A Primal-Dual Algorithm for Distributed Resource Allocation

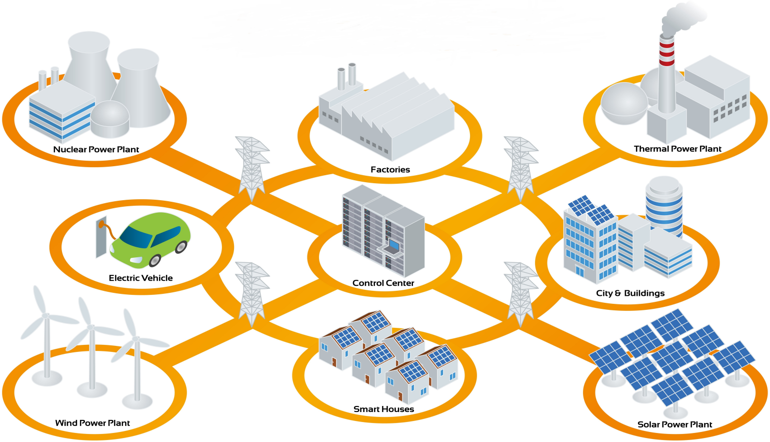
Dongsheng Ding

joint work with

Mihailo R. Jovanović

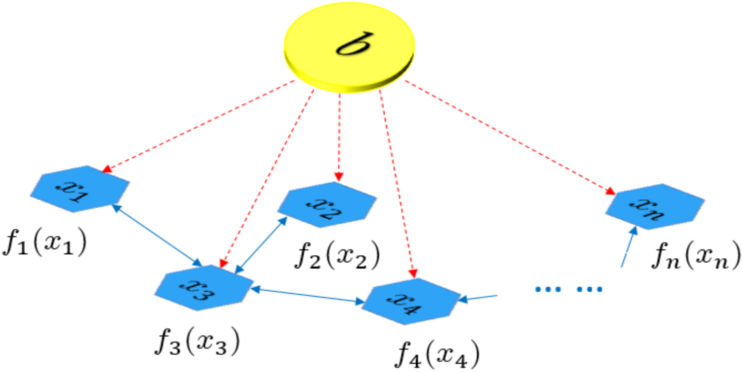
American Control Conference, Milwaukee, WI, June 29, 2018

Economic dispatch problem



Source: Solar+Power, July, 2015

Distributed resource allocation



- x_i – nodes in a connected, undirected graph with Laplacian L
- $f_i(x_i)$ – a local cost function for node x_i

Problem formulation

$$
\begin{array}{lll}
\underset{x}{\text{minimize}} & f(x) & \\
\text{subject to} & \mathbb{1}^T x - b = 0 & \text{resource constraint} \\
 & x \in \Omega & \text{set constraint}
\end{array}
$$

- $x = [\, x_1 \ \cdots \ x_n \,]^T$, $x_i \in \mathbb{R}$
- $f(x) = \sum_i f_i(x_i)$ – convex, continuously differentiable
- b – globally known
- $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, Ω_i – closed, convex

Challenges

- **non-smoothness**
  - set constraint: $x \in \Omega$
- **global information**
  - resource constraint: $\mathbb{1}^T x - b = 0$
- **convergence analysis of distributed algorithms**

PROPOSED DISTRIBUTED ALGORITHM

Distributed primal-dual algorithm

$$\dot{x} = -L\,(\nabla f(x) + \nabla M_{\mu g}(x + \mu y)) \quad \text{primal descent}$$

$$\dot{y} = \mu\,(\nabla M_{\mu g}(x + \mu y) - y) \quad \text{dual ascent}$$

- **key features**
  - $\mathbb{1}^T x(0) - b = 0$ – initialization
  - $\mathbb{1}^T x(t) - b = 0, \forall t \geq 0$ – feasible anytime
  - $\nabla f(x), \nabla M_{\mu g}(x + \mu y)$ – distributed computations

