A Unified Distributed Algorithm for *Non*-Games
Non-cooperative, Non-convex, and Non-differentiable

Jong-Shi Pang[*] and **Meisam Razaviyayn**[†]

presented at

Workshop on Optimization for Modern Computation

Peking University, Beijing, China

10:10 – 10:45 PM, Thursday September 04, 2014

[*]Department of Industrial and Systems Engineering, University of Southern California, Los Angeles

[†]Visiting Research Assistant, Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis

The Non-cooperative Game $\mathcal{G}$

• An n-player non-cooperative game $\mathcal{G}$ wherein each player $i = 1, \cdots, n$, anticipating the rivals' strategy tuple $x^{-i} \triangleq \left(x^j\right)_{i \neq j=1}^n \in \mathcal{X}^{-i} \triangleq \prod_{i \neq j=1}^n \mathcal{X}^j$, solves the optimization problem:

$$\underset{x^i \in \mathcal{X}^i}{\text{minimize}} \; \theta_i(x^i, x^{-i})$$

• $\mathcal{X}^i \subseteq \mathbb{R}^{n_i}$ is a closed convex set

• $\theta_i : \mathbf{\Omega} \to \mathbb{R}$ is a locally Lipschitz continuous and directionally differentiable function defined on $\mathbf{\Omega} \triangleq \prod_{i=1}^n \Omega^i$ where each Ω^i is an open convex set containing $\mathcal{X}^i$

• **A key structural assumption for convergence of distributed algorithm:** each $\theta_i(x) = f_i(x) + g_i(x^i)$, with $f_i(x)$, dependent on *all* players' strategy profile $x \triangleq \left(x^i\right)_{i=1}^n$, being twice continuously differentiable but not necessarily convex, and $g_i(x^i)$, dependent on player i's strategy profile x^i only is convex but not necessarily differentiable.

Quasi-Nash equilibrium: Definition and existence

Definition. A player profile $x^* \triangleq \left(x^{*,i}\right)_{i=1}^n$ is a **QNE** if for every $i = 1, \cdots, n$:

$$\theta_i(\bullet, x^{*,-i})'(x^{*,i}; x^i - x^{*,i}) \geq 0, \quad \forall\, x^i \in \mathcal{X}^i.$$

Existence. Suppose each $\theta_i(x) = f_i(x) + g_i(x)$ with $\nabla_{x^i} f_i$ continuously differentiable on Ω and $g_i(\bullet, x^{-i})$ convex on $\mathcal{X}^i$ that is compact and convex.

Proof by a fixed-point argument applied to the map:

$\Phi : x \triangleq \left(x^i\right)_{i=1}^n \in \boldsymbol{\mathcal{X}} \triangleq \prod_{i=1}^n \mathcal{X}^i \mapsto \Phi(x) \triangleq (\Phi_i(x))_{i=1}^n \in \boldsymbol{\mathcal{X}}$, where, for $i = 1, \cdots, n$,

$$\Phi_i(x) \triangleq \underset{z^i \in \mathcal{X}^i}{\operatorname{argmin}} \left[f_i(z^i, x^{-i}) + g_i(z^i, x^{-i}) + \frac{\alpha}{2} \| z^i - x^i \|^2 \right],$$

with $\alpha > 0$ such that the minimand is strongly convex in z^i for fixed x^{-i}.

Remark. For existence, $g_i(x)$ can be fully dependent on the player profile x; but for convergence of distributed algorithm, $g_i(x^i)$ is only player dependent.

The unified algorithm

The main idea:

• Employing player-convex surrogate objective functions and the information from the most current iterate, non-overlapping groups of players update, in parallel, their strategies from the solution of sub-games.

• Thus the algorithm is a mixture of the classical block Gauss-Seidel and Jacobi iterations, applied in a way consistent with the game-theoretic setting of the problem.

