Alternating Inertial and Overrelaxed Algorithms for Distributed Generalized Nash Equilibrium Seeking in Multi-Player Games

Author: **Zhangcheng Feng**, Wenying Xu and Jinde Cao

Southeast University

ICCM 2022
2022.7.30.

OUTLINE

- Background and Motivation
- Main Results
- Conclusions and related works

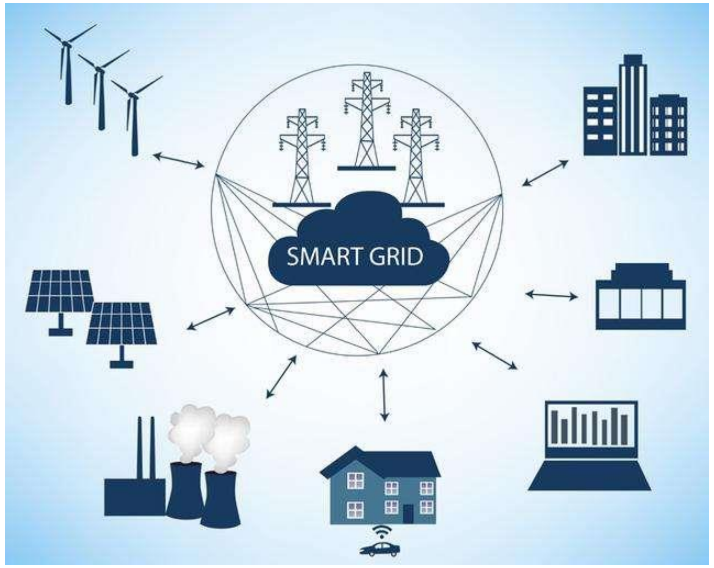
Background —— Example: Smart Grid

- **Smart grid**

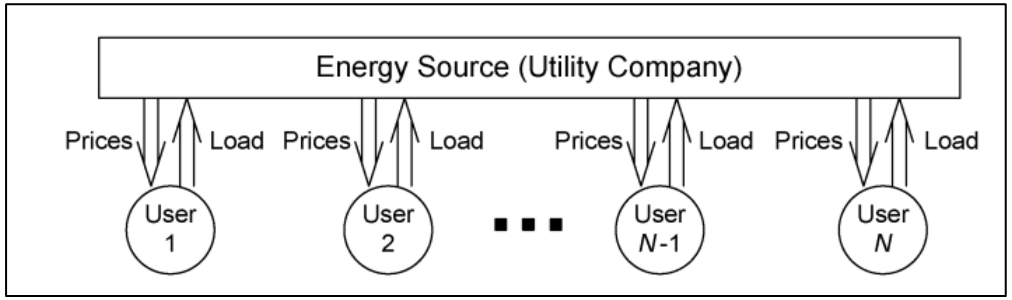
  ➢Economic Dispatch (ED)

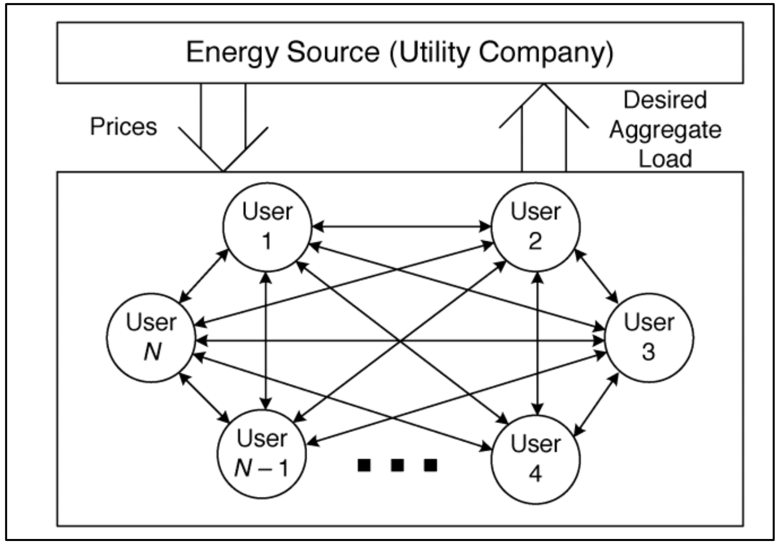
$$\max_{P_i} \quad \text{Profits}_i(P_i)$$

