Distributed generalized Nash equilibrium seeking in aggregative games on time-varying networks

Giuseppe Belgioioso Angelia Nedić Sergio Grammatico

Network Dynamics in the Social, Economic, and Financial Sciences, Torino, Italy.

November 7, 2019







Outline



Mathematical setup

Part I - GNE seeking under partial-decision information

Part II - Convergence analysis: Monotone operator/Fixed-point theory

Conclusion and outlook

Outline



Mathematical setup

Part I - GNE seeking under partial-decision information

 ${\sf Part\ II-Convergence\ analysis:\ Monotone\ operator/Fixed-point\ theory}$

Conclusion and outlook



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

```
orall i: \left\{egin{array}{ll} \operatorname{argmin} & J_i(x_i, \mathbf{x}_{-i}) & \longleftarrow & \operatorname{cost function} \ x_i \in \mathbb{R}^n \end{array}
ight.
```



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

```
orall i: egin{array}{lll} \operatorname{argmin} & J_i(x_i,x_{-i}) & \longleftarrow & \operatorname{cost function} \ & & \operatorname{s.t.} & x_i \in \Omega_i \subset \mathbb{R}^n & \longleftarrow & \operatorname{local constraint} \end{array}
```



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

$$orall i: egin{array}{lll} \operatorname{argmin} & J_i(x_i,x_{-i}) & \longleftarrow & \operatorname{cost function} \ & & & & \longleftarrow & \operatorname{local constraint} \ & & & & \longleftarrow & \operatorname{local constraint} \ & & & & \longleftarrow & \operatorname{coupling constraint} \end{array}$$



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

$$orall i: egin{array}{lll} \operatorname{argmin} & J_i(x_i,x_{-i}) & \longleftarrow & \operatorname{convex} & \operatorname{in} x_i \ & \operatorname{s.t.} & x_i \in \Omega_i \subset \mathbb{R}^n & \longleftarrow & \operatorname{local} & \operatorname{constraint} \ & x_i \in \mathcal{X}_i(x_{-i}) & \longleftarrow & \operatorname{coupling} & \operatorname{constraint} \end{array}$$



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

$$orall i: egin{array}{lll} \operatorname{argmin} & J_i(x_i,x_{-i}) & \longleftarrow & \operatorname{convex} & \operatorname{in} x_i \ & \operatorname{s.t.} & x_i \in \Omega_i \subset \mathbb{R}^n & \longleftarrow & \operatorname{closed, convex} \ & x_i \in \mathcal{X}_i(x_{-i}) & \longleftarrow & \operatorname{coupling constraint} \end{array}$$



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

$$\forall i: \begin{cases} \underset{x_i \in \mathbb{R}^n}{\operatorname{argmin}} & J_i(x_i, x_{-i}) & \longleftarrow \text{ convex in } x_i \\ \text{s.t.} & x_i \in \Omega_i \subset \mathbb{R}^n & \longleftarrow \text{ closed, convex} \\ & \sum_{j=1}^N A_j x_j - b_j \leq 0 & \longleftarrow \text{ affine coupling con.} \end{cases}$$



- $ightharpoonup N \gg 1$ agents/players, each with cost function and constraint set
- ► Game = { inter-dependent optimization problems }

$$\forall i: \begin{cases} \underset{x_i \in \mathbb{R}^n}{\operatorname{argmin}} & J_i(x_i, \frac{1}{N} \sum_{j=1}^N x_j) & \longleftarrow \text{ convex in } x_i \\ \text{s.t.} & x_i \in \Omega_i \subset \mathbb{R}^n & \longleftarrow \text{ closed, convex} \\ & \sum_{j=1}^N A_j x_j - b_j \leq 0 & \longleftarrow \text{ affine coupling con.} \end{cases}$$

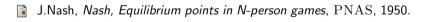
Generalized Nash equilibrium (GNE) is desired



A GNE is a feasible set of strategies $col(x_1^*, \dots, x_N^*)$ such that

$$\forall i: \quad J_i(x_i^*, \mathbf{x}_{-i}^*) \leq J_i(y, \mathbf{x}_{-i}^*), \quad \forall y \in \mathcal{X}(x_{-i}^*),$$

Everyone's choice is optimal given the choices of others.