Two key families:

• The player groups: $\boldsymbol{\sigma}^{\nu} \triangleq \left\{ \sigma_1^{\nu}, \cdots, \sigma_{\kappa_\nu}^{\nu} \right\}$ consists of κ_ν pairwise disjoint subsets of the players' labels, for some integer $\kappa_\nu > 0$. Players in each group σ_k^{ν} solve a sub-game; all such sub-games in iteration ν are solved in parallel.

$$\boldsymbol{\mathcal{N}}_{\nu} \triangleq \bigcup_{k=1}^{\kappa_\nu} \sigma_k^{\nu} \text{ not necessarily equal to } \{1, \cdots, n\};$$

i.e., some players may not update in an iteration.

• Given $x^\nu \in \boldsymbol{\mathcal{X}}$, the bivariate surrogate objectives: $\left\{ \widehat{\theta}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; x^\nu) \, : \, i \in \sigma_k^\nu \right\}_{k=1}^{\kappa_\nu}$ in lieu of the original objectives $\left\{ \{\theta_i\}_{i \in \sigma_k^\nu} \right\}_{k=1}^{\kappa_\nu}$.

• **The subgames**, denoted $\mathcal{G}_\nu^{\sigma_k^\nu}$ for $k = 1, \cdots \kappa_\nu$: the optimization problems of the players in σ_k^ν are

$$\left\{ \underset{x^i \in \mathcal{X}^i}{\text{minimize}} \; \widehat{\theta}_i^{\sigma_k^\nu} \left(\underbrace{x^i, x^{\sigma_k^\nu; -i}}_{\substack{\text{subgame variables} \\ x^{\sigma_k^\nu}}} \; ; \; \underbrace{x^\nu}_{\substack{\text{input to subgame} \\ \text{at iteration } \nu}} \right) \right\}_{i \in \sigma_k^\nu} .$$

• **The new iterate** for a step size $\tau_{\sigma_k^\nu} \in (0, 1]$

$$x^{\nu+1;\sigma_k^\nu} \triangleq x^{\nu;\sigma_k^\nu} + \tau_{\sigma_k^\nu} \left(\underbrace{\widehat{x}^{\,\nu;\sigma_k^\nu}}_{\text{solution to subgame}} - x^{\nu;\sigma_k^\nu} \right).$$

• Need **directional derivative consistency** at limit x^∞ of generated sequence:

$$\theta_i(\bullet, x^{\infty,-i})'(x^{\infty,i}; x^i - x^{\infty,i}) \geq \widehat{\theta}_i^{\sigma_k^t}(\bullet, x^{\infty,\sigma_k^t;-i}; x^\infty)'(x^{\infty,i}; x^i - x^{\infty,i}), \quad \forall \, x^i \in \mathcal{X}^i.$$

An illustration. A 10-player game with the grouping:

$$\boldsymbol{\sigma}^\nu = \{\{1,2\},\{3,4,5\},\{6,7,8,9\}\}$$

so that $\kappa_\nu = 3$ and $\boldsymbol{\mathcal{N}}_\nu = \{1,\cdots,9\}$, leaving out the 10th-player.

Players 1 and 2 update their strategies by solving a subgame $\mathcal{G}_\nu^{\{1,2\}}$ defined by the surrogate objective functions $\widehat{\theta}_1^{\{1,2\}}(\bullet; x^\nu)$ and $\widehat{\theta}_2^{\{1,2\}}(\bullet; x^\nu)$.

In parallel, players 3, 4, and 5 update their strategies by solving a subgame $\mathcal{G}_\nu^{\{3,4,5\}}$ using the surrogate objective functions

$$\left\{ \widehat{\theta}_3^{\{3,4,5\}}(\bullet; x^\nu),\, \widehat{\theta}_4^{\{3,4,5\}}(\bullet; x^\nu),\, \widehat{\theta}_5^{\{3,4,5\}}(\bullet; x^\nu) \right\};$$

similarly for players 6 through 9.

The 10th player is not performing an update in the current iteration ν according to the given grouping.

Special cases: player groups

• Block Jacobi $\mathcal{N}_\nu = \{1, \cdots, n\}$ and σ_k^ν may contains multiple elements.

