Distributed algorithms for network optimization under non-sparse constraints

Ashish Cherukuri and **Jorge Cortés**

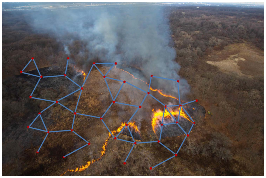
Mechanical and Aerospace Engineering
University of California, San Diego

Allerton Conference on Communication, Contol, and Computing
Monticello, Sep 28-30, 2016

Need for network optimization is pervasive

Optimizing agent operation given limited network resources

- **power networks:** generation, transmission, distribution, consumption
- **wireless communication networks:** throughput, routing, topology
- **sensor&robotic networks:** data gathering, fusion, estimation, life



Grid of the future: from vertical to flat

Integration of renewables and distributed energy resources (DERs)

From **small number** of **large** generators to **large number** of **smaller** generators

- advent of renewables, distributed energy generation
- large-scale grid optimization problems, highly dynamic
- traditional top-down approaches impractical, inefficient

Rethinking of operational&infrastructure design for efficiency and emission targets

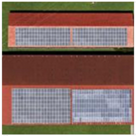
Optimized coordination for allowing&dispatching power flows originating from any point, handle dynamic loads, robust against failures, privacy, plug-and-play

Nuclear Shutdowns Put Belgians and Britons on Blackout Alert - IEEE

Multiple reactor shutdowns in Belgium and the U.K. are a reminder of the potential brittleness of power systems reliant on a small number of large generators

SPECTRUM.IEEE.ORG

September 22, 2014



How Rooftop Solar Can Stabilize the Grid

Following Germany's lead, California gives advanced inverters a bigger role in the grid

SPECTRUM.IEEE.ORG | BY PETER FAIRLEY

January 21, 2015

Network optimization with non-sparse constraints

Network of n agents communicating over connected undirected graph

- **convex** cost function: $f_i : \mathbb{R} \to \mathbb{R}$, $\forall i$
- **local** constraint: $x_i^m \leq x_i \leq x_i^M$, $\forall i$
- **global** constraint: $Ax = b$, with $b \in \mathbb{R}^m$ and non-sparse $A \in \mathbb{R}^{m \times n}$

Network optimization problem

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{subject to} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Objective: distributed algorithmic solution under

- **local exchanges:** only neighbors communicate with each other
- **information:** i knows f_i, x_i^m, x_i^M and $([A]_k, b_k)$ for k such that $[A]_{k,i} \neq 0$

Sample scenario: I

Economic dispatch

*Group of n **power generators** aim to **meet power demand** while **minimizing total cost** of generation and respecting **individual generator constraints***

Economic dispatch problem

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{i=1}^n f_i(P_i) \\ \text{subject to} \quad & \textstyle\sum_{i=1}^n P_i = L \\ & P^m \leq P \leq P^M \end{aligned}$$

- load constraint is global and generator constraints are local
- $m = 1$, $A = [1, \dots, 1]$, and $b = L$

Sample scenario: II

Sensitivity analysis-based optimal power flow[1]

Given operating point, group of n ***power generators*** *seek to determine* ***cost-effective change in generation*** *to* ***meet change in demand*** *while accounting for* ***flow constraints***

Linearized optimal power flow

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{i=1}^{N_g} f_i(\Delta P_i^g) \\ \text{subject to} \quad & \textstyle\sum_{i=1}^{N_g} \Delta P_i^g = \sum_{i=1}^{N_l} \Delta P_j^d + \Lambda^\top \Delta P^g \\ & \underline{P}^f \leq \Psi \begin{bmatrix} \Delta P^g \\ \Delta P^d \end{bmatrix} \leq \overline{P}^f \\ & \underline{P}^g \leq \Delta P^g \leq \overline{P}^g \end{aligned}$$

- change in losses and flows represented using shift factors
- power balance and flow constraints are global as Λ and Ψ are non-sparse

[1] K. E. Van Horn, A. D. Domínguez-García, and P. W. Sauer. "Measurement-based real-time security-constrained economic dispatch," *IEEE Transaction on Power Systems*, vol. 31, no. 5, pp. 3548-3560, 2016.

Outline

Exact reformulation using consensus

- Decision variable for agent i is copy of network state $\hat{x}^i \in \mathbb{R}^n$
- Collective decision variable $\hat{x} = (\hat{x}^1; \hat{x}^2; \dots; \hat{x}^n) \in (\mathbb{R}^n)^n$
- $(\tilde{A}_i, \tilde{b}_i)$ are submatrices formed by rows k of A and b where $[A]_{k,i} \neq 0$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(\hat{x}_i^i) \\ \text{s.t.} \quad & \tilde{A}_i \hat{x}^i = \tilde{b}_i, \forall i \\ & x_i^m \leq \hat{x}_i^i \leq x_i^M, \forall i \\ & (\mathsf{L} \otimes I_n)\hat{x} = \mathbf{0}_{n^2} \end{aligned}$$

L is graph Laplacian

All constraints are local (computable using information exchange with neighbors) in the reformulated problem!