$$s.t. \quad \sum_i P_i = Total\ Demand$$



  ➢Demand-Side Management

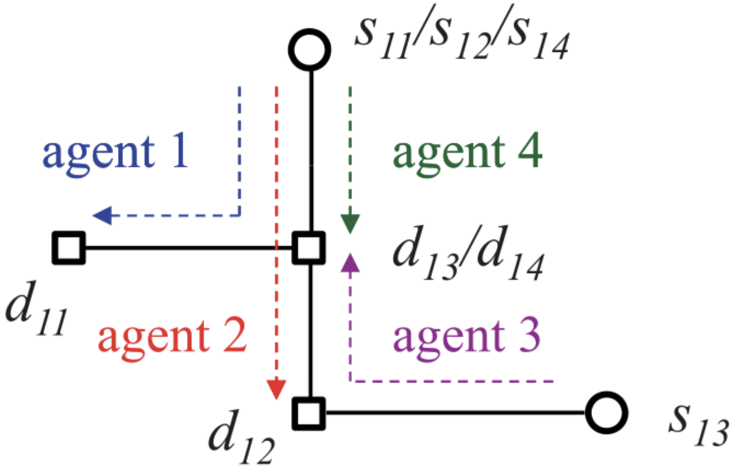




Mohsenian-Rad A H, Wong V W S, Jatskevich J, et al. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid[J]. IEEE transactions on Smart Grid, 2010, 1(3): 320-331.

Background —— Other Examples

- Ad-hoc wireless communication network



- Competitive economy, environmental pollution control……

Tekin C, Liu M, Southwell R, et al. Atomic congestion games on graphs and their applications in networking[J]. IEEE/ACM Transactions on Networking, 2012, 20(5): 1541-1552.

Motivating Features of Partial-Information

- Interdependence among decision makers
    - *Access decision information of opponents*
- Large-scale networks (thousands of decision makers)
    - *Locality of information*

Power outages

Traffic jam

OUTLINE

- Background and Motivation
- Main Results
- Conclusions and related works

Main Results —— Game Theoretic-Setup

- N agents (decision makers)
- Each agent i has decision variable (strategy) x_i

$$\boldsymbol{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \text{ ("all")}; \qquad (\forall i)\ \boldsymbol{x}_{-i} := \begin{pmatrix} \vdots \\ x_{i-1} \\ x_{i+1} \\ \vdots \end{pmatrix} \text{ ("others")}$$

- Each agent i has its individual objective:

$$\min_{x_i} \quad J_i(x_i, \boldsymbol{x}_{-i}) \longleftarrow \textit{Cost function}$$

$$s.t. \quad x_i \in \mathrm{constraints}_i(\boldsymbol{x}_{-i})$$

- "Game": = { inter-dependent optimization problems }

Generalized Nash Equilibrium Problem

- **Game:**

$$\forall i, P_i(\boldsymbol{x}_{-i}): \begin{cases} \min\limits_{x_i} \quad J_i(x_i, \boldsymbol{x}_{-i}) \\ s.t. \quad x_i \in \Omega_i \\ \qquad\quad A_i x_i + \sum\limits_{j \neq i} A_j x_j \leq b \end{cases}$$

Local constraint

Shared coupling constraint

——"Generalized"

- **Generalized Nash equilibrium (GNE):**

$$\boldsymbol{x}^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_N^* \end{pmatrix} \in \begin{pmatrix} \operatorname{argmin} P_1(\boldsymbol{x}_{-1}^*) \\ \vdots \\ \operatorname{argmin} P_N(\boldsymbol{x}_{-N}^*) \end{pmatrix}$$

Lagrangian Method —— Decouple Constraints

- **Lagrangian functions:**

$$L_i(x_i, \boldsymbol{x}_{-i}; \lambda_i) = J_i(x_i, \boldsymbol{x}_{-i}) + \lambda_i^\top (A\boldsymbol{x} - b)$$

- **KKT conditions:**

$$\begin{cases} \nabla_{x_i} L_i(x_i, \boldsymbol{x}_{-i}; \lambda_i) = 0 \\ A\boldsymbol{x} - b \le 0 \\ \lambda_i^\top (A\boldsymbol{x} - b) = 0 \\ \lambda_i \ge 0 \end{cases}$$

operator theory

$\Rightarrow$

(N —— normal cone)

$$\begin{cases} \mathbf{0} \in F(\boldsymbol{x}) + A^\top \lambda + N_\Omega(\boldsymbol{x}) \\ \mathbf{0} \in -(A\boldsymbol{x} - b) + N_{\ge 0}(\lambda) \end{cases}$$

- **variational GNE (v-GNE)** —— $\lambda_1 = \lambda_2 = \cdots = \lambda$

- **Pseudo-gradient mapping:**

$$F(\boldsymbol{x}) := \begin{pmatrix} \nabla_{x_1} J_1(x_1, \boldsymbol{x}_{-1}) \\ \vdots \\ \nabla_{x_N} J_N(x_N, \boldsymbol{x}_{-N}) \end{pmatrix}$$

Main Results —— Assumptions

Assumptions:

- $J_i(\cdot, \boldsymbol{x}_{-i})$ continuous differentiable, convex
- Ω_i non-empty, compact and convex
- $\boldsymbol{\Omega} \cap \{\boldsymbol{x}: A\boldsymbol{x} \leq b\}$ non-empty, Slater's constraint qualification
- F Lipschitz continuous, *strongly monotone*

$$\langle x-y, F(x)-F(y)\rangle \geq \mu \| x-y \|^2 (\mu > 0)$$

v-GNE **exists**

Partial-Decision Information Setup I

- **Problem:**

  Agent i **does not know** $\boldsymbol{x}_{-i}$ $\Rightarrow$ cannot compute $J_i(x_i, \boldsymbol{x}_{-i})$ and $\nabla_{x_i} J_i(x_i, \boldsymbol{x}_{-i})$