Kia, Syst. Control Lett. '17

Cherukuri & Cortés, Autom. '16

Nedić, Olshevsky, Shi, arXiv:1706.05441

Non-smooth optimization

- **indicator function of Ω**

$$I_{\Omega_i}(z_i) = \begin{cases} 0, & z_i \in \Omega_i \\ \infty, & \text{otherwise} \end{cases}$$

$$g(z) = \sum_{i=1}^{n} g_i(z_i) = \sum_{i=1}^{n} I_{\Omega_i}(z_i)$$

- **Non-smooth optimization with equality constraints**

$$\begin{aligned} \underset{x,z}{\text{minimize}} \quad & f(x) + g(z) \\ \text{subject to} \quad & \mathbb{1}^T x - b = 0 \\ & x - z = 0 \end{aligned}$$

Proximal operator

- **proximal operator**

$$\mathbf{prox}_{\mu g}(v) \;:=\; \operatorname*{argmin}_{z}\; g(z) + \frac{1}{2\mu}\|z - v\|^2$$

- **Moreau envelope**

$$M_{\mu g}(v) \;:=\; g(\mathbf{prox}_{\mu g}(v)) + \frac{1}{2\mu}\|\mathbf{prox}_{\mu g}(v) - v\|_2^2$$

continuously differentiable, even when g is not

$$\nabla M_{\mu g}(v) \;=\; \frac{1}{\mu}\left(v - \mathbf{prox}_{\mu g}(v)\right)$$

Parikh & Boyd, FnT in Optimization '14

Augmented Lagrangian

$$
\begin{aligned}
\mathcal{L}_\mu(x,\, z;\, \lambda, y) &= \mathcal{L}(x, z;\, \lambda, y) + \frac{1}{2\mu}((\mathbb{1}^T x - b)^2 + \|x - z\|^2) \\
&= f(x) + \lambda(\mathbb{1}^T x - b) + \frac{1}{2\mu}(\mathbb{1}^T x - b)^2 \\
&\quad + g(z) + y^T(x - z) + \frac{1}{2\mu}\|x - z\|^2
\end{aligned}
$$

Non-smooth

- **Lagrangian**

$$
\mathcal{L}(x, z;\, \lambda, y) = f(x) + g(z) + \lambda(\mathbb{1}^T x - b) + y^T(x - z)
$$

Partial minimization over z

$$\begin{aligned}\mathcal{L}_\mu(x,\, z;\ \lambda, y) \;&=\; f(x) \,+\, \lambda(\mathbb{1}^T x - b) \,+\, \frac{1}{2\mu}(\mathbb{1}^T x - b)^2 \\ &\quad + \, g(z) \,+\, \underbrace{y^T(x-z) \,+\, \frac{1}{2\mu}\|x-z\|^2}_{\frac{1}{2\mu}\|z-(x+\mu y)\|^2 \,-\, \frac{\mu}{2}\|y\|^2}\end{aligned}$$

$$\begin{aligned}&=\; f(x) \,+\, \lambda(\mathbb{1}^T x - b) \,+\, \frac{1}{2\mu}(\mathbb{1}^T x - b)^2 \\ &\quad + \, \underbrace{g(z) \,+\, \frac{1}{2\mu}\|z-(x+\mu y)\|^2}_{\text{relates to } z} \,-\, \frac{\mu}{2}\|y\|^2\end{aligned}$$

