Dynamic Network Energy Management via Proximal Message Passing

Matt Kraning, Eric Chu, Javad Lavaei, and **Stephen Boyd**

Google, 2/20/2013

Outline

Smart grid

- embed intelligence in energy systems to
  - do more with less
  - reduce CO2 emissions
  - handle uncertainties in generation (wind, solar, ...)
  - exploit new demand response capabilities
  - handle shift towards EVs
  - extend life of current infrastructure

Smart grid

- embed intelligence in energy systems to
  - do more with less
  - reduce CO2 emissions
  - handle uncertainties in generation (wind, solar, . . .)
  - exploit new demand response capabilities
  - handle shift towards EVs
  - extend life of current infrastructure
- cf. current system
  - load is what it is; generation scheduled to match it
  - systems built with large margins for max load

Smart grid critical technologies: The big picture

- physical layer
  - photovoltaics, switches, storage, fuel cells, ...
- infrastructure/plumbing
  - smart enabled stuff, communication protocols, security, ...
- **algorithms** (our focus)
  - real-time decision making
- economics layer
  - markets, investment, regulation, ...

Coordinating devices on the smart grid

- **setting**: a network of smart devices, that can adjust/change/defer their power consumption/generation
- **goal**: coordinate device behavior (generation/consumption) over time

Coordinating devices on the smart grid

- **setting**: a network of smart devices, that can adjust/change/defer their power consumption/generation
- **goal**: coordinate device behavior (generation/consumption) over time

- **method**: use optimization to coordinate devices
- **algorithm**: use proximal message passing to solve optimization problem

Device coordination via optimization

- devices exchange energy at nodes, in multiple time periods
  - generators
  - loads (fixed, deferrable, curtailable)
  - energy storage systems
  - transmission lines
- each device has dynamic constraints, cost function over time
- to coordinate devices, **minimize total cost subject to power balance at each node, in each time period**
- solving this optimization problem gives
  - (optimal) device power schedules
  - locational marginal prices at each node in each time period

This talk: Proximal message passing algorithm

- decentralized method to solve dynamic energy management problem
- each device schedules its own consumption/generation profile
- devices coordinate via simple message exchanges with neighbors
- can be viewed as sophisticated (location, time varying) price discovery mechanism
- can handle **enormous** problems

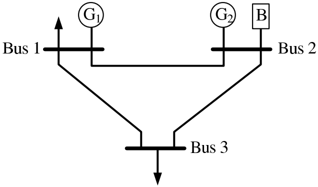
Outline

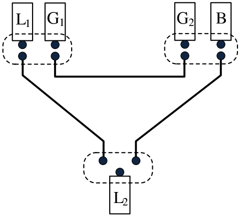
Formal network model

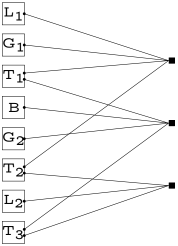
- a *network* consists of
  - a set of *terminals* $\mathcal{T}$
  - a set of *devices* $\mathcal{D}$
  - a set of *nets* $\mathcal{N}$
- $\mathcal{D}$ and $\mathcal{N}$ are partitions of $\mathcal{T}$, *i.e.*, each terminal is in exactly one device and one net
- can represent network as bipartite graph with
  - $\mathcal{D}$ and $\mathcal{N}$ the two vertex classes
  - $\mathcal{T}$ as the edges connecting them

Example

- (left) 3 buses, 2 generators, 1 battery, 2 loads, 3 transmission lines
- (middle, right) network model: 11 terminals, 3 nets, 8 devices







Terminals

- power flows into or out of terminals on each device (negative power corresponds to power generation)
- each terminal $t \in \mathcal{T}$ has a *power schedule*

  $$p_t = (p_t(1), \ldots, p_t(T)) \in \mathbf{R}^T$$

  giving power flow over time periods $\tau = 1, \ldots, T$
- set of all terminal power schedules denoted by $p \in \mathbf{R}^{|\mathcal{T}| \times T}$