John F. Nash Jr. Nobel Econ. 94





▶ via (primal-dual) projected pseudo-gradient dynamics

$$\forall i: \quad x_i^{k+1} = \operatorname{proj}_{\Omega_i} \left(x_i - \alpha_i \nabla_{x_i} J_i(x_i^k, x_{-i}^k) \right)$$

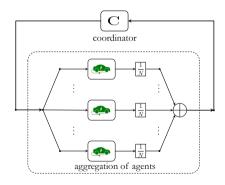
- Facchinei, Kanzow, Generalized Nash equilibrium problems, 4OR, 2007.
- [a] Kannan, Shanbhag, Distributed computation of equilibria in monotone Nash games via iterative regularization techniques SICON, 2012.
- Paccagnan, Gentile, Parise, Kamgarpour, Lygeros Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints, IEEE CDC, 2016.
- Belgioioso, Grammatico, Semi-decentralized Nash equilibrium seeking in aggregative games with separable coupling constraints, IEEE L-CSS, 2017.



 \blacktriangleright At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k$



▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$





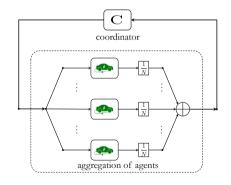
▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$

(S.1) Agents: local strategy update

$$x_i^{k+1} = \mathrm{proj}_{\Omega_i} \left[x_i^k - \gamma (\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k) \right]$$

$$\begin{split} \sigma^{k+1} &= \frac{1}{N} \sum_{j=1}^{N} x_j^{k+1} \\ \lambda^{k+1} &= \operatorname{proj}_{\mathbb{R}_{\geq 0}} \left[\lambda^k + \gamma (2Ax^{k+1} - Ax^k - b) \right] \end{split}$$

- Paccagnan et al., IEEE CDC, 2016.
- Belgioioso, Grammatico, ECC, 2018.





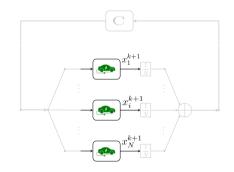
▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$

(S.1) Agents: local strategy update

$$x_i^{k+1} = \mathrm{proj}_{\Omega_i} \left[x_i^k - \gamma (\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k) \right]$$

$$\begin{split} \sigma^{k+1} &= \frac{1}{N} \sum_{j=1}^{N} x_j^{k+1} \\ \lambda^{k+1} &= \operatorname{proj}_{\mathbb{R}_{\geq 0}} \left[\lambda^k + \gamma (2Ax^{k+1} - Ax^k - b) \right] \end{split}$$

- Paccagnan et al., IEEE CDC, 2016.
- Belgioioso, Grammatico, ECC, 2018.





▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$

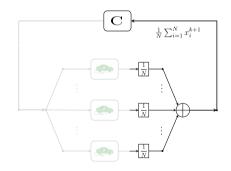
(S.1) Agents: local strategy update

$$x_i^{k+1} = \mathrm{proj}_{\Omega_i} \left[x_i^k - \gamma(\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k) \right]$$

$$\begin{split} \sigma^{k+1} &= \frac{1}{N} \sum_{j=1}^{N} x_j^{k+1} \\ \lambda^{k+1} &= \operatorname{proj}_{\mathbb{R}_{\geq 0}} \left[\lambda^k + \gamma (2A\boldsymbol{x}^{k+1} - A\boldsymbol{x}^k - b) \right] \end{split}$$









▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$

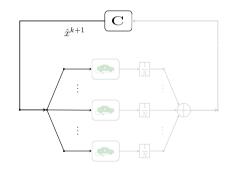
(S.1) Agents: local strategy update

$$x_i^{k+1} = \operatorname{proj}_{\Omega_i} \left[x_i^k - \gamma (\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k) \right]$$

$$\begin{split} \sigma^{k+1} &= \frac{1}{N} \sum_{j=1}^{N} x_j^{k+1} \\ \lambda^{k+1} &= \operatorname{proj}_{\mathbb{R}_{\geq 0}} \left[\lambda^k + \gamma (2Ax^{k+1} - Ax^k - b) \right] \end{split}$$









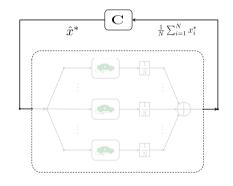
▶ At each stage k all the agents know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Central coordinator!}$

(S.1) Agents: local strategy update

$$x_i^{k+1} = \mathrm{proj}_{\Omega_i} \left[x_i^k - \gamma (\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k) \right]$$

$$\begin{split} \sigma^{k+1} &= \frac{1}{N} \sum_{j=1}^{N} x_j^{k+1} \\ \lambda^{k+1} &= \operatorname{proj}_{\mathbb{R}_{>0}} \left[\lambda^k + \gamma (2Ax^{k+1} - Ax^k - b) \right] \end{split}$$

- Paccagnan et al., IEEE CDC, 2016.
- Belgioioso, Grammatico, ECC, 2018.