Exact reformulation using consensus

- Decision variable for agent i is copy of network state $\hat{x}^i \in \mathbb{R}^n$
- Collective decision variable $\hat{x} = (\hat{x}^1; \hat{x}^2; \dots; \hat{x}^n) \in (\mathbb{R}^n)^n$
- $(\tilde{A}_i, \tilde{b}_i)$ are submatrices formed by rows k of A and b where $[A]_{k,i} \neq 0$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(\hat{x}_i^i) \\ \text{s.t.} \quad & \tilde{A}_i \hat{x}^i = \tilde{b}_i, \forall i \\ & x_i^m \leq \hat{x}_i^i \leq x_i^M, \forall i \\ & (\mathsf{L} \otimes I_n)\hat{x} = \mathbf{0}_{n^2} \end{aligned}$$

L is graph Laplacian

Proposition

Original problem and consensus-based formulation have the same optimizers

Exact reformulation using consensus

- Decision variable for agent i is copy of network state $\hat{x}^i \in \mathbb{R}^n$
- Collective decision variable $\hat{x} = (\hat{x}^1; \hat{x}^2; \ldots; \hat{x}^n) \in (\mathbb{R}^n)^n$
- $(\tilde{A}_i, \tilde{b}_i)$ are submatrices formed by rows k of A and b where $[A]_{k,i} \neq 0$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(\hat{x}_i^i) \\ \text{s.t.} \quad & \tilde{A}_i \hat{x}^i = \tilde{b}_i, \forall i \\ & x_i^m \leq \hat{x}_i^i \leq x_i^M, \forall i \\ & (\mathsf{L} \otimes I_n)\hat{x} = \mathbf{0}_{n^2} \end{aligned}$$

L is graph Laplacian

Distributed implementation:

- size of the interchanged messages is order n
- either communication complexity or time complexity suffers

Exact reformulation using auxiliary variables

- for $k \in \{1, \ldots, m\}$, let $y^k \in \mathbb{R}^n$ be auxiliary variable for k-th constraint
- decision variable for agent i is $(x_i, \{y_i^k\}_{k=1}^m)$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & \operatorname{diag}([A]_k)x + \mathsf{L}y^k = \frac{b_k}{\mathbf{1}_n^\top e^k} e^k, \ \forall k \\ & x^m \leq x \leq x^M \end{aligned}$$

$e^k \in \mathbb{R}^n$ is defined by $e_i^k = \begin{cases} 1, & \text{if } [A]_{k,i} \neq 0 \\ 0, & \text{otherwise} \end{cases}$

All constraints are local in the reformulated problem!

Exact reformulation using auxiliary variables

- for $k \in \{1, \dots, m\}$, let $y^k \in \mathbb{R}^n$ be auxiliary variable for k-th constraint
- decision variable for agent i is $(x_i, \{y_i^k\}_{k=1}^m)$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \leq x \leq x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & [A]_{k,i} x_i + \sum_{j \in \mathcal{N}_i} (y_i^k - y_j^k) = \frac{b_k}{\mathbf{1}_n^\top e^k} e_i^k, \ \forall k, i \\ & x^m \leq x \leq x^M \end{aligned}$$

$e^k \in \mathbb{R}^n$ is defined by $e_i^k = \begin{cases} 1, & \text{if } [A]_{k,i} \neq 0 \\ 0, & \text{otherwise} \end{cases}$

All constraints are local in the reformulated problem!

Exact reformulation using auxiliary variables

- for $k \in \{1, \dots, m\}$, let $y^k \in \mathbb{R}^n$ be auxiliary variable for k-th constraint
- decision variable for agent i is $(x_i, \{y_i^k\}_{k=1}^m)$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \le x \le x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & \operatorname{diag}([A]_k)x + \mathsf{L}y^k = \frac{b_k}{\mathbf{1}_n^\top e^k} e^k, \ \forall k \\ & x^m \le x \le x^M \end{aligned}$$

$e^k \in \mathbb{R}^n$ is defined by $e_i^k = \begin{cases} 1, & \text{if } [A]_{k,i} \neq 0 \\ 0, & \text{otherwise} \end{cases}$

Proposition

Original problem and reformulation have same optimizers

Key fact: $\mathbf{1}_n^\top \left(\operatorname{diag}([A]_k)x + \mathsf{L}y^k = \frac{b_k}{\mathbf{1}_n^\top e^k} e^k\right)$ yields $[A]_k x = b_k$

Exact reformulation using auxiliary variables

- for $k \in \{1, \dots, m\}$, let $y^k \in \mathbb{R}^n$ be auxiliary variable for k-th constraint
- decision variable for agent i is $(x_i, \{y_i^k\}_{k=1}^m)$

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & Ax = b \\ & x^m \le x \le x^M \end{aligned}$$

Exact reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & \operatorname{diag}([A]_k)x + \mathsf{L}y^k = \frac{b_k}{\mathbf{1}_n^\top e^k} e^k, \ \forall k \\ & x^m \le x \le x^M \end{aligned}$$