Devices

- devices model general power system elements
  - generators
  - loads (deferrable, curtailable, fixed)
  - transmission lines
  - energy storage systems
  - other energy sinks, sources, and converters
- $p_d \in \mathbf{R}^{|d| \times T}$ is the set of $|d|$ power schedules for terminals in device d
- device objective function $f_d(p_d) : \mathbf{R}^{|d| \times T} \to \mathbf{R} \cup \{+\infty\}$
  - $+\infty$ used to encode device constraints
  - can also have private variables *e.g.*, state of charge for a battery

Nets

- nets are ideal (lossless, uncapacitated) energy exchange points
- $p_n \in \mathbf{R}^{|n| \times T}$ is the set of $|n|$ power schedules for terminals in net n
- semantics of nets: **power balance holds at all times**

$$\sum_{t \in n} p_t(\tau) = 0, \quad \tau = 1, \ldots, T, \quad n \in \mathcal{N}$$

Average net power imbalance

- for terminal t corresponding to net n, we define

$$\bar{p}_t = \frac{1}{|n|} \sum_{t' \in n} p_{t'}$$

  i.e., $\bar{p}_t$ averages terminal power profiles over its net
  - $\bar{p}_d = \{\bar{p}_t \mid t \in d\}$
- net power balance can be written as $\bar{p} = 0$

Dynamic optimal power flow problem

dynamic optimal power flow problem (D-OPF):

$$\begin{array}{ll} \text{minimize} & f(p) = \sum_{d \in \mathcal{D}} f_d(p_d) \\ \text{subject to} & \bar{p} = 0 \end{array}$$

- variables are terminal power schedules $p \in \mathbf{R}^{|\mathcal{T}| \times T}$
- net power balance equality constraints are linear
- other constraints, objective terms contained in device objectives
- optimal dual variables give (scaled) locational marginal prices (LMP), which are time-varying
- when all device objective functions are convex, D-OPF can be solved globally and effeciently (in principle)

Outline

Generator

- single terminal device with power schedule $p_{\rm gen}$
- cost function $\sum_{\tau=1}^{T} \phi_{\rm gen}(-p_{\rm gen}(\tau))$
- min/max power constraints: $P^{\min} \leq -p_{\rm gen} \leq P^{\max}$
- ramp-rate constraints:

$$R^{\min} \leq -\left(p_{\rm gen}(\tau+1) - p_{\rm gen}(\tau)\right) \leq R^{\max}$$

- can include other costs and constraints, *e.g.*,
  - turning on and off
  - power change costs

Transmission line

- two terminal device with power schedules p_1 and p_2
- zero cost function
- capacity constraint: $|p_1 - p_2| \leq C^{\max}$
- line loss constraint: $p_1 + p_2 = \ell(p_1, p_2)$
- $\ell(p_1, p_2) \geq 0$ is loss function ($\ell(0,0) = 0$, typically convex)

Energy storage system

- single terminal device with power schedule p_{ess}
- zero cost function
- charging/discharging rate limits $-D^{\max} \leq p_{\text{ess}} \leq C^{\max}$
- local storage state variables

$$q(\tau) = q^{\text{init}} + \sum_{t=1}^{\tau} p_{\text{ess}}(t), \quad \tau = 1, \ldots, T$$

- capacity limits $0 \leq q(\tau) \leq Q^{\max}$, $\tau = 1, \ldots, T$
- more sophisticated models can include storage cycling penalty, state-dependent charging and discharging rate limits, efficiencies

Loads

- single terminal device with power schedule p_{load}
- **fixed (non-smart) load**: $p_{\text{load}} = l$, $l \in \mathbf{R}^T$ is given load profile
- **deferrable load**: total energy consumption E in the time interval $[A, D]$:

$$\sum_{\tau=A}^{D} p_{\text{load}}(\tau) = E, \qquad 0 \leq p_{\text{load}} \leq L^{\max}$$