Outline



Mathematical setup

Part I - GNE seeking under partial-decision information

Part II - Convergence analysis: Monotone operator/Fixed-point theory

Conclusion and outlook

Literature Overview: GNE seeking in partial-decision information



► Each agent is endowed with estimates of each other strategy

$$\forall i: \quad \hat{\pmb{x}}_i = \operatorname{col}(\hat{x}_{i,1}, \dots, \hat{x}_{i,N}), \quad \hat{x}_{i,i} = x_i,$$

via (primal-dual) projected pseudo-gradient dynamics + consensus dynamics

$$\forall i: \begin{cases} x_i^{k+1} = \operatorname{proj}_{\Omega_i} \left(x_i - \alpha_i \left(\nabla_{x_i} J_i(x_i^k, \hat{x}_{i,-i}^k) + \beta_i \sum_{j=1}^N w_{ij} (x_i^k - \hat{x}_{j,i}^k) \right) \\ \hat{x}_{i,-i}^{k+1} = x_{i,-i}^k - \beta_i \sum_{j=1}^N w_{ij} (\hat{x}_{i,-i}^k - \hat{x}_{j,-i}^k) \end{cases}$$

- T. Tatarenko, A. Nedić, IEEE CDC, 2018.
- L. Pavel, IEEE TAC, 2019.
- M. Bianchi, G. Belgioioso, S. Grammatico, ARXIV, 2019.

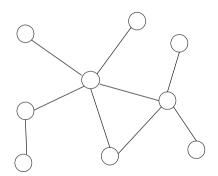


At each stage k the agents do not know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k$



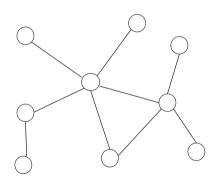


At each stage k the agents do not know $\sigma^k = \frac{1}{N} \sum_{i=1}^N x_i^k \to \text{Local estimates}$!



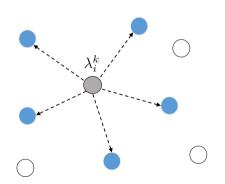
Undirected, Time-Varying,
 Q- repeatedly connected





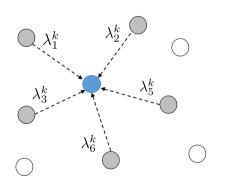
- ► Undirected, Time-Varying, Q— repeatedly connected
- ► Local variables of agent *i*:
 - (i) $x_i \leftarrow \text{decision}$
 - (ii) $\lambda_i \leftarrow \text{dual variable}$
 - (iii) $\sigma_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i x_i$
 - (iv) $z_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i \lambda_i$
 - (v) $y_i \leftarrow \text{est. of } \frac{1}{N} \sum_i A_i x_i b_i$





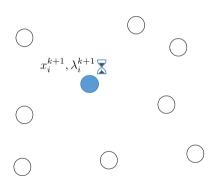
- ► Undirected, Time-Varying, Q— repeatedly connected
- ► Local variables of agent *i*:
 - (i) $x_i \leftarrow \text{decision}$
 - (ii) $\lambda_i \leftarrow \text{dual variable}$
 - (iii) $\sigma_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i x_i$
 - (iv) $z_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i \lambda_i$
 - (v) $y_i \leftarrow \text{est. of } \frac{1}{N} \sum_i A_i x_i b_i$





- Undirected, Time-Varying,Q- repeatedly connected
- ► Local variables of agent *i*:
 - (i) $x_i \leftarrow \text{decision}$
 - (ii) $\lambda_i \leftarrow \text{dual variable}$
 - (iii) $\sigma_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i x_i$
 - (iv) $z_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i \lambda_i$
 - (v) $y_i \leftarrow \text{est. of } \frac{1}{N} \sum_i A_i x_i b_i$





- Undirected, Time-Varying,Q- repeatedly connected
- ightharpoonup Local variables of agent i:
 - (i) $x_i \leftarrow \text{decision}$
 - (ii) $\lambda_i \leftarrow \text{dual variable}$
 - (iii) $\sigma_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i x_i$
 - (iv) $z_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i \lambda_i$
 - (v) $y_i \leftarrow \text{est. of } \frac{1}{N} \sum_i A_i x_i b_i$



- Undirected, Time-Varying,Q- repeatedly connected
- ► Local variables of agent *i*:
 - (i) $x_i \leftarrow \text{decision}$
 - (ii) $\lambda_i \leftarrow \text{dual variable}$
 - (iii) $\sigma_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i x_i$
 - (iv) $z_i \leftarrow \text{estimate of } \frac{1}{N} \sum_i \lambda_i$
 - (v) $y_i \leftarrow \text{est. of } \frac{1}{N} \sum_i A_i x_i b_i$
- ► Iteration *k* ...