- **Solution:**

  Agent i **estimates** $\boldsymbol{x}_{-i}$

  ➢ x_i^j := estimation by agent i of x_j ($x_j \rightarrow x_1^j, \cdots, x_N^j$)

  ➢ x_i^i := x_i

  ➢ $\boldsymbol{x}_i^{-i}$:= estimation by agent i of $\boldsymbol{x}_{-i}$

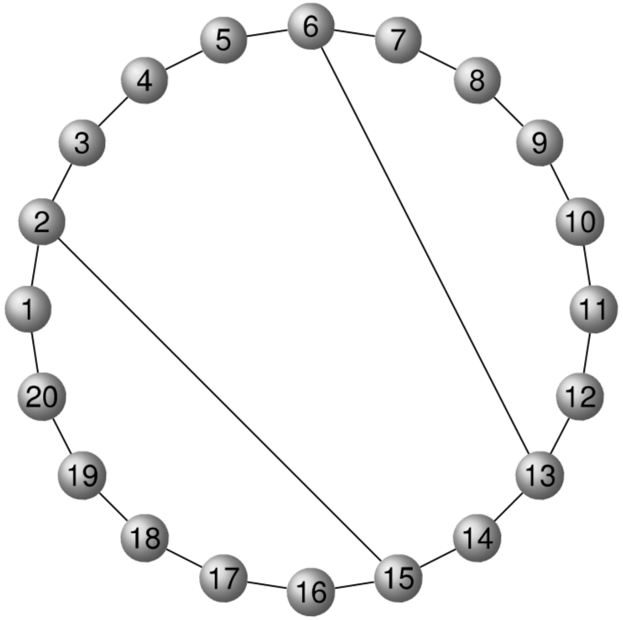


Fig. 2. Communication graph among all firms, where an edge from i to j means that firm i and j can exchange information.

- **Graph properties:**

  Agents exchange decision and estimation variables with neighbors on a undirected, connected graph

---

Pavel L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach[J]. IEEE Transactions on Automatic Control, 2019, 65(4): 1584-1597.

Partial-Decision Information Setup II

- **Pseudo-gradient mapping (PGM):**

$$F(\boldsymbol{x}) := \begin{pmatrix} \nabla_{x_1} J_1(x_1, \boldsymbol{x}_{-1}) \\ \vdots \\ \nabla_{x_N} J_N(x_N, \boldsymbol{x}_{-N}) \end{pmatrix}$$

- **Extended PGM (EPGM):**

$$\boldsymbol{F}(\boldsymbol{x}) := \begin{pmatrix} \nabla_{x_1} J_1(x_1, \boldsymbol{x}_1^{-1}) \\ \vdots \\ \nabla_{x_N} J_N(x_N, \boldsymbol{x}_N^{-N}) \end{pmatrix}$$

- **Property and assumption:**
  - $\boldsymbol{F}(\mathbf{1}_N \otimes \boldsymbol{x}) = F(\boldsymbol{x})$ (all estimates equal)
  - EPGM $\boldsymbol{F}(\boldsymbol{x})$ is Lipschitz continuous, not monotone

- **Contribution:**
  - Operator splitting + inertia\overrelaxation idea
  - Convergent rate & computation consumption
  - Mild assumption
  - Numerical simulation

Distributed Alternating Inertial v-GNE Seeking

- **Algorithm 1**: (P=projection operator, W=$[w_{ij}] \in R^{N\times N}$—weighted adjacency matrix)

$$
\begin{aligned}
x_{i,k+1} &= P_{\Omega_i}(\tilde{x}_{i,k} - \gamma_i(\nabla_{x_i} J_i(\tilde{x}_{i,k}, \tilde{\mathbf{x}}_{i,k}^{-i}) + A_i^T \tilde{\lambda}_{i,k} + c \sum_{j\in\mathcal{N}_i} w_{ij}(\tilde{x}_{i,k} - \tilde{\mathbf{x}}_{j,k}^{i}))) \\
\mathbf{x}_{i,k+1}^{-i} &= \tilde{\mathbf{x}}_{i,k}^{-i} - \gamma_i c \sum_{j\in\mathcal{N}_i} w_{ij}(\tilde{\mathbf{x}}_{i,k}^{-i} - \tilde{\mathbf{x}}_{j,k}^{-i}) \\
z_{i,k+1} &= \tilde{z}_{i,k} + \xi_i \sum_{j\in\mathcal{N}_i} w_{ij}(\tilde{\lambda}_{i,k} - \tilde{\lambda}_{j,k}) \\
\lambda_{i,k+1} &= P_{\mathbf{R}_+^m}(\tilde{\lambda}_{i,k} + \kappa_i(A_i(2x_{i,k+1} - \tilde{x}_{i,k}) - b_i \\
&\qquad - \sum_{j\in\mathcal{N}_i} w_{ij}(2(z_{i,k+1} - z_{j,k+1}) - (\tilde{z}_{i,k} - \tilde{z}_{j,k})) - \sum_{j\in\mathcal{N}_i} w_{ij}(\tilde{\lambda}_{i,k} - \tilde{\lambda}_{j,k})))
\end{aligned}
$$

