Asynchronous Newton-Raphson Consensus for Distributed Convex Optimization

F. Zanella, D.Varagnolo, A. Cenedese, G. Pillonetto, L. Schenato

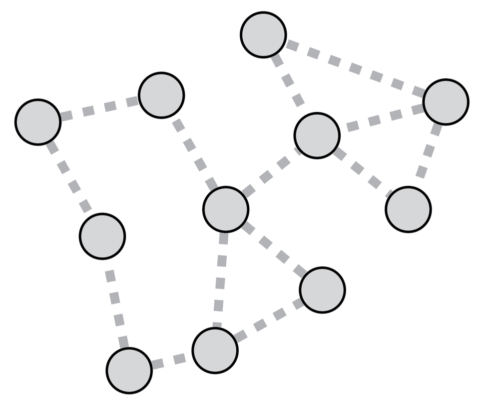
Department of Information Engineering, University of Padova KTH Royal Institute of Technology

Problem description

Aim:

- develop a **distributed convex optimization algorithm** suitable for

$$x^* := \arg\min_x f(x) \qquad \text{with} \qquad f(x) = \sum_{i=1}^{N} f_i(x)$$



Muulti-Agents Scenario:

- use only local information
- exchange messages only with neighbors
- use asynchronous communications

Assumption: $f_i : \mathbb{R} \mapsto \mathbb{R}$ smooth, closed, proper and strictly convex

Derivation of the algorithm

Simplified case – **quadratics**:

$$f_i(x) = \frac{1}{2} a_i \big(x - b_i\big)^2 \qquad \Rightarrow \qquad x^* = \frac{\sum_{i=1}^{N} a_i b_i}{\sum_{i=1}^{N} a_i} = \frac{\frac{1}{N}\sum_{i=1}^{N} a_i b_i}{\frac{1}{N}\sum_{i=1}^{N} a_i}$$

a well known structure:

parallel of two average consensus

Question: *how to extend this to the general case?* First attempt:

