}_t^k(x, \mathbf{z}) := \psi_t^{*,k}(\mathbf{z})(x)$. is an optimal control law.
- Information state of the dynamic program is polynomial in number of agents.

Linear Quadratic Model and Main Results

- Mean-field of type $k \in \mathcal{K}$ (empirical mean): $Z_t^k = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} X_t^i$.
- Dynamics of agent $i \in \mathcal{N}^k, k \in \mathcal{K}$: $X_{t+1}^i = A_t^k X_t^i + B_t^k U_t^i + W_t^i + D_t^k \mathbf{Z}_t$.
- Information structure: $U_t^i = g_t^i(X_t^i, \mathbf{Z}_{1:t})$.
- Optimization problem: $\min_{\mathbf{g}} \mathbb{E}^{\mathbf{g}} \left[\mathbf{Z}_t{}^\intercal P_t \mathbf{Z}_t + \sum_{k=1}^K \frac{\sum_{i \in \mathcal{N}^k} X_t^{i\intercal} Q_t^k X_t^i + U_t^{i\intercal} R^k U_t^i}{|\mathcal{N}^k|} \right]$.

- **Theorem 2:** The optimal strategy is unique and is linear in local state and the mean-field of the system. In particular,

$$U_t^i = L_t^k (X_t^i - Z_t^k) + H_t^k \mathbf{Z}_t \quad (2)$$

  where above gains are obtained by solution of $K+1$ standard Riccati equations.
- The Riccati equations do not depend on the number of agents.
- The above optimal decentralized strategy obtains centralized performance.

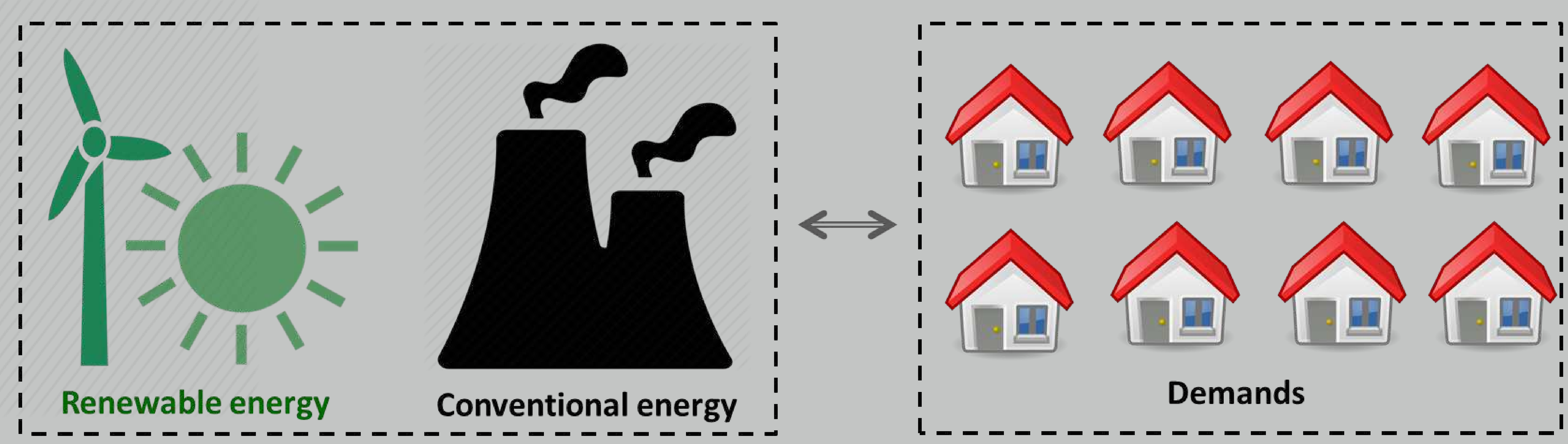
Generalizations

- Infinite horizon, Infinite population, Noisy observations, Mean-field of actions.