$\alpha_k = 0$ *when* k *is even*, $\alpha_k = \alpha$ *when* k *is odd*.

$$
\begin{aligned}
\tilde{x}_{i,k} &= x_{i,k} + \alpha_k(x_{i,k} - x_{i,k-1}) \\
\tilde{\mathbf{x}}_{i,k}^{-i} &= \mathbf{x}_{i,k}^{-i} + \alpha_k(\mathbf{x}_{i,k}^{-i} - \mathbf{x}_{i,k-1}^{-i}) \\
\tilde{z}_{i,k} &= z_{i,k} + \alpha_k(z_{i,k} - z_{i,k-1}) \\
\tilde{\lambda}_{i,k} &= \lambda_{i,k} + \alpha_k(\lambda_{i,k} - \lambda_{i,k-1})
\end{aligned}
$$

1. Alternating inertial step
2. Local strategy x_i update (projected-pseudo-gradient)
3. Local estimation $\boldsymbol{x}_i^{-i}$ update
4. Local auxiliary variable z_i update
5. Dual variable λ_i update

[1] Pavel L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach[J]. IEEE Transactions on Automatic Control, 2019, 65(4): 1584-1597.
[2] Iutzeler F, Hendrickx J M. A generic online acceleration scheme for optimization algorithms via relaxation and inertia[J]. Optimization Methods and Software, 2019, 34(2): 383-405.

Distributed Alternating Overrelaxed v-GNE Seeking

- **Algorithm 2**：

$$
\begin{aligned}
\tilde{x}_{i,k} &= P_{\Omega_i}\Big(x_{i,k} - \tau_i\big(\nabla_{x_i} J_i(x_{i,k}, \mathbf{x}_{i,k}^{-i}) + A_i^\top \lambda_{i,k} + c \sum_{j \in \mathcal{N}_i} w_{ij}(x_{i,k} - \mathbf{x}_{j,k}^{i})\big)\Big) \\
\tilde{\mathbf{x}}_{i,k}^{-i} &= \mathbf{x}_{i,k}^{-i} - \tau_i c \sum_{j \in \mathcal{N}_i} w_{ij}(\mathbf{x}_{i,k}^{-i} - \mathbf{x}_{j,k}^{-i}) \\
\tilde{z}_{i,k} &= z_{i,k} + \nu_i \sum_{j \in \mathcal{N}_i} w_{ij}(\lambda_{i,k} - \lambda_{j,k}) \\
\tilde{\lambda}_{i,k} &= P_{\mathbf{R}_+^m}\Big(\lambda_{i,k} + \sigma_i\big(A_i(2\tilde{x}_{i,k} - x_{i,k}) - b_i \\
&\quad - \sum_{j \in \mathcal{N}_i} w_{ij}(2(\tilde{z}_{i,k} - \tilde{z}_{j,k}) - (z_{i,k} - z_{j,k})) - \sum_{j \in \mathcal{N}_i} w_{ij}(\lambda_{i,k} - \lambda_{j,k})\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\tau_k = 1 \textit{ when } k \textit{ is even}, \ \tau_k = \tau \textit{ when } k \textit{ is odd}. \\
&x_{i,k+1} = \tilde{x}_{i,k} + (\tau_k - 1)(\tilde{x}_{i,k} - x_{i,k}) \\
&\mathbf{x}_{i,k+1}^{-i} = \tilde{\mathbf{x}}_{i,k}^{-i} + (\tau_k - 1)(\tilde{\mathbf{x}}_{i,k}^{-i} - \mathbf{x}_{i,k}^{-i}) \\
&z_{i,k+1} = \tilde{z}_{i,k} + (\tau_k - 1)(\tilde{z}_{i,k} - z_{i,k}) \\
&\lambda_{i,k+1} = \tilde{\lambda}_{i,k} + (\tau_k - 1)(\tilde{\lambda}_{i,k} - \lambda_{i,k})
\end{aligned}
$$

1. Local strategy x_i update (projected-pseudo-gradient)
2. Local estimation $\boldsymbol{x}_i^{-i}$ update
3. Local auxiliary variable z_i update
4. Dual variable λ_i update
5. Alternating overrelaxed step

[1] Pavel L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach[J]. IEEE Transactions on Automatic Control, 2019, 65(4): 1584-1597.

[2] Iutzeler F, Hendrickx J M. A generic online acceleration scheme for optimization algorithms via relaxation and inertia[J]. Optimization Methods and Software, 2019, 34(2): 383-405.