• Point Jacobi $\kappa_\nu = n$; thus $\sigma_k^\nu = \{k\}$ for $k = 1, \cdots n$: each player i solves an optimization problem:

$$\underset{x^i \in \mathcal{X}^i}{\text{minimize}} \ \widehat{\theta}_i(x^i; x^\nu).$$

• Block Gauss-Seidel $\kappa_\nu = 1$ for all ν: only the players in the block σ_1^ν update their strategies that immediately become the inputs to the new iterate $x^{\nu+1}$ while all other players $j \notin \sigma_1^\nu$ keep their strategies at the current iterate $x^{\nu,j}$.

• Point Gauss-Seidel $\kappa_\nu = 1$ and σ_1^ν is a singleton.

• Above are deterministic player groups; also consider randomized player groups: Let $\{\sigma_1, \cdots \sigma_K\}$ be a partition of $\{1, \cdots, n\}$. At iteration ν, the subset $\boldsymbol{\sigma}^\nu \subseteq \{\sigma_1, \ldots, \sigma_K\}$ of player groups is chosen randomly and independently from the previous iterations, so that

$$\Pr(\sigma_i \in \boldsymbol{\sigma}^\nu) = p_{\sigma_i} > 0,$$

There is a positive probability p_{σ_i}, same at all iterations ν, for the subset σ_i of players to be chosen to update their strategies.

Special cases: surrogate objectives

- Standard convex case Suppose $\theta_i(\bullet, x^{-i})$ is convex. For $i \in \sigma_k^\nu$, let

$$\widehat{\theta}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; z) \triangleq \theta_i(x^{\sigma_k^\nu}, z^{-\sigma_k^\nu}) + \underbrace{\frac{\alpha_i}{2} \| x^i - z^i \|^2}_{\text{regularization}}, \quad \text{for some positive scalar } \alpha_i$$

- Mixed convexity and differentiability Suppose $\theta_i(\bullet, x^{-i}) = g_i(\bullet, x^{-i}) + f_i(\bullet, x^{-i})$, where $g_i(\bullet, x^{-i})$ is convex and $f_i(\bullet, x^{-i})$ is differentiable. Let

$$\widehat{\theta}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; z) \triangleq \underbrace{g_i(x^{\sigma_k^\nu}, z^{-\sigma_k^\nu}) + \underbrace{f_i(z) + \sum_{j \in \sigma_k^\nu} \nabla_{z^j} f_i(z)^T (x^j - z^j)}_{\text{partial linearization}} + \frac{\alpha_i}{2} \| x^i - z^i \|^2}_{\text{convex in } x^{\sigma_k^\nu} \text{ for fixed } z}$$

- Newton-type quadratic approximation Suppose $\nabla_{x^i} \theta_i(\bullet, x^{-i})$ exists. Let

$$\widehat{\theta}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; z) \triangleq \underbrace{\theta_i(z) + \sum_{j \in \sigma_k^\nu} \nabla_{x^j} \theta_i(z)^T (x^j - z^j) + \tfrac{1}{2} \sum_{j,j' \in \sigma_k^\nu} (x^{j'} - z^{j'})^T B^{\sigma_k^\nu; j,j'} (x^j - z^j)}_{\text{quadratic in } x^{\sigma_k^\nu} \text{ for fixed } z},$$

$B^{\sigma_k^\nu; j,j'}$ approximates mixed partial derivatives of $\theta_i(\bullet, z^{-\sigma_k^\nu})$ w.r.t. x^j and $x^{j'}$.

Convergence analysis

Two approaches

• Contraction — showing that the sequence $\{x^\nu\}_{\nu=1}^\infty$ contracts in the vector sense by means of the assumption of a spectral radius condition of a key matrix

• Potential — relying on the existence of a potential function that decreases at each iteration.

Think about a system of linear equations: $Ax = b$

• (Generalized) diagonal dominance yields convergence under contraction.