(S.1) Communication over network

$$\hat{\sigma}_{i}^{k} = \sum_{j=1}^{N} w_{ij} \sigma_{j}^{k}, \quad \hat{z}_{i}^{k} = \sum_{j=1}^{N} w_{ij} z_{j}^{k}, \quad \hat{y}_{i}^{k} = \sum_{j=1}^{N} w_{ij} y_{j}^{k},$$

(S.2) Local Primal-Dual update

$$\begin{split} \tilde{x}_i^k &= \operatorname{Proj}_{\Omega_i} \left[x_i^k - \alpha_i \left(\nabla_{x_i} J_i(x_i^k, \hat{\sigma}_i^k) + A_i^{\top} \hat{z}_i^k \right) \right], \\ \tilde{\lambda}_i^k &= \operatorname{proj}_{\mathbb{R}_+} \left(\lambda_i^k + \beta_i (y_i^k - \lambda_i^k + \hat{z}_i^k) \right), \end{split}$$

(S.3) Local Krasnosel'skii-Mann Process

$$x_i^{k+1} = x_i^k + \gamma^k (\tilde{x}_i^k - x_i^k),$$

$$\lambda_i^{k+1} = \lambda_i^k + \gamma^k (\tilde{\lambda}_i^k - \lambda_i^k),$$

(S.4) Local Dynamic Tracking

$$\begin{split} \sigma_i^{k+1} &= \hat{\sigma}_i^k + x_i^{k+1} - x_i^k, \\ z_i^{k+1} &= \hat{z}_i^k + \lambda_i^{k+1} - \lambda_i^k. \\ y_i^{k+1} &= \hat{y}_i^k + C_i(2\tilde{x}_i^k - x_i^k) - C_i(2\tilde{x}_i^{k-1} - x_i^{k-1}). \end{split}$$



Algorithm. Local updating rules:

- 1. Communication and distributed averaging
- 2. Inexact projected-pseudogradient dynamics
- 3. Krasnosels'kii-Mann process
- 4. Dynamic Tracking of the aggregation quantities

Outline



Mathematical setup

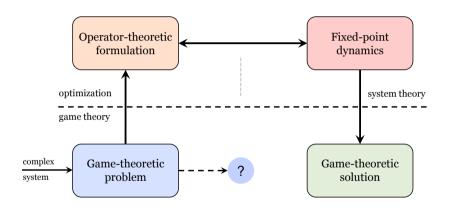
Part I - GNE seeking under partial-decision information

Part II - Convergence analysis: Monotone operator/Fixed-point theory

Conclusion and outlook

Operator theoretic approach: schematic outline

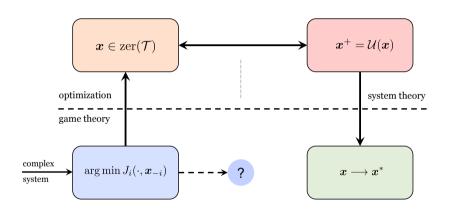






Operator theoretic approach: schematic outline







Operator theoretic approach - partial information



How to estimate on-line aggregative quantities?

Operator theoretic approach - partial information



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques



Nedić et al., ${\rm SIAM}$, 2017. Xu, ${\rm CDC}$, 2015. Zanella, ${\rm CDC}$, 2011. Xi, ${\rm TAC}$, 2018.



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques

Nedić et al., SIAM, 2017. Xu, CDC, 2015. Zanella, CDC, 2011. Xi, TAC, 2018.

Distributed Optimization. Goal: $\min \sum_{i=1}^{N} f_i(x)$ over a network



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques

Nedić et al., SIAM, 2017. Xu, CDC, 2015. Zanella, CDC, 2011. Xi, TAC, 2018.

Distributed Optimization. Goal: min $\sum_{i=1}^{N} f_i(x)$ over a network

Distributed Gradient Descent
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \nabla f_i(x_i^k)$$



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques

Nedić et al., SIAM, 2017, Xu, CDC, 2015, Zanella, CDC, 2011, Xi, TAC, 2018.

Distributed Optimization. Goal: min $\sum_{i=1}^{N} f_i(x)$ over a network

Distributed Gradient Descent
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \nabla f_i(x_i^k)$$

Ideal (non-distributed) scheme
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \frac{1}{N} \sum_{j=1}^{N} \nabla f_i(x_i^k)$$



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques

Nedić et al., SIAM, 2017, Xu, CDC, 2015, Zanella, CDC, 2011, Xi, TAC, 2018.