Sketch of Convergence Analysis

$$\omega_k = \mathrm{col}\,(\boldsymbol{x}_k, \boldsymbol{z}_k, \boldsymbol{\lambda}_k)$$

- **Algorithm 1**:

$$\begin{cases} \omega_{k+1} = T(\omega_k), & \textit{if } k \textit{ is even} \\ \omega_{k+1} = T(\omega_k + \alpha(\omega_k - \omega_{k-1})), & \textit{if } k \textit{ is odd} \end{cases}$$

- **Algorithm 2**:

$$\begin{cases} \omega_{k+1} = T(\omega_k), & \textit{if } k \textit{ is even} \\ \omega_{k+1} = T(\omega_k) + (\tau - 1)\,(T(\omega_k) - \omega_k), & \textit{if } k \textit{ is odd} \end{cases}$$

$\boldsymbol{x}^* = \mathbf{1}_N \otimes \boldsymbol{x}^*$, $\boldsymbol{x}^*$ v-GNE

game theory | *operator theory*

$\mathrm{col}\,(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{\lambda}^*) \in zer\,(\mathfrak{A} + \mathfrak{B})$ $\longleftrightarrow$ $\mathrm{col}\,(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{\lambda}^*) \in \mathrm{fix}\,(T)$

fixed-point iteration

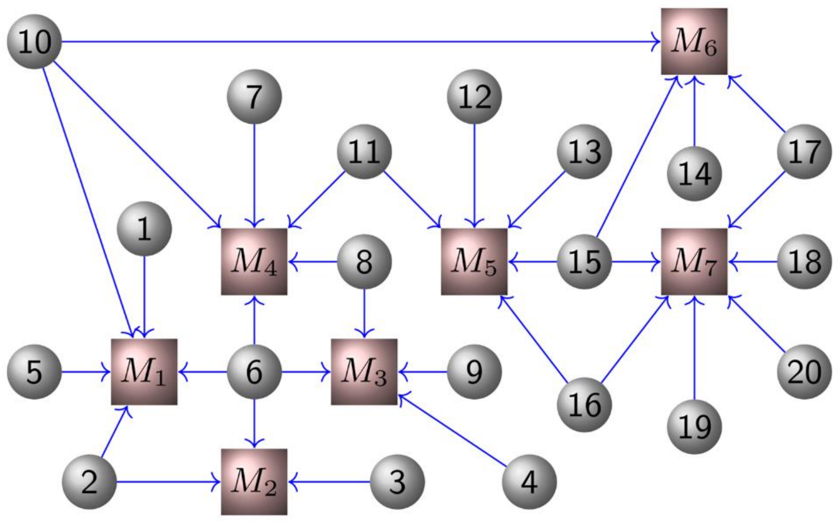
$\boldsymbol{x}_k \to \boldsymbol{x}^*$

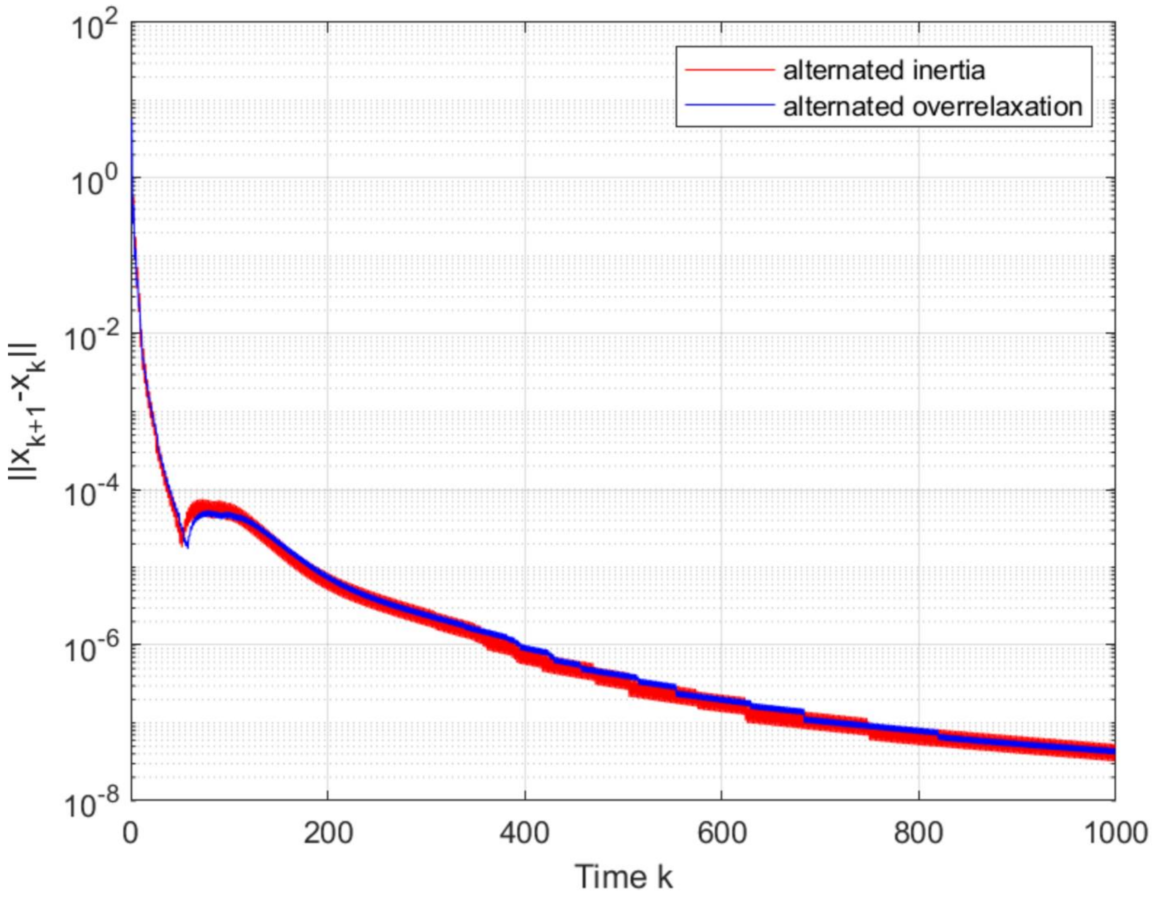
$$T = (Id - \Phi^{-1}\mathfrak{B}) \circ (Id + \Phi^{-1}\mathfrak{A})^{-1}$$