Manifold of minimizers

$$\underset{z}{\text{argmin}}\ \mathcal{L}_\mu(x,\ z;\ \lambda, y)$$

$$\Updownarrow$$

$$\underset{z}{\text{argmin}}\ g(z)\ +\ \frac{1}{2\mu}\|z - (x + \mu y)\|^2$$

- **explicit minimizer of $\mathcal{L}_\mu(x,\ z;\ \lambda, y)$ over z**

$$z_\mu^\star(x,\ y)\ =\ \mathbf{prox}_{\mu g}(x + \mu y)$$

Proximal augmented Lagrangian

$$\mathcal{L}_\mu(x;\lambda,y) \;:=\; \mathcal{L}_\mu(x,z;\lambda,y)\Big|_{z\,=\,z_\mu^\star(x,y)}$$

$$\begin{aligned}\mathcal{L}_\mu(x;\lambda,y) \;&=\; f(x) \,+\, \lambda(\mathbb{1}^T x - b) \,+\, \frac{1}{2\mu}(\mathbb{1}^T x - b)^2 \,-\, \frac{\mu}{2}\|y\|^2 \\ &\quad + g(\mathbf{prox}_{\mu g}(x+\mu y)) \,+\, \frac{1}{2\mu}\|\mathbf{prox}_{\mu g}(x+\mu y) - (x+\mu y)\|^2 \\ &= f(x) \,+\, \lambda(\mathbb{1}^T x - b) \,+\, \frac{1}{2\mu}(\mathbb{1}^T x - b)^2 \,-\, \frac{\mu}{2}\|y\|^2 \\ &\quad + M_{\mu g}(x+\mu y)\end{aligned}$$

continuously differentiable

Dhingra, Khong, Jovanović, arXiv:1610.04514

Arrow-Hurwicz-Uzawa gradient flow

- **primal-descent dual-ascent**

$$\dot{x} = -\left(\nabla f(x) + \nabla M_{\mu g}(x+\mu y) + \underbrace{\mathbb{1}\lambda + \frac{1}{\mu}\mathbb{1}(\mathbb{1}^T x - b)}_{\text{Vectors with the same element}}\right)$$

$$\dot{\lambda} = \mathbb{1}^T x - b$$

$$\dot{y} = \mu\left(\nabla M_{\mu g}(x+\mu y) - y\right).$$

- **downside**
  - broadcast $\mathbb{1}^T x$ and λ to every x_i
  - λ needs to access every x_i

Cherukuri, Gharesifard, Cortés, SIAM J. Control Optim. ’17

Distributed primal-dual algorithm

$$\dot{x} = -L\,(\nabla f(x) + \nabla M_{\mu g}(x + \mu y)) \quad \text{primal descent}$$

$$\dot{y} = \mu\,(\nabla M_{\mu g}(x + \mu y) - y) \quad \text{dual ascent}$$

- **key features**
  - $\mathbb{1}^T x(0) - b = 0$ – initialization
  - $\mathbb{1}^T x(t) - b = 0, \forall t \geq 0$ – feasible anytime
  - $\nabla f(x), \nabla M_{\mu g}(x + \mu y)$ – distributed computations

Kia, Syst. Control Lett. '17

Cherukuri & Cortés, Autom. '16

Nedić, Olshevsky, Shi, arXiv:1706.05441

Optimality

- **set of stationary points**

$$\bar{\Omega} = \left\{ (x,y) \,\middle|\, \begin{array}{l} L\left(\nabla f(x) + \nabla M_{\mu g}(x+\mu y)\right) \;=\; 0 \\ \nabla M_{\mu g}(x+\mu y) - y \;=\; 0,\; \mathbb{1}^T x - b \;=\; 0 \end{array} \right\}$$

- **KKT conditions**

$$\Omega^\star = \left\{ (x,z,\lambda,y) \,\middle|\, \begin{array}{l} \nabla f(x) + \mathbb{1}\lambda + y \;=\; 0,\; x - z \;=\; 0 \\ y \in \partial g(z),\; \mathbb{1}^T x - b \;=\; 0 \end{array} \right\}$$

  - for any $(\bar{x}, \bar{y}) \in \bar{\Omega}$, there exists $z^\star$ and $\lambda^\star$ such that $(\bar{x}, z^\star, \lambda^\star, \bar{y}) \in \Omega^\star$