Prototype

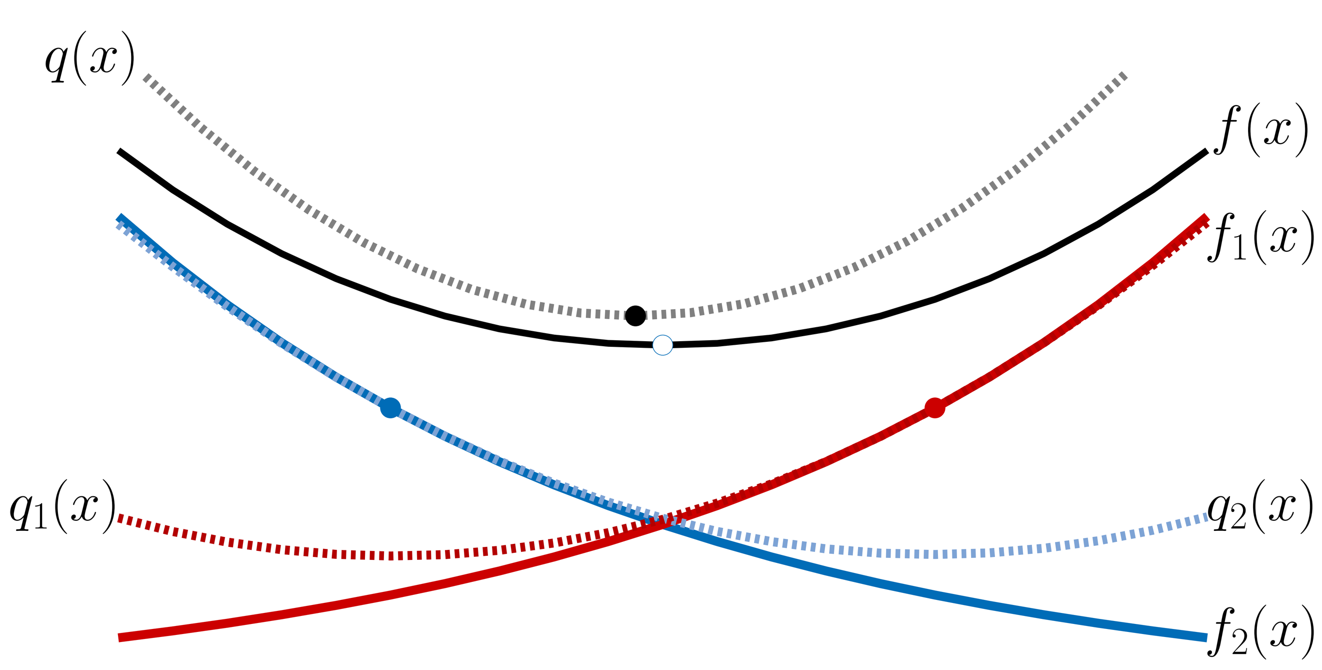
let P = average consensus matrix

1: initialize as follows:
randomly select the $x_i(0)$'s
compute $y_i(0) = f_i''\big(x_i(0)\big)x_i(0) - f_i'\big(x_i(0)\big)$ and $z_i(0) = f_i''\big(x_i(0)\big)$

2: run the average consensus
$\boldsymbol{y}^+ = P\boldsymbol{y}$
$\boldsymbol{z}^+ = P\boldsymbol{z}$

3: when converged, compute $\boldsymbol{x} = \dfrac{\boldsymbol{y}}{\boldsymbol{z}}$

Proposition: the $\boldsymbol{x}$ computed by the previous algorithm corresponds to the exact Newton direction $\Rightarrow$ *the previous procedure indicates where to move* $\Rightarrow$ updating $\boldsymbol{x}$ through $\dfrac{\boldsymbol{y}}{\boldsymbol{z}}$ gets us closer to x^*:



→ **directions for generalizations:**

1. track the changing $x_i(k)$
2. make the local estimation step $\boldsymbol{x}(k+1) = \dfrac{\boldsymbol{y}(k+1)}{\boldsymbol{z}(k+1)}$ milder
3. make communications asynchronous

The main algorithm: NRC

let
$g_i(k) := f_i''\big(x_i(k)\big)x_i(k) - f_i'\big(x_i(k)\big)$
$h_i(k) := f_i''\big(x_i(k)\big)$
let also $S(k)$, $E(k)$, $P(k)$ be, e.g.,

$$S(k) = \begin{bmatrix} 0 &&&&& \\ & 1 &&&& \\ && 0 &&& \\ &&& 0 && \\ &&&& 0 & \\ &&&&& 0 \end{bmatrix} \quad E(k) = \begin{bmatrix} 0 &&&&& \\ & 1 &&&& \\ && 0 &&& \\ &&& 0 && \\ &&&& 1 & \\ &&&&& 0 \end{bmatrix} \quad P(k) = \begin{bmatrix} 1 &&&&& \\ & 1-\alpha &&& \alpha & \\ && 1 &&& \\ &&& 1 && \\ & \alpha &&& 1-\alpha & \\ &&&&& 1 \end{bmatrix}$$

1: initialize as follows:
randomly select the $x_i(0)$'s
set $y_i(0) = z_i(0) = g_i(-1) = h_i(-1) = 0$

(main algorithm)

2: **for** $k = 1, 2, \ldots$ **do**

3: $\quad \boldsymbol{y}(k+1) = P(k)\Big(\boldsymbol{y}(k) + E(k)\big(\boldsymbol{g}(k) - \boldsymbol{g}(k-1)\big)\Big)$

4: $\quad \boldsymbol{z}(k+1) = P(k)\Big(\boldsymbol{z}(k) + E(k)\big(\boldsymbol{h}(k) - \boldsymbol{h}(k-1)\big)\Big)$

5: $\quad \boldsymbol{x}(k+1) = \boldsymbol{x}(k) + \varepsilon S(k)\left(-\boldsymbol{x}(k) + \dfrac{\boldsymbol{y}(k+1)}{\boldsymbol{z}(k+1)}\right)$

6: **end for**

Why the algorithm works (for sufficiently small ε):

- $y_i(k) \approx \dfrac{1}{N}\sum_{i=1}^{N}\big(f_i''(x_i)x_i - f_i'(x_i)\big)$
- $z_i(k) \approx \dfrac{1}{N}\sum_{i=1}^{N} f_i''(x_i)$
- $\dot{x}_{\text{ave}} \approx -\dfrac{f'(x_{\text{ave}})}{f''(x_{\text{ave}})}$

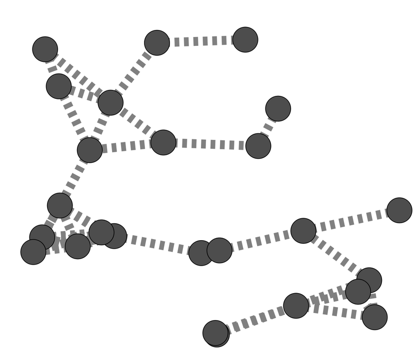
Convergence properties

• **Theorem:** ***(uniform activation)*** if on the long run all the agents are activated the same number of times then NRC is ***globally convergent***