• Symmetry of A yields the potential function: $P(x) \triangleq \frac{1}{2}x^T Ax - b^T x$.

Contraction approach

• An integer $T > 0$ and a fixed family $\left\{\widehat{\boldsymbol{\sigma}}^{\,t} \triangleq \left\{\sigma_1^t, \cdots, \sigma_{\kappa_t}^t\right\}\right\}_{t=1}^T$ of index subsets of the players' labels that partitions $\{1, \cdots, n\}$.

• Families of bivariate surrogate functions

$$\widehat{\boldsymbol{\theta}}^{\,t} = \left\{ \widehat{\theta}_i^{\sigma_k^t} : i \in \sigma_k^t \right\}_{k=1}^{\kappa_t}, \quad \text{for } t = 1, \cdots, T,$$

such that for every pair $(x^{\sigma_k^t;-i}; z)$, the function $\widehat{\theta}_i^{\sigma_k^t}(\bullet, x^{\sigma_k^t;-i}; z)$ is convex.

• For each set σ_k^t, let $\mathcal{G}_t^{\sigma_k^t}$ denote the subgame consisting of the players $i \in \sigma_k^t$ with objective functions $\widehat{\theta}_i^{\sigma_k^t}(\bullet; z)$ for certain (known) iterate z to be specified.

• Let $\kappa_\nu = \kappa_t$ and $\sigma_k^\nu = \sigma_k^t$ for $\nu \equiv t$ modulo T and for all $k = 1, \cdots, \kappa_\nu$; thus, for each $i = 1, \cdots, n$, $\widehat{\theta}_i^{\sigma_k^\nu} = \widehat{\theta}_i^{\sigma_k^t}$ where $\nu \equiv t$ modulo T and σ_k^t is the unique index set containing i.

• Thus, each player i and the members in σ_k^t will update their strategy tuple exactly once every T iterations through the solution of the subgame $\mathcal{G}_t^{\sigma_k^t}$.

• Finally, we take each step size $\tau_{\sigma_k^\nu} = 1$.

A further illustration

Consider a 12-player game with $T = 3$ and with $\widehat{\boldsymbol{\sigma}}^{\,1} = \{\{1,2\},\{3,5,6\}\}$, $\widehat{\boldsymbol{\sigma}}^{\,2} = \{\{4,7,8\}\}$, and $\widehat{\boldsymbol{\sigma}}^{\,3} = \{\{9\},\{10,11,12\}\}$.

Starting with $x^0 = \left(x^{0;i}\right)_{i=1}^{12} = \mathbf{x}^{(0)}$, we obtain after one iteration

$$x^1 = \left(x^{1;\{1,2\}}, x^{1;\{3,5,6\}}, x^{0;4}, x^{0;\{7\text{thru}12\}} \right).$$

The sub-vectors $x^{1;\{1,2\}}$ and $x^{1;\{3,5,6\}}$ mean that the players 1 and 2 update their strategies by solving a 2-player subgame and simultaneously the players 3, 5, and 6 update their strategies by solving a 3-player subgame. The remaining players 4, 7 through 12 do not update their strategy in this first iteration.

The next two iterations yield, respectively,

$$\begin{aligned} x^2 &= \left(x^{1;\{1,2\}}, x^{1;\{3,5,6\}}, x^{2;\{4,7,8\}}, x^{0;\{9\text{thru}12\}} \right) \\ x^3 &= \left(x^{1;\{1,2\}}, x^{1;\{3,5,6\}}, x^{2;\{4,7,8\}}, x^{3;\{9\}}, x^{3;\{10,11,12\}} \right). \end{aligned}$$

The update of x^2 employs x^1 in defining the player objectives $\widehat{\theta}_4^{\{4,7,8\}}(\bullet; x^1)$, $\widehat{\theta}_7^{\{4,7,8\}}(\bullet; x^1)$, and $\widehat{\theta}_8^{\{4,7,8\}}(\bullet; x^1)$. Similarly, the update of x^3 employs x^2.