[1] Pavel L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach[J]. IEEE Transactions on Automatic Control, 2019, 65(4): 1584-1597.
[2] Bauschke H H, Combettes P L. Convex analysis and monotone operator theory in Hilbert spaces[M]. New York: Springer, 2011.

Numerical Simulation I

- **Nash-Gournot game** —— 20 firms participate in 7 markets’ competition





Numerical Simulation II

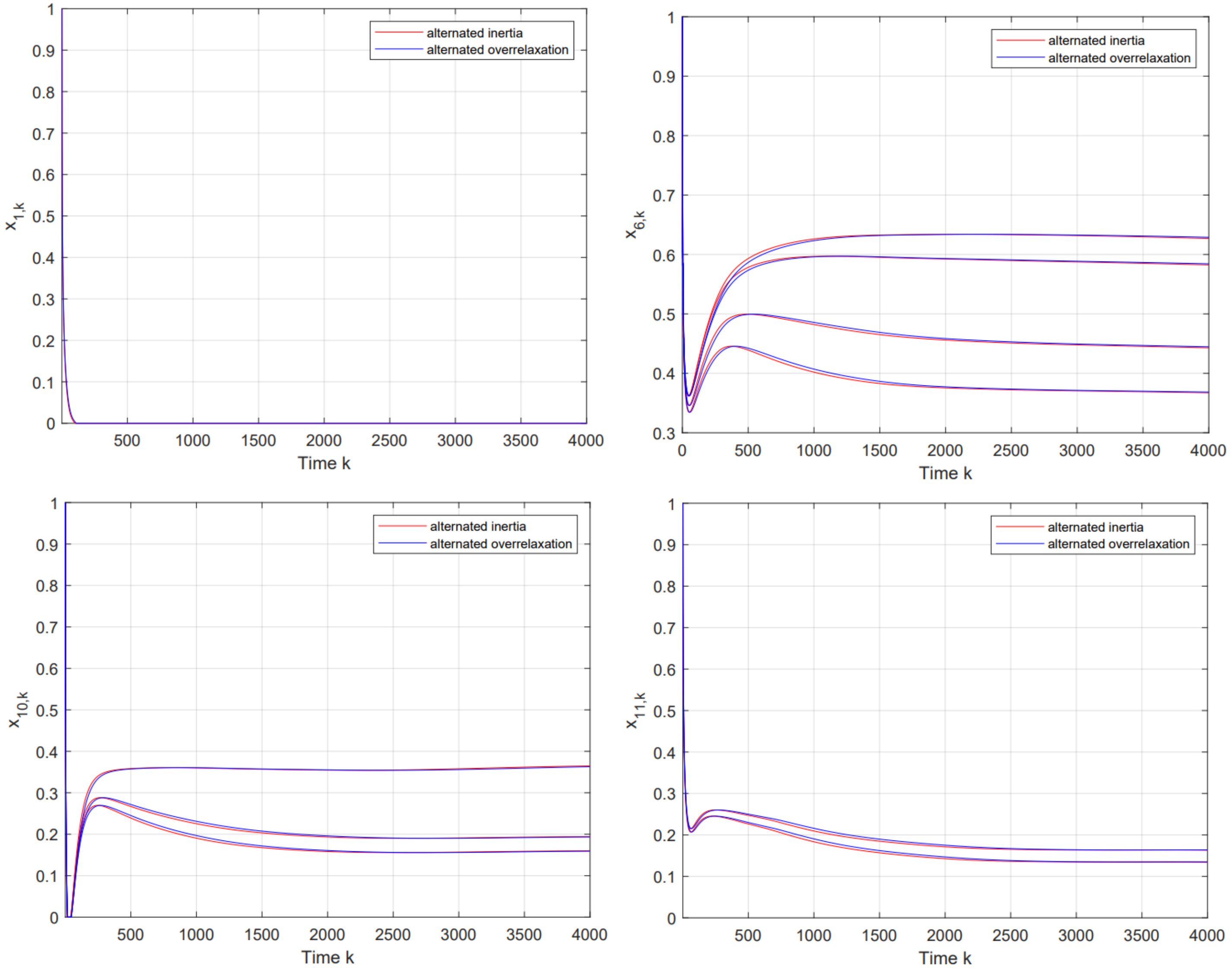