- **curtailable load**: pay penalty for failing to meet load profile l:

$$\alpha \sum_{\tau=1}^{T} (l(\tau) - p_{\text{load}}(\tau))_+$$

Electric vehicle charging

- single terminal device with power schedule $p_{\rm ev}$
- desired minimum state of charge profile $q^{\rm des} \in \mathbf{R}^T$
- can only be charged in time interval $[A, D]$
- charging constraints $0 \leq p_{\rm ev} \leq C^{\max}$
- charge level given by

$$q(\tau) = q^{\rm init} + \sum_{\tau'=A}^{\tau} p_{\rm ev}(\tau'),$$

- shortfall cost function

$$\alpha \sum_{\tau=A}^{D} (q^{\rm des}(\tau) - q(\tau))_+,$$

- can add terminal constraint, $q(D) = Q^{\rm cap}$

Outline

Proximal message passing algorithm

repeat until convergence:

1. *Proximal power schedule update.*

$$p_d^{k+1} := \operatorname*{argmin}_{p_d} \left(f_d(p_d) + \rho/2 \left\| p_d - (p_d^k - \bar{p}_d^k - u_d^k) \right\|_2^2 \right)$$

   in parallel, for each device

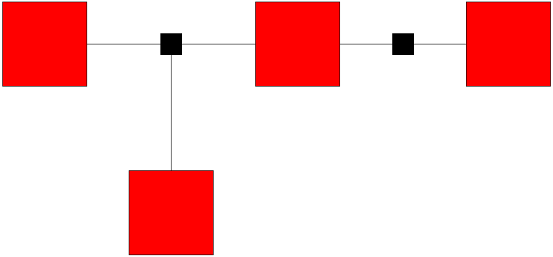
   ($\rho > 0$; RHS is **proximal operator** of f_d at $p_d^k - \bar{p}_d^k - u_d^k$)

2. *Scaled price update.*

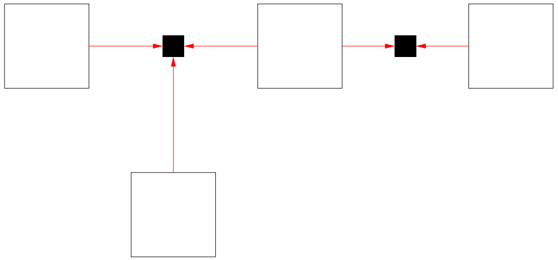
$$u_n^{k+1} := u_n^k + \bar{p}_n^{k+1}$$

   in parallel, for each net

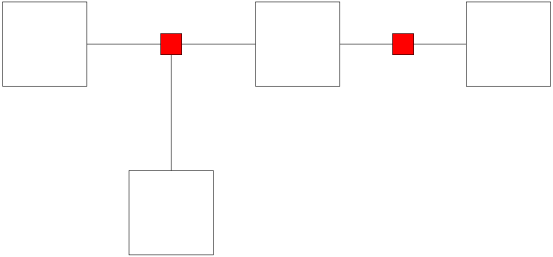
Devices compute new tentative power profiles



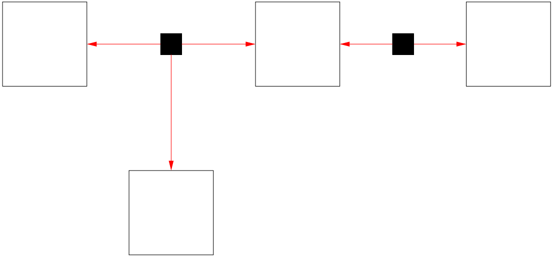
Devices send tentative power profiles to neighboring nets



Nets compute power imbalance; update prices



Nets send updated prices, power imbalance to neighboring devices



Proximal message passing algorithm

- each device only has knowledge of its own objective function
- for each device class, need to implement prox operator
- all message passing is local, between devices and adjacent nets
- no global coordination other than iteration synchronization