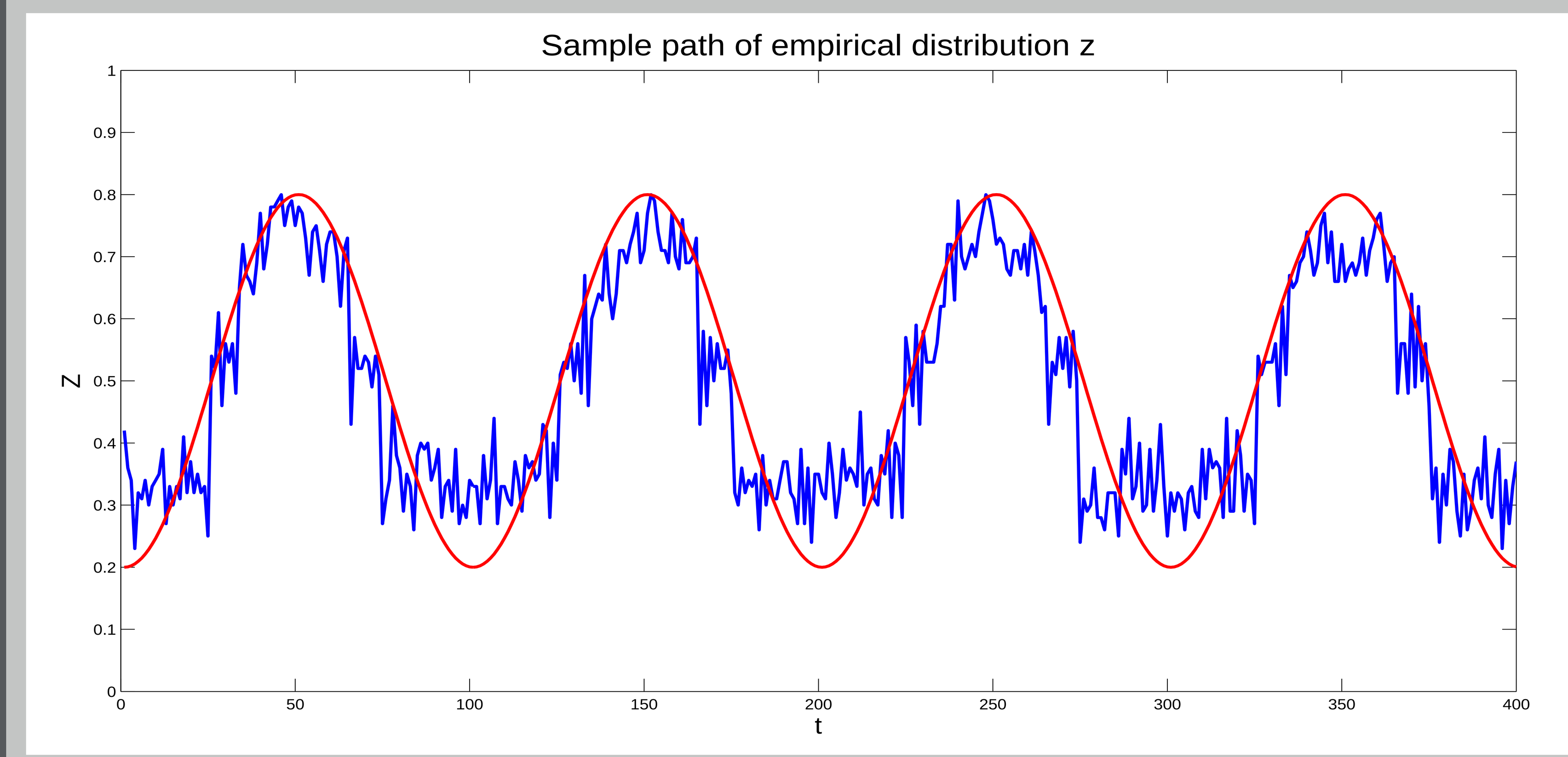
Salient Features

- The size of the information state of the dynamic program increases polynomially with the number of agents rather than exponentially.
- In linear quadratic mean-field teams, the computational complexity of the solution is independent of the number of agents.
- Since the obtained strategies are based on common observed data, each agent can independently compute the optimal strategy in a decentralized manner.
- Mean-field of the system can be computed and communicated easily.
- In infinite population, mean-field becomes deterministic and computable.