Distributed Optimization. Goal: min $\sum_{i=1}^{N} f_i(x)$ over a network

Distributed Gradient Descent $x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \nabla f_i(x_i^k)$

Ideal (non-distributed) scheme
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \frac{1}{N} \sum_{j=1}^{N} \nabla f_i(x_i^k)$$

DIGing (Nedić et al., SIAM, 2017)
$$x_{i}^{k+1} = \sum_{j=1}^{N} w_{ij} x_{i}^{k} - \alpha_{i} y_{i}(k),$$

$$y_{i}^{k+1} = \sum_{j=1}^{N} w_{ij} y_{i}^{k} + \nabla f_{i}(x_{i}^{k+1}) - \nabla f_{i}(x_{i}^{k}),$$



How to estimate on-line aggregative quantities?

Idea: combine splitting methods with gradient tracking techniques

Nedić et al., SIAM, 2017, Xu, CDC, 2015, Zanella, CDC, 2011, Xi, TAC, 2018.

Distributed Optimization. Goal: min $\sum_{i=1}^{N} f_i(x)$ over a network

Distributed Gradient Descent
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i \nabla f_i(x_i^k)$$

Ideal (non-distributed) scheme
$$x_i^{k+1} = \sum_{j=1}^N w_{ij} x_i^k - \alpha_i \frac{1}{N} \sum_{j=1}^N \nabla f_i(x_i^k)$$

DIGing (Nedić et al., SIAM, 2017)
$$x_i^{k+1} = \sum_{j=1}^{N} w_{ij} x_i^k - \alpha_i y_i(k),$$

$$y^{\infty} = \frac{1}{N} \sum_i \nabla f_i(x^{\infty})$$



How to deal with the inexactness of the estimates?



How to deal with the inexactness of the estimates?

Krasnosel'skii-Mann fixed-point iteration with errors

$$egin{aligned} oldsymbol{x}^{k+1} &= oldsymbol{x}^k + \gamma^k (\mathcal{T}(oldsymbol{x}^k) - oldsymbol{x}^k + e^k) \ e^k &= \mathcal{T}(\hat{oldsymbol{x}}^k) - \mathcal{T}(oldsymbol{x}^k) \end{aligned}$$

P. Combettes, Quasi-Fejerian Analysis of some optimization algorithms, ELSEVIER, 2001



How to deal with the inexactness of the estimates?

Krasnosel'skii-Mann fixed-point iteration with errors

$$egin{aligned} oldsymbol{x}^{k+1} &= oldsymbol{x}^k + \gamma^k (\mathcal{T}(oldsymbol{x}^k) - oldsymbol{x}^k + e^k) \ e^k &= \mathcal{T}(\hat{oldsymbol{x}}^k) - \mathcal{T}(oldsymbol{x}^k) \end{aligned}$$

P. Combettes, Quasi-Fejerian Analysis of some optimization algorithms, ELSEVIER, 2001

Theorem. If $\sum_{k=0}^{\infty} \gamma^k ||e^k|| < \infty$, T non-expansive, then $\mathbf{x}^k \to \mathbf{x}^* \in \mathrm{fix}(T)$.

Outline



Mathematical setup

Part I - GNE seeking under partial-decision information

 ${\sf Part\ II-Convergence\ analysis:\ Monotone\ operator/Fixed-point\ theory}$

Conclusion and outlook

Conclusion and outlook



- ▶ Problem: GNE seeking in aggregative games under partial decision-information
- ► Framework: Monotone operator/fixed-point theory
- ► Our result: Full-distributed algorithm on time-varying networks

G. Belgioioso, A. Nedić, S. Grammatico, *Distributed generalized Nash equilibrium seeking in aggregative games on time-varying networks, arXiv* (available at https://arxiv.org/abs/1907.00191), 2019.

Conclusion and outlook



- ► Same Framework: Monotone operator/fixed-point theory
- ▶ Different Viewpoint: The network dynamics is fixed

$$m{x}(k+1) = egin{bmatrix} ext{prox}_{f_1} & & & \\ & & \ddots & \\ & & & ext{prox}_{f_N} \end{bmatrix} W(k) m{x}(k)$$

ightharpoonup Problem: Under what conditions on W and f_i 's does the system converge?

C. Cenedese, G. Belgioioso, Y.Kawano, S. Grammatico, M. Cao, Asynchronous and time-varying proximal type dynamics multi-agent network games, arXiv (available at https://arxiv.org/abs/1909.11203), 2019.

Thank you for your attention!