After three iterations, we have completed a full cycle where all players have updated their strategies exactly once, obtaining the new iterate $\mathbf{x}^{(1)} = x^3$.

The next cycle of updates is then initiated according to the same partition $\left\{\widehat{\boldsymbol{\sigma}}^1, \widehat{\boldsymbol{\sigma}}^2, \widehat{\boldsymbol{\sigma}}^3\right\}$ and employs the same family of bivariate surrogate functions.

- **group 1:** $\underbrace{\underbrace{\widehat{\theta}_1^{\{1,2\}}, \widehat{\theta}_2^{\{1,2\}}}_{\text{2-person subgame}}, \underbrace{\widehat{\theta}_3^{\{3,5,6\}}, \widehat{\theta}_5^{\{3,5,6\}}, \widehat{\theta}_6^{\{3,5,6\}}}_{\text{3-person subgame}}}_{\text{2 subgames solve in parallel}};$ **parallel**

- **group 2:** $\underbrace{\widehat{\theta}_4^{\{4,7,8\}}, \widehat{\theta}_7^{\{4,7,8\}}, \widehat{\theta}_8^{\{4,7,8\}}}_{\text{3-person subgame}};$ **single game**

- **group 3:** $\underbrace{\underbrace{\widehat{\theta}_9^{\{9\}}}_{\text{single-player opt}}, \underbrace{\widehat{\theta}_{10}^{\{10,11,12\}}, \widehat{\theta}_{11}^{\{10,11,12\}}, \widehat{\theta}_{12}^{\{10,11,12\}}}_{\text{3-person subgame}}}_{\text{2 subgames solved in parallel}}$ **parallel: single opt. + game**

group 1 $\xrightarrow{\text{sequential}}$ **group 2:** $\xrightarrow{\text{sequential}}$ **group 3**

Set-up for assumptions

• Assume $\widehat{\theta}_i^{\sigma_k^t}(x^{\sigma_k^t}; z) = g_i(x^i) + \widehat{f}_i^{\sigma_k^t}(x^{\sigma_k^t}; z)$, where g_i is convex and the (surrogate objective) $\widehat{f}_i^{\sigma_k^t}(\bullet; z)$ is twice continuously differentiable.

• $\widehat{f}_i^{\sigma_k^t}(x^{\sigma_k^t}; z)$ is strongly convex in $x^{\sigma_k^t}$ uniformly in z; i.e., $\exists\ \gamma_{k;ii}^t > 0$ such that for all $x^i \in \mathcal{X}^i$, all $u^{\sigma_k^t} \in \mathcal{X}^{\sigma_k^t}$ and all $z \in \boldsymbol{\mathcal{X}}$,

$$\left(x^i - u^i \right)^T \nabla^2_{u^i u^i} \widehat{f}_i^{\sigma_k^t}(u^{\sigma_k^t}; z) \left(x^i - u^i \right) \geq \gamma_{k;ii}^t \, \| \, x^i - u^i \, \|^2.$$

• Further assume that each function $\nabla_{u^i} \widehat{f}_i^{\sigma_k^t}$ is continuously differentiable in both arguments with bounded derivatives. Let

$$\gamma_{k;ij}^t \triangleq \sup_{u^{\sigma_k^t} \in \mathcal{X}^{\sigma_k^t};\, z \in \boldsymbol{\mathcal{X}}} \left\| \nabla^2_{u^j u^i} \widehat{f}_i^{\sigma_k^t}(u^{\sigma_k^t}; z) \right\| < \infty, \quad \forall\, i \neq j \text{ in } \sigma_k^t$$

$$\widetilde{\gamma}_{k;i\ell}^t \triangleq \sup_{u^{\sigma_k^t} \in \mathcal{X}^{\sigma_k^t};\, z \in \boldsymbol{\mathcal{X}}} \left\| \nabla^2_{z^\ell u^i} \widehat{f}_i^{\sigma_k^t}(u^{\sigma_k^t}; z) \right\| < \infty, \quad \forall\, i \in \sigma_k^t \text{ and } \ell = 1, \cdots, n.$$

Let $\boldsymbol{\Gamma} \triangleq \text{blkdiag}\left[\Gamma^t \right]_{t=1}^T$, where each $\Gamma^t \triangleq \text{blkdiag}\left[\Gamma_k^t \right]_{k=1}^{\kappa_t}$ and $\Gamma_k^t \triangleq \left[\gamma_{k;ij}^t \right]_{i,j \in \sigma_k^t}$.