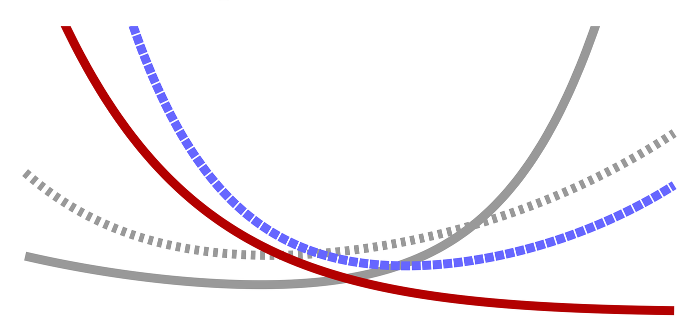
• **Theorem:** ***(persistent activation)*** if every agent is activated at least once in every suff. large time windows then NRC is ***locally convergent***

Performance

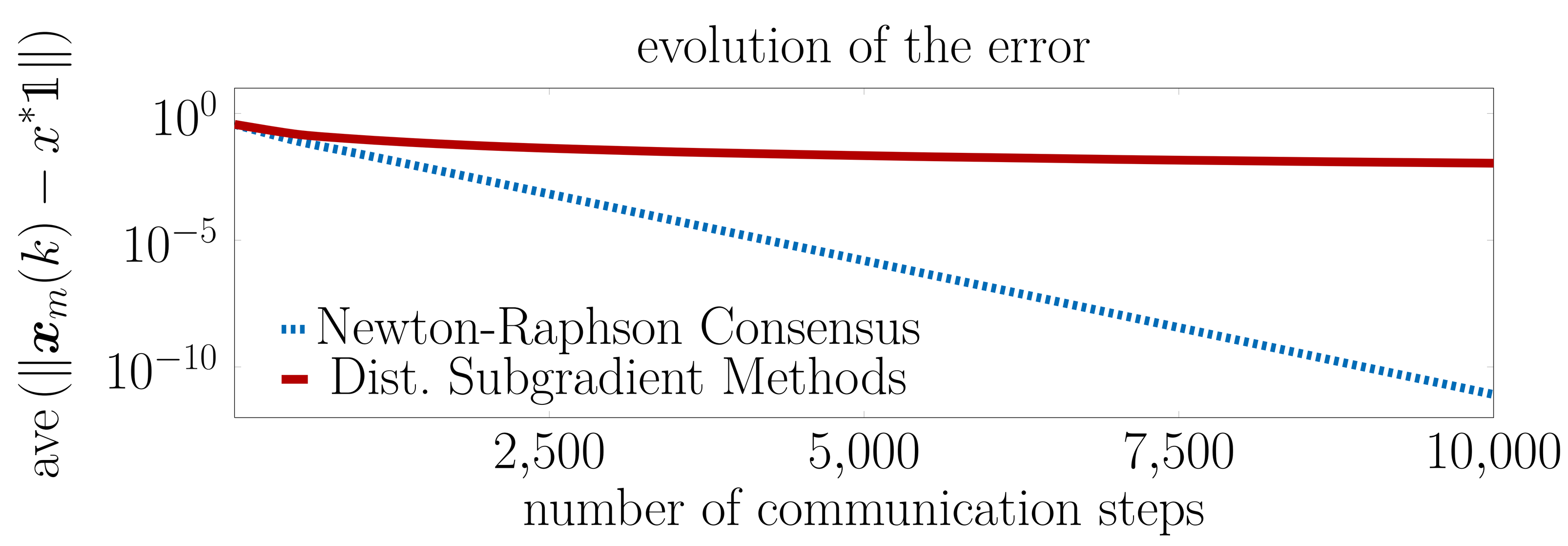
communication graph



examples of local costs



evolution of the error



References

- Zanella, Varagnolo, Cenedese, Pillonetto, Schenato, *Newton-Raphson Consensus for Distributed Convex Optimization*, IEEE TAC 2012
- Zanella, Varagnolo, Cenedese, Pillonetto, Schenato, *Asynchronous N.-R. Consensus for Distributed Convex Optimization*, NecSys 2012

{fzanella | cenedese | giapi | schenato }@dei.unipd.it damiano@kth.se