Coordinate transformation

- **eliminate the average**

$$x \;=\; U\xi + \mathbb{1}\frac{b}{n}$$

  - $L = V\Lambda V^T = \begin{bmatrix} U & \frac{1}{n}\mathbb{1} \end{bmatrix} \begin{bmatrix} \Lambda_0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U^T \\ \frac{1}{n}\mathbb{1}^T \end{bmatrix} = U\Lambda_0 U^T$

- **transformed primal-dual algorithm**

$$\begin{aligned} \dot{\xi} \;&=\; -\,\Lambda_0 U^T \left(\nabla f(x) + \nabla M_{\mu g}(x + \mu y) \right) \\ \dot{y} \;&=\; \mu \left(\nabla M_{\mu g}(x + \mu y) \;-\; y \right) \end{aligned}$$

  - $\xi(0) \in \mathbb{R}^{n-1}$, $y(0) \in \mathbb{R}^n$

Convergence analysis

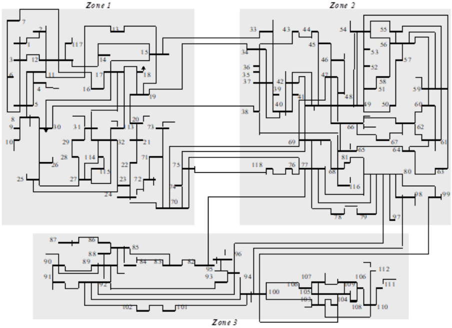
- **assumption**
  - $f(x)$ – strongly convex
  - $\nabla f(x)$ – Lipschitz continuous

- **Lyapunov stability**

$$V(\tilde{\xi},\, \tilde{y}) \;=\; \frac{1}{2}\,\tilde{\xi}^T\,\Lambda_0^{-1}\,\tilde{\xi} \;+\; \frac{1}{2}\,\|\tilde{y}\|^2$$

  - $\tilde{\xi} = \xi - \bar{\xi},\ \tilde{y} = y - \bar{y}$
  - $V(\tilde{\xi},\, \tilde{y}) > 0,\ \forall\, \tilde{\xi},\, \tilde{y} \neq 0 \ \ \&\ \ \dot{V} \leq 0$
  - largest invariant set – LaSalle-type argument

Economic dispatch example



IEEE 118-bus, IIT, Chicago

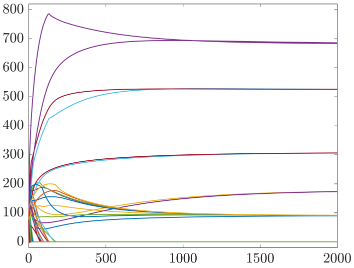
- $n = 54$ – NO. of units
- $b = 4200$ MVA – load
- $x_i \geq 0$ – power injection of unit i
- $f_i(x_i) = a_i + b_i x_i + c_i x_i^2$ – cost function
- ring network with edges $(1, 11), (11, 21), (21, 31), (31, 41), (41, 51)$

Results

- relative error

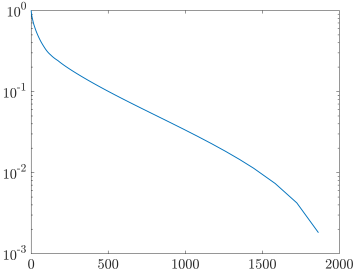
$$\sqrt{\frac{\|x-\bar{x}\|^2+\|y-\bar{y}\|^2}{\|x(0)-\bar{x}\|^2+\|y(0)-\bar{y}\|^2}}$$

$x_i(t), i = 1, \cdots, 54$



Time t

relative error



Time t

Summary

- **results**
  - PAL for the resource allocation
  - a distributed primal-dual algorithm
  - global convergence
- **future work**
  - convergence rate analysis
  - multiple resources allocation
  - dynamic resource allocation

THANK YOU!