Let $\widetilde{\boldsymbol{\Gamma}} \triangleq \left[\widetilde{\Gamma}^{ts} \right]_{t,s=1}^T$, where each $\widetilde{\Gamma}^{ts} \triangleq \left[\widetilde{\Gamma}_{k,k'}^{ts} \right]_{(k,k')=(1,1)}^{(\kappa_t,\kappa_s)}$ with $\widetilde{\Gamma}_{k,k'}^{ts} \triangleq \left[\left(\widetilde{\gamma}_{k;ij}^t \right) \right]_{i \in \sigma_k^t}^{j \in \sigma_{k'}^s}$.

The comparison matrix: $\overline{\mathbf{\Gamma}} \triangleq$ blkdiag $\left[\overline{\Gamma}^t\right]_{t=1}^T$, where each $\overline{\Gamma}^t \triangleq$ blkdiag $\left[\overline{\Gamma}_k^t\right]_{k=1}^{\kappa_t}$ and $\overline{\Gamma}_k^t \triangleq \left[\overline{\gamma}_{k;ij}^t\right]_{i,j\in\sigma_k^t}$, where $\left(\overline{\Gamma}_k^t\right)_{ij} \triangleq \begin{cases} \gamma_{k;ii}^t & \text{if } i = j \\ -\gamma_{k;ij}^t & \text{otherwise} \end{cases}$ for $i, j \in \sigma_k^t$.

Key assumption: The matrix $\overline{\mathbf{\Gamma}} - \widetilde{\mathbf{\Gamma}}$, which has all off-diagonal entries non-positive (thus a Z-matrix), is also a P-matrix (thus a Minkowski matrix).

Writing $\widetilde{\mathbf{\Gamma}} = \widetilde{\mathbf{L}} + \widetilde{\mathbf{D}} + \widetilde{\mathbf{U}}$ as the sum of the strictly lower triangular, diagonal, and strictly upper triangular parts, respectively, we have

- $\overline{\mathbf{\Gamma}} - \widetilde{\mathbf{L}}$ is invertible and has a nonnegative inverse,
- the spectral radius of the (nonnegative) matrix $\left[\overline{\mathbf{\Gamma}} - \widetilde{\mathbf{L}}\right]^{-1}\left(\widetilde{\mathbf{D}} + \widetilde{\mathbf{U}}\right)$ is less than unity, or equivalently,
- $\exists$ positive scalars $d_{k;ij}^t$ and $\widetilde{d}_{k;i\ell}^t$ such that

$$\gamma_{k;ii}^t d_{k;ii}^t > \sum_{i\neq j\in\sigma_k^t} \gamma_{k;ij}^t\, d_{k;ij}^t + \sum_{\ell=1}^n \widetilde{\gamma}_{k;i\ell}^t\, \widetilde{d}_{k;i\ell}^t, \quad \forall\, t = 1, \cdots, T,\ k = 1, \cdots, \kappa_t,\ i \in \sigma_k^t.$$

Potential Games

Definition. A family of functions $\{\theta_i(x)\}_{i=1}^n$ on the set $\boldsymbol{\mathcal{X}}$ admits

- an **exact potential function** $P : \boldsymbol{\Omega} \to \mathbb{R}$ if P is continuous such that for all i, all $x^{-i} \in \Omega^{-i}$, and all y^i and $z^i \in \Omega^i$,

$$P(y^i, x^{-i}) - P(z^i, x^{-i}) = \theta_i(y^i, x^{-i}) - \theta_i(z^i, x^{-i});$$