Convergence

if device objectives are closed convex proper and D-OPF has solution

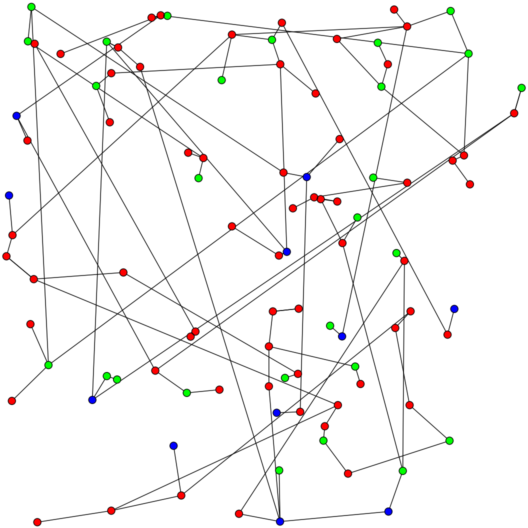
- *residual convergence*: $\bar{p}^k \to 0$ (power balance achieved)
- *objective convergence*: $\sum_{d \in \mathcal{D}} f_d(p_d^k) \to f^\star$ (operation is optimal)
- *dual variable convergence*: $\rho u^k \to y^\star$ (optimal prices found)

Outline

Numerical examples

- 140 examples: 20 each of 7 different sizes
- $|\mathcal{N}|$ ranges from 100 to 100000
- $|\mathcal{D}|$ ranges from 200 to 200000
- $|\mathcal{T}|$ ranges from 300 to 300000
- $T = 96$ (24 hour period, 15-minute intervals)
- number of variables in D-OPF ranges from 30k to 30M
- network topology (transmission line connections) chosen as random geometric graph, plus some long lines

Example network with $|\mathcal{N}| = 100$ (30k variables)



Devices

- to each net, we attach a randomly chosen single terminal device
  - generator
  - battery
  - fixed load
  - deferrable load
  - curtailable load
- device parameters chosen so that problem is feasible but challenging

Prox functions

- prox functions are easy to implement when f_d is separable in time
  - fixed load
  - curtailable load
  - transmission line

  (prox evaluation times measured in ns)

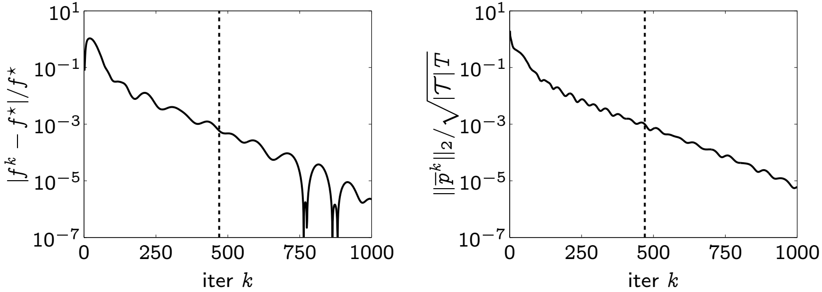
- for others, use CVXGEN to generate custom `C` code to solve QPs
  - generator
  - battery
  - deferrable load

  (prox evaluation times measured in μs)

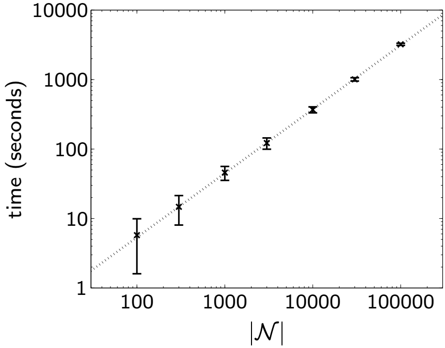
Serial multithreaded implementation

- examples run on 3.4Ghz 4-core Xeon with 8 (hyper)threads
- each prox function assigned to one of 8 threads using OpenMP
- maximum time for prox function evaluation in each iteration is ≈ 1 ms, so we can estimate fully decentralized run time

Convergence for $|\mathcal{N}| = 3000$ (1M variables)



Solve time scaling (serial)



- fit exponent is 0.923
- with parallel computation, sub second solve time for **any size** network

