Message Passing for Dynamic Network Energy Management

Matt Kraning, Eric Chu, Javad Lavaei, and Stephen Boyd
Information Systems Laboratory, Department of Electrical Engineering, Stanford University

Smart grid optimization

- large network of dynamic devices, connected by lossy, capacitated lines attempting to jointly optimize network objective subject to device and line constraints
- network objective is the sum of device objective functions
- device objective functions encode both operation costs and constraints for that device
- variables for each device consist of consumption or generation for every time period as well as local variables for that device
- **we develop a decentralized message passing algorithm to efficiently solve this problem for networks of very large size by distributing computation across all devices in the network**

Model definitions

- a *network* consists of a set of *terminals* ($\mathcal{T}$), *devices* ($\mathcal{D}$), and *nets* ($\mathcal{N}$)
- each terminal $t \in \mathcal{T}$ is uniquely associated to one device and one net, and possesses a *power schedule*, $p_t \in \mathbf{R}^T$, where T is a given time horizon
- $p_t(\tau)$ is the amount of energy consumed by device d through terminal t in time period τ, where $t \in d$ ($p_t(\tau) < 0$ corresponds to energy generation)
- each device $d \in \mathcal{D}$ has a local objective function f_d that maps its associated terminals' power schedules into $\mathbf{R} \cup +\infty$
- each net $n \in \mathcal{N}$ ensures that the power schedules of its associated terminals balance in each time period, *i.e.*, $\bar{p} = 0$ where

$$\bar{p}_t \equiv \frac{1}{|n|} \sum_{t' \in n} p_{t'}, \quad t \in n, \quad \forall t \in \mathcal{T}$$

- the **optimal power scheduling problem** (OPSP) is to minimize the network objective, $f = \sum_{d \in \mathcal{D}} f_d$, subject to power balance across all nets

Intuition

- devices model generators, loads, energy storage systems, and other power sources, sinks, and converters
- nets are lossless, uncapacitated transporters of power; we can model losses and capacities by the addition of a device
- terminals are ports on a device through which power flows, either into or out of the device
- the objective function associated with a device is used to measure the cost (which can be negative, representing revenue) associated with a particular way of operating that device and includes constraints by setting its value to $+\infty$

Device examples

- a **generator** generates power $-p_{\rm gen}$ over a specified range $P^{\rm min} \le -p_{\rm gen} \le P^{\rm max}$, with ramp rate constraints $|Dp_{\rm gen}| \le R^{\rm max}$, and a quadratic cost function
- a **load** can be one of three different kinds:
  - a **fixed** load must meet a certain power profile $l \in \mathbf{R}^T$ so that $p_{\rm load} = l$
  - a **deferrable** load must exceed a minimum consumption level E over a period of time, *i.e.*, $\sum_{\tau=A}^{D} p_{\rm load}(\tau) \ge E$
  - a **curtailable** load does not have hard constraints on power requirements but linearly penalizes shortfall from a desired load profile $l \in \mathbf{R}^T$ according to $\alpha(l - p_{\rm load})_+$, for some $\alpha > 0$
- a **battery** can take in or deliver energy and has charge level $q(\tau) = q^{\rm init} + \sum_{t=1}^{\tau} p_{\rm bat}(t)$, $0 \le q \le Q^{\rm max}$, with charging and discharging rate limits $-D^{\rm max} \le p_{\rm bat} \le C^{\rm max}$
- a **transmission line** has a (flow) capacity $C^{\rm max}$ and a loss function $\ell(p_1, p_2)$ such that $p_1 + p_2 + \ell(p_1, p_2) = 0$ and $|p_1 - p_2| \le C^{\rm max}$

Prox-average message passing

- the number of variables in the OPSP is lower bounded by $|\mathcal{N}|T$, making a centralized solver computationally impractical for networks with millions or billions of devices
- rewrite the OPSP as

$$\begin{array}{ll} \text{minimize} & \sum_{d \in \mathcal{D}} f_d(p_d) + \sum_{n \in \mathcal{N}} \mathbf{1}\{\bar{z}_n = 0\} \\ \text{subject to} & p = z \end{array}$$

- form the augmented Lagrangian

$$L_\rho(p, z, u) = \sum_{d \in \mathcal{D}} f_d(p_d) + \sum_{n \in \mathcal{N}} \mathbf{1}\{\bar{z}_n = 0\} + (\rho/2)\|p - z + u\|_2^2$$