- a **generalized potential function** $P : \boldsymbol{\Omega} \to \mathbb{R}$ if P is continuous such that for all i, all $x^{-i} \in \Omega^{-i}$, and all y^i and $z^i \in \Omega^i$,

$$\theta_i(y^i, x^{-i}) > \theta_i(z^i, x^{-i}) \quad \Rightarrow \quad P(y^i, x^{-i}) - P(z^i, x^{-i}) \geq \xi_i(\theta_i(y^i, x^{-i}) - \theta_i(z^i, x^{-i})),$$

for some forcing functions $\xi_i : \mathbb{R}_+ \to \mathbb{R}_+$, i.e., $\lim_{\nu\to\infty} \xi_i(t_\nu) = 0 \Rightarrow \lim_{\nu\to\infty} t_\nu = 0$.

Example Generalized $\not\Rightarrow$ exact:

$$\begin{array}{llcl} \underset{x_1 \in \mathbb{R}}{\text{minimize}} & \theta_1(x_1, x_2) \triangleq x_1 & \mid & \underset{x_2 \in \mathbb{R}}{\text{minimize}} \quad \theta_2(x_1, x_2) \triangleq x_1 x_2 + x_2 \\ \text{subject to} & -2 \leq x_1 \leq 2 & \mid & \text{subject to} \quad 1 \leq x_2 \leq 3. \end{array}$$

Generalized potential function: $P(x_1, x_2) = x_1 x_2 + x_2$. $\square$

The potential function, if it exists, is employed to gauge the progress of the algorithm.

How to recognize the existence of a potential?

The convex case. Suppose that $\theta_i(\bullet, x^{-i})$ is convex. Recalling its subdifferential, $\partial_{x^i}\theta_i(\bullet, x^{-i})$, we define the multifunction

$$\mathbf{\Theta}(x) \triangleq \prod_{i=1}^{n} \partial_{x^i}\theta_i(x), \quad x \in \boldsymbol{\mathcal{X}}.$$

Among the following four statements, it holds that **(a)** $\Leftrightarrow$ **(b)** $\Rightarrow$ **(c)** $\Leftrightarrow$ **(d)**:

(a) $\Theta(x)$ is maximally cyclically monotone on $\Omega \triangleq \prod_{i=1}^{n} \Omega^i$;

(b) $\exists$ a convex function $\psi(x)$ such that $\partial\psi(x) = \mathbf{\Theta}(x)$ for all $x \in \Omega$;

(c) $\exists$ a convex function $\psi(x)$ on Ω and continuous functions $A_i(x^{-i})$ on Ω^{-i} such that $\theta_i(x) = \psi(x) + A_i(x^{-i})$ for all $x \in \Omega$ and all $i = 1, \cdots, n$;

(d) the family $\{\theta_i(x)\}_{i=1}^{n}$ admits a convex exact potential function $P(x)$.

If $\nabla_{x^i}\theta_i(x)$ is differentiable, the existence of a (differentiable) potential is related to the symmetry of the Jacobian of the vector function $(\nabla_{x^i}\theta_i(x))_{i=1}^{n}$.

Player selection rule is

• **Essentially covering** if $\exists$ an integer $T \geq 1$ such that

$$\mathcal{N}_\nu \cup \mathcal{N}_{\nu+1} \cup \ldots \cup \mathcal{N}_{\nu+T-1} = \{1, 2, \ldots, n\}, \quad \forall\, \nu = 1, 2, \ldots,$$

so that within every T iterations, all players will have updated their strategies at least once.

[Unlike partitioning, the above index sets may overlap, resulting in some players updating their strategies more than once during these T iterations.]

• **Randomized** if the players are chosen randomly, identically, and independently from the previous iterations so that

$$\Pr(j \in \mathcal{N}_\nu) = p_j \geq p_{\min} > 0,\ \forall\, j = 1, 2, \ldots, n,\ \forall \nu = 1, 2, \ldots.$$

Postulates on objectives and their surrogates:

• Each $\theta_i(x) = f_i(x) + g_i(x^i)$ for some differentiable function f_i and convex function g_i.