Example 1: Demand Response

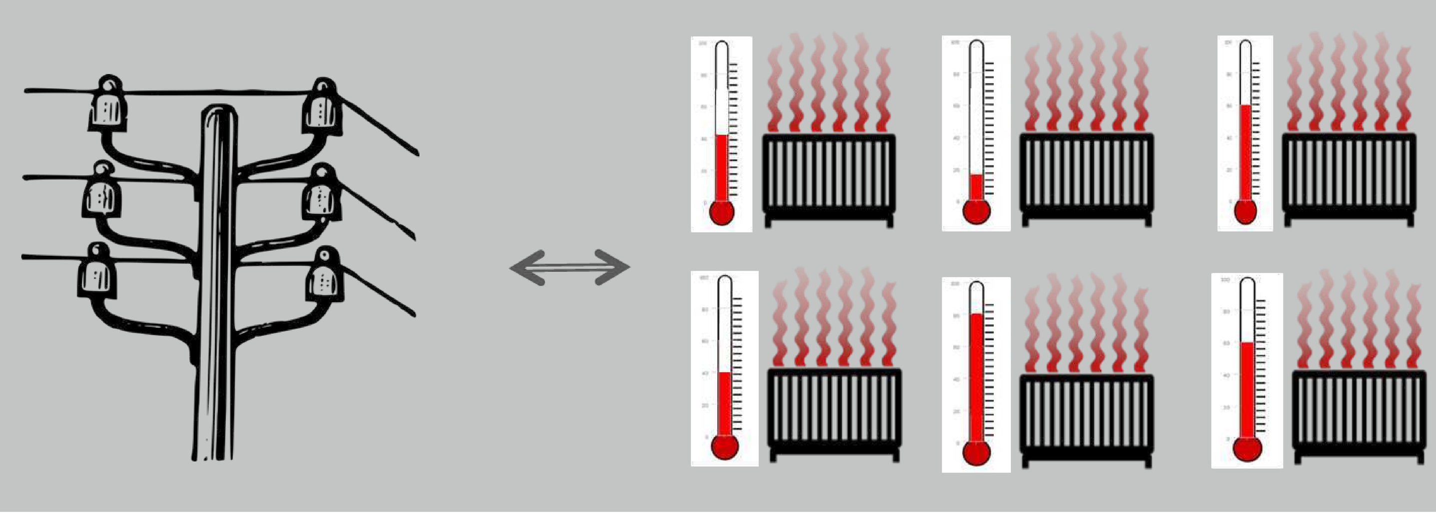


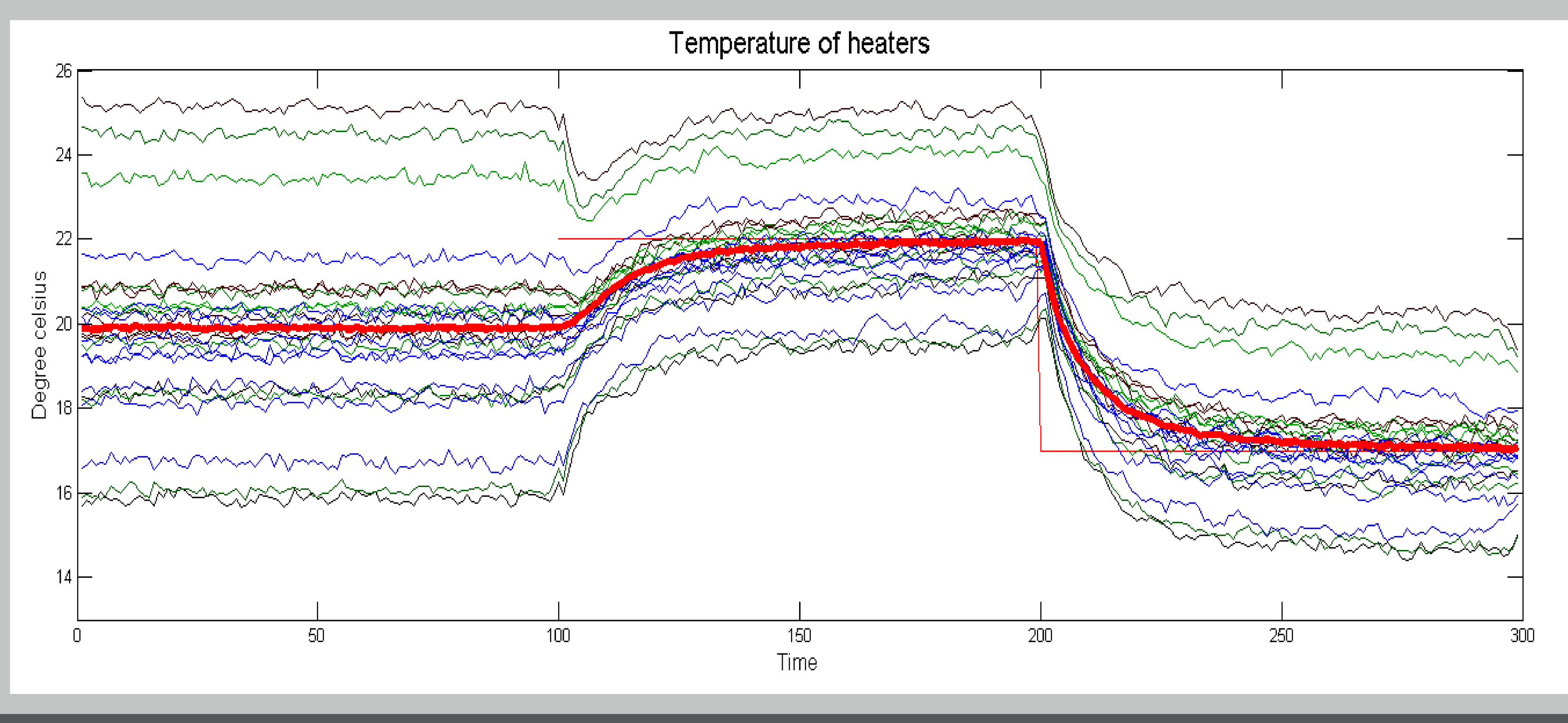
- $X_t^i \in \mathcal{X} = \{OFF, ON\}$, $\quad Z_t = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_t^i = OFF)$
- Dynamics: $\mathbb{P}(X_{t+1}^i | X_t^i, U_t^i) =: [P(u_t^i)]_{x_t^i x_{t+1}^i}$
- Actions: $U_t^i \in \mathcal{U} = \{FREE, OFF, ON\}$, $\quad$ Cost of action: $C(U_t^i)$
- Objective: $\min_{\mathbf{g}} \mathbb{E}^{\mathbf{g}} \left[\sum_{t=1}^\infty \beta^t \left(\frac{1}{n} \sum_{i=1}^n C(U_t^i) + D(Z_t \parallel \zeta_t) \right) \right]$.



Example 2: Temperature Control

- Control the average temperature with minimum intervention at heaters.





References

[1] Jalal Arabneydi and Aditya Mahajan.
Team optimal control of coupled subsystems with mean-field sharing.
Conference on Decision and Control (CDC), 2014.

[2] Jalal Arabneydi and Aditya Mahajan.
Team optimal solution of finite number of mean-field coupled lqg subsystems.
Conference on Decision and Control (CDC), 2015.

Email:jalal.arabneydi@mail.mcgill.ca
Email:aditya.mahajan@mcgill.ca