- derive the **prox-average message passing equations** via ADMM

$$\begin{aligned} p_d^{k+1} &:= \mathbf{prox}_{f_d,\rho}(p_d^k - \bar{p}_d^k - u_d^k), \quad d \in \mathcal{D}, \\ u_n^{k+1} &:= u_n^k + \bar{p}_n^{k+1}, \quad n \in \mathcal{N}, \end{aligned}$$

  where the first and second steps can be carried out in parallel by all devices and nets in the network
- when all device objective functions are convex, prox-average message passing converges to optimal value of OPSP
- primal and dual residuals

$$r^k = \bar{p}^k, \quad s^k = \rho((p^k - \bar{p}^k) - (p^{k-1} - \bar{p}^{k-1}))$$

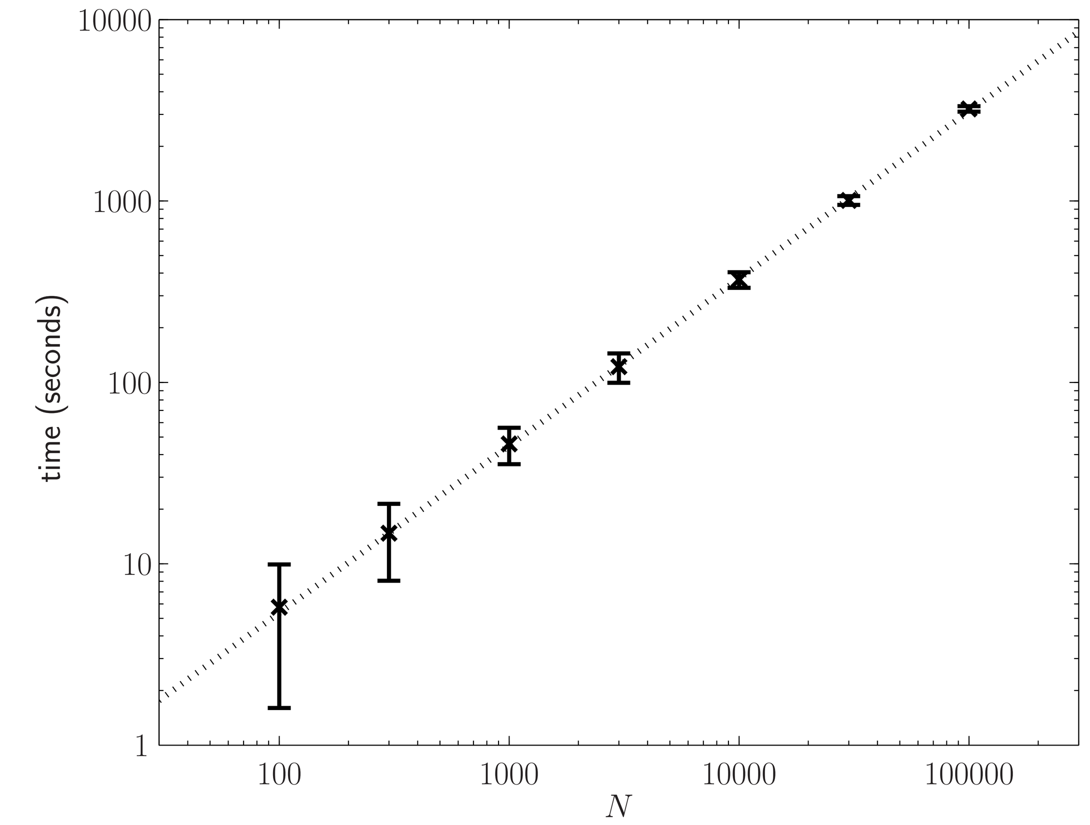
  lead to stopping criteria when

$$\|r^k\|_2/\sqrt{|\mathcal{T}|T} \le \epsilon_{\rm pri}, \quad \|s^k\|_2/\sqrt{|\mathcal{T}|T} \le \epsilon_{\rm dual}$$

- using recently developed methods, device prox functions can be computed in milliseconds on embedded processors, allowing the entire network to execute prox-average iterations at kilohertz rates

Numerical example

- generate random networks with a range $N = 100, \ldots, 100000$ of nets
  - corresponds to problems with between 3 thousand and 30 million variables
- solve times are on a 8-core 3.4Ghz Intel Xeon processor



- 52 minutes required to solve a problem with 30 million variables
- with decentralized computing, the solve time for **all** network examples would be less than 500 ms

Conclusions and extensions

- decentralized computation allows for sub second solve times independent of network size
- can be easily applied to hierarchical models
- when combined with receding horizon control, could be used for real-time network operation