Outline

Handling uncertainty via receding horizon control

- in every time period
  - each device forecasts its own future costs/constraints over some horizon
  - devices coordinate (optimize) using forecasts to obtain **consumption/generation plan**
  - devices execute first period consumption/generation in plan

Handling uncertainty via receding horizon control

- in every time period
  - each device forecasts its own future costs/constraints over some horizon
  - devices coordinate (optimize) using forecasts to obtain **consumption/generation plan**
  - devices execute first period consumption/generation in plan

- reacts to changes in constraint/objective forecasts
- same method used in chemical process control, supply chain optimization, . . .
- forecasts do not need to be accurate

Handling AC power flow

- assume voltage magnitudes are fixed
- introduce voltage phase angle profile θ_t for each terminal
- add phase angle consistency constraint for each net $n = \{t_1, \ldots, t_{|n|}\}$:

$$\theta_{t_1} = \theta_{t_2} = \cdots = \theta_{t_{|n|}}$$

- local device objectives include phase angle constraints
- proximal message passing readily extended to include phase angles

Handling non-convexities

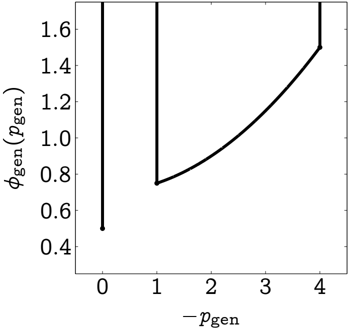
- with non-convex device objectives, D-OPF is (nominally) hard
- one approach: form convex relaxation of D-OPF (RD-OPF)

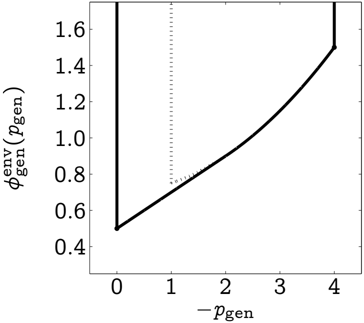
$$\begin{array}{ll} \text{minimize} & f^{\text{env}}(p) = \sum_{d \in \mathcal{D}} f_d^{\text{env}}(p_d) \\ \text{subject to} & \bar{p} = 0, \end{array}$$

  where f_d^{env} is convex envelope of f_d
- RD-OPF is convex optimization problem
  - readily solved
  - gives lower bound on D-OPF optimal value
  - provides good starting point for local optimization
  - in some cases, relaxation is tight

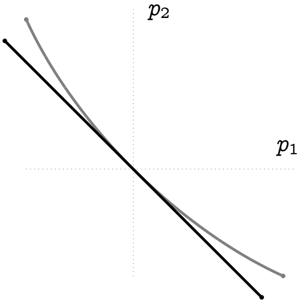
Relaxed generator

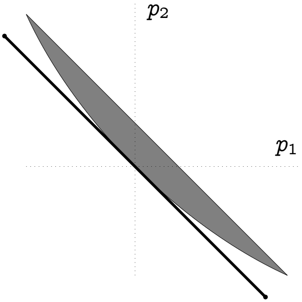
- *left*: (nonconvex) generator with power range, option to turn off
- *right*: its relaxation





Relaxed transmission line





black: lossless, capacitated line; *gray*: AC power loss

Summary and vision

- we've developed a completely decentralized method for optimal power exchange/consumption/generation on a smart grid
- decentralized computation allows for sub second solve times **independent** of network size

Summary and vision

- we've developed a completely decentralized method for optimal power exchange/consumption/generation on a smart grid
- decentralized computation allows for sub second solve times **independent** of network size
- when combined with receding horizon control, can be used for real-time network operation

Summary and vision

- we've developed a completely decentralized method for optimal power exchange/consumption/generation on a smart grid
- decentralized computation allows for sub second solve times **independent** of network size
- when combined with receding horizon control, can be used for real-time network operation
- we envision a plug-and-play system that is robust, self-healing (internet of power)