Figure 4. The trajectories of local decisions $x_{i,k}$ of firms 1, 6, 10 and 11 by Algorithms 1 and 2, respectively.

Numerical Simulation III

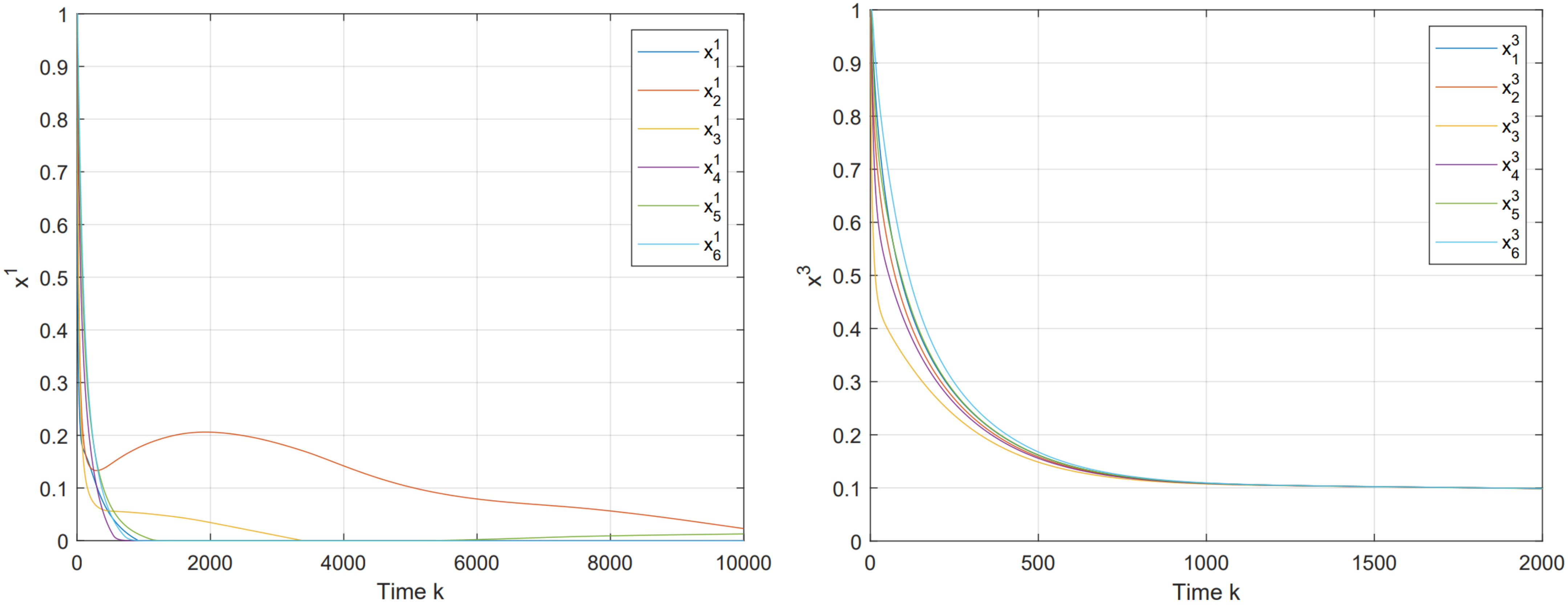


Figure 5. The trajectories of the estimate variable $\mathbf{x}_j^1$ from firms 1–6 generated by Algorithm 1 (left); and the trajectories of the estimate variable $\mathbf{x}_j^3$ from firm 1–6 generated by Algorithm 2 (right).