$e^k \in \mathbb{R}^n$ is defined by $e_i^k = \begin{cases} 1, & \text{if } [A]_{k,i} \neq 0 \\ 0, & \text{otherwise} \end{cases}$

Distributed implementation:

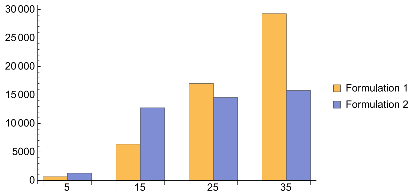
- size of the interchanged messages is of order $m + 1$
- scalable implementation when m and n independent

Comparison

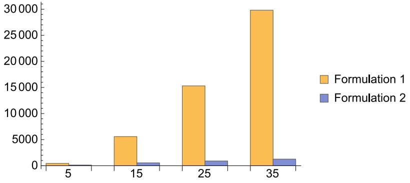
Economic dispatch problem

$$\min\left\{\sum_{i=1}^{n} c_i P_i^2 \;\middle|\; \sum_{i=1}^{n} P_i = L\right\}$$

- four cases, number of generators (n): 5, 15, 25, 35
- same primal-dual dynamics for both formulations



No. of steps to convergence for different network sizes



Volume of communication at each iteration for different network sizes

Method with auxiliary variables can be generalized

Network optimization problems with "separable" inequality constraints can be reformulated in a similar way

Original problem

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^n g_i(x_i, \{x_j\}_{j \in \mathcal{N}_i}) \leq 0 \end{aligned}$$

Reformulation

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & \operatorname{diag}([g_1(\cdot), \dots, g_n(\cdot)]) + \mathsf{L}y \leq \mathbf{0}_n \end{aligned}$$

For the reformulation:

- decision variable for agent i is (x_i, y_i)
- constraints are local: for each i,

$$g_i(x_i, \{x_j\}_{j \in \mathcal{N}_i}) + \sum_{j \in \mathcal{N}_i} (y_i - y_j) \leq 0$$

Outline

Motivation for perturbation analysis

Alternative approach to make network optimization problem ‘distributed’

- sparsify matrix A by zeroing some entries
- bound distance between optimizer of original and approximated problems
- bound distance between optimal values

Perturbation analysis: general constraints

Proposition (Arbitrary convex optimization problem)

Let f be $\mathcal{C}^2$ with $0 \prec \nabla^2 f$, $\mathcal{F}_1$ and $\mathcal{F}_2$ compact, and

$$x_1^* = \operatorname{argmin}\{f(x) \mid x \in \mathcal{F}_1\} \qquad x_2^* = \operatorname{argmin}\{f(x) \mid x \in \mathcal{F}_2\}$$

Then,

$$\|x_1^* - x_2^*\| \leq \sqrt{\frac{3}{2m}}\Big(Md(\mathcal{F}_1,\mathcal{F}_2)^2 + 2Gd(\mathcal{F}_1,\mathcal{F}_2)\Big)^{1/2} + Md(\mathcal{F}_1,\mathcal{F}_2)$$

- conservative bound, not Lipschitz with respect to distance between constraint sets $\mathcal{F}_1$ and $\mathcal{F}_2$
- analysis is oblivious to geometry of $\mathcal{F}_1$ and $\mathcal{F}_2$

$d(\mathcal{F}_1,\mathcal{F}_2)$ is the Hausdorff distance between sets and

$$G = \max\{\|\nabla f(x)\| \mid x \in \mathcal{F}_1 \cup \mathcal{F}_2\}$$
$$m = \min\{\|\nabla^2 f(x)\| \mid x \in \mathcal{F}_1 \cup \mathcal{F}_2\}$$
$$M = \max\{\|\nabla^2 f(x)\| \mid x \in \mathcal{F}_1 \cup \mathcal{F}_2\}$$

Perturbation analysis: affine constraints

Proposition

For $x_0 \in \mathbb{R}^n$, $A_1, A_2 \in \mathbb{R}^{m\times n}$ *of full row-rank,* $b_1, b_2 \in \mathbb{R}^m$, *let*

$$x_1^* = \operatorname{argmin}\{\|x - x_0\|^2 \mid A_1 x = b_1\} \qquad x_2^* = \operatorname{argmin}\{\|x - x_0\|^2 \mid A_2 x = b_2\}$$

Then,

$$\|x_1^* - x_2^*\| \leq \alpha\|A_1 - A_2\| + \beta\|b_1 - b_2\|,$$

- Lipschitz bound that uses the affine nature of constraints
- still, perturbation of same magnitude to different entries of A_1 gives the same error bound, which is not desirable

---

$$\alpha = (\|x_0\| + \|b_2\|)\tilde{\alpha}(A_1, A_2)$$
$$\beta = \|A_1^\top (A_1 A_1^\top)^{-1}\|$$

Summary

Conclusions

- global affine constraints to local affine constraints
- exact reformulations and their comparison
- relaxations via perturbation analysis

Future work

- extend perturbation analysis to general objective functions
- determine entries of A that affect least the optimizer accuracy
- design algorithms to identify "optimal" sparse A
- characterize trade-off between communication cost and accuracy of solution