• Correspondingly, $\widehat{\theta}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; z) = g_i(x^i) + \widehat{f}_i^{\sigma_k^\nu}(x^{\sigma_k^\nu}; z)$, where the family $\left\{ \widehat{f}_i^{\sigma_k^\nu}(\bullet; x^\nu) \right\}_{i \in \sigma_k^\nu}$ admits an exact potential function $\widehat{f}_{\sigma_k^\nu}(\bullet; x^\nu)$ satisfying

• **Strong convexity:** there exists a constant $\eta > 0$ such that

$$\widehat{f}_{\sigma_k^\nu}(\widetilde{x}^{\,\sigma_k^\nu}; y) \,\geq\, \widehat{f}_{\sigma_k^\nu}(x^{\sigma_k^\nu}; y) + \nabla_{x^{\sigma_k^\nu}} \widehat{f}_{\sigma_k^\nu}(x^{\sigma_k^\nu}; y)^T \left(\, \widetilde{x}^{\,\sigma_k^\nu} - x^{\sigma_k^\nu} \,\right) + \frac{\eta}{2} \| \, \widetilde{x}^{\,\sigma_k^\nu} - x^{\sigma_k^\nu} \, \|^2$$

for all $x, \widetilde{x}^{\sigma_k^\nu} \in \mathcal{X}^{\,\sigma_k^\nu}$, and y in $\boldsymbol{\mathcal{X}}$.

• **Gradient consistency:** $\nabla_{x^i} f_i(x)^T (u^i - x^i) = \left(\nabla_{x^i} \widehat{f}_i^{\,\sigma_k^\nu}(\bullet, x^{\sigma_k^\nu;-i}; x)|_{x^i} \right)^T (u^i - x^i)$
for all u^i, $x^i \in \mathcal{X}^{\,i}$, $x^{-i} \in \mathcal{X}^{-i}$ and $i \in \sigma_k^\nu$.

Convergence with constant step-size. Assume

- an exact potential function P exists;

- a scalar $L > 0$ exists such that $\|\nabla f_i(x) - \nabla f_i(x')\| \leq L\|x - x'\|$ for all x, $x' \in \boldsymbol{\mathcal{X}}$ and all $i = 1, \cdots, n$;

- a constant step-size $\tau \in (0, 2\eta/L)$ is employed.

Then, for an essentially covering player selection rule, every limit point of the iterates generated by the unified algorithm is a QNE of the game $\mathcal{G}$. Same holds with probability one for the randomized player selection rule.

Generalized potential games: 2 more restrictions:

- **Point Gauss-Seidel**, i.e., each σ_k^ν is a singleton;

- **Tight upper-bound assumption:**

$$\widehat{\theta}_{\sigma^\nu}(x^{\sigma^\nu}; y) \geq \theta_{\sigma^\nu}(x^{\sigma^\nu}; y^{-\sigma^\nu}) \quad \text{and} \quad \widehat{\theta}_{\sigma^\nu}(x^{\sigma^\nu}; x) = \theta_{\sigma^\nu}(x^{\sigma^\nu}; x^{-\sigma^\nu}), \quad \forall x, y \in \boldsymbol{\mathcal{X}}.$$

Concluding remarks

• We have introduced and analyzed the convergence of a unified distributed algorithm for computing a QNE of a multi-player game with non-smooth, non-convex player objective functions and with decoupled convex constraints.

• The algorithm employs a family of surrogate objective functions to deal with the non-convexity and non-differentiability of the original objective functions and solves subgames in parallel involving deterministic or randomized choice of non-overlapping groups of players.

• The convergence analysis is based on two approaches: contraction and potential; the former relies on a spectral condition while the latter assumes the existence of a potential function.

• Extension of the algorithm and analysis to games with coupled convex constraints can be done by introducing multipliers (or prices) of such constraints that are updated in an outer iteration.

• Non-convex constraints are presently being researched.

Thank you!