Numerical Simulation IV —— Comparison

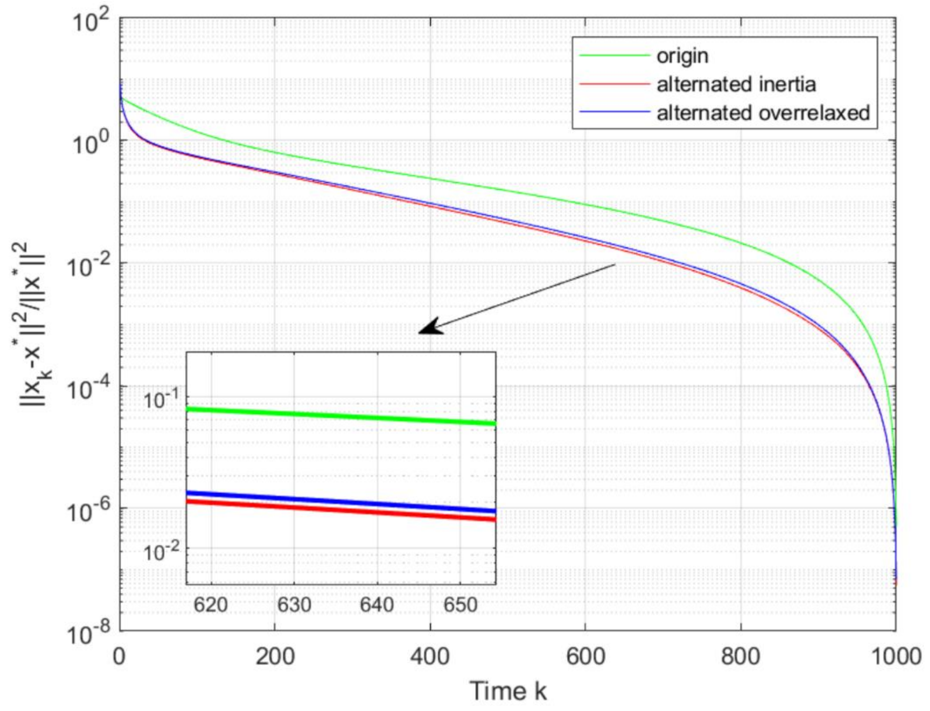


Figure 6. Relative error $\|x_k - x^*\|^2 / \|x^*\|^2$ generated by ([21], [Algorithm 1]), Algorithms 1 and 2.

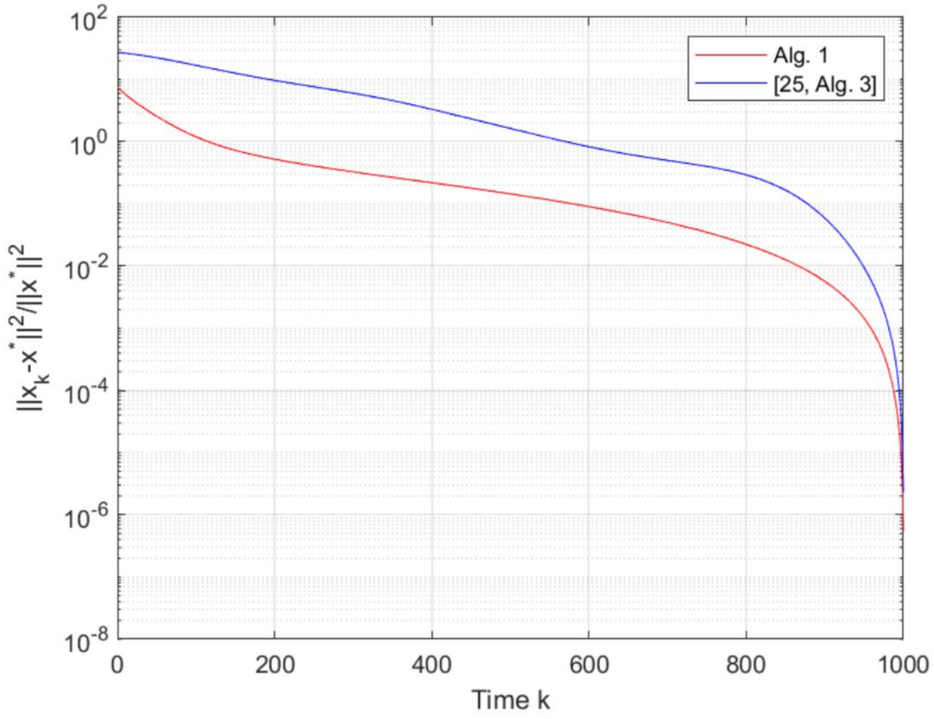


Figure 7. Relative error $\|x_k - x^*\|^2 / \|x^*\|^2$ generated by Algorithm 1 and ([25], [Alg. 3]).

OUTLINE

- Background and Motivation
- Main Results
- Conclusions and related works

Conclusion

- Distributed projected-gradient algorithms with alternating inertia/overrelaxed
  - ➢ guarantee the **convergence** to the v-GNE under **mild assumptions**
  - ➢ **fast convergence rate** and **low computation cost**
  - ➢ **faster** than proximal-point algorithm

- Operator splitting technique is key to analysis of convergence

Related Works

- Time-varying topology
  - **directed, disconnected** (Q-strongly-connected)
  - weighted matrix is **not doubly stochastic**

- Dynamic game
  - **time-varying** objective function
  - imperfect feedback —— **zero-order** algorithms

- Applications

Thank you for your attention!