Novel Power Grid Reduction Method based on L1 Regularization

Ye Wang, Meng Li, Xinyang Yi, Zhao Song, Michael Orshansky, Constantine Caramanis

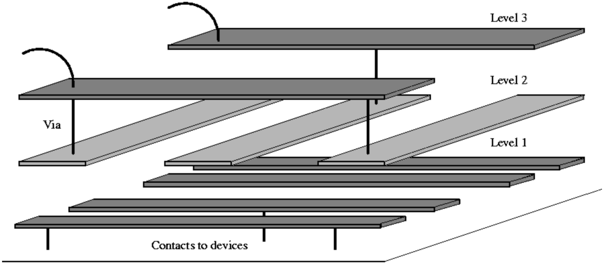
University of Texas at Austin

lhywang@utexas.edu

June 3, 2015

Power Grid Reduction: Motivation

- Large-scale power grids



- Simulation time determined by size of power grid: number of ports and elements (R,L,C)
- Speed up simulation/analysis by reducing grid size
  - Focus only on steady-state analysis (DC), hope to extend to other types of analysis (Transient, AC)

Power Grid Reduction: Formulation

- Port relation (Ohm's law) by L_G

$$L_G v = i$$

- Admittance matrix (Graph Laplacian)

$$L_G(i,j) = \begin{cases} \sum_{k,k\neq i} \omega_{ik}, & \text{if } i = j, \\ -\omega_{ij}, & \text{if } i \neq j \text{ and } \{i,j\} \in E, \\ 0, & \text{otherwise.} \end{cases}$$

- Goal: want a sparse approximation $L_{G'}$
  - With far fewer non-zeros
  - Preserve similar port relation

State-of-the-Art Methods

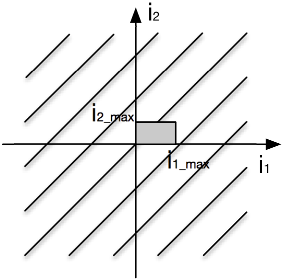
- Krylov subspace based methods [FL04,LS06]
  - Project the original system onto a low-rank Krylov subspace for efficiency
- Time-Constant Equilibration Reduction (TICER) [She99]
  - Eliminate low-degree nodes by connecting their neighbors
- Algebraic multigrid methods [SAN03]
  - Reduce the number of nodes and edges simultaneously
- Sampling based spectral sparsification approach [ZFZ14]
  - In time $O(m \log n/\epsilon^2)$, find an ϵ-power approximation G' of $O(n \log n/\epsilon^2)$ edges satisfies:

$$(1-\epsilon)v^T L_G v \leq v^T L_{G'} v \leq (1+\epsilon)v^T L_G v, \ \forall v \in \mathbb{R}^n.$$

- **They all try to build sparsifiers preserving $L_G v = i$ for all $i \in \mathbb{R}^n$... Is that necessary?**

Our Key Observation

- In practice, currents delivered from ports do not vary unboundedly
  - Peak values of currents can be estimated from system-level description or transistor-level simulation
- The actual space is a small subset of the entire space



- **How to utilize the range information not explored before?**
  - For more sparsity and accuracy of reduced power grids

Our Main Contribution

- Propose an efficient method that **allows using range information** for better sparsification
- Leverage recent advances of ℓ^1 regularization to drive sparsity
- We call it graph Sparsification by ℓ^1 regularization on Laplacian (**SparseLL**)

First Attempt for Sparsification

- Objective function: averaged error in the given range

$$\min_{L_{G'}} \int_{\Omega_V} [\|(L_G - L_{G'})v\|_2^2] dv$$

  - Allow to incorporate the range information
- Constraints: sparsity specified by ℓ^0-norm (number of non-zeros)

$$\|L_{G'}\|_0 \leq m_0$$

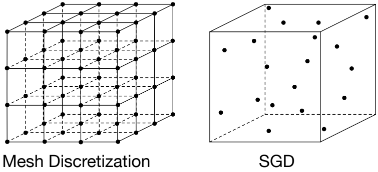
- Non-linear and non-convex in both objective and constraints, hard to solve...

Reformulation as Stochastic Optimization

- Integral discretization by deterministic mesh requires exponential number of samples

$$\int_{\Omega_V} [\|(L_G - L_{G'})v\|_2^2] dv \approx \frac{1}{N} \sum_{i=1}^{N} \|(L_G - L_{G'})v_i\|_2^2$$

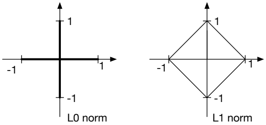
- Randomized discretization leverages fast convergence from stochastic gradient descent (SGD)



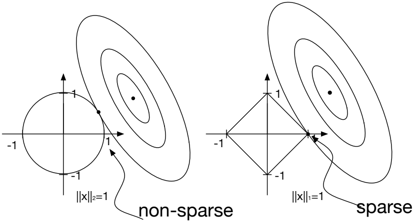
  - Sample $v_i \sim \Omega_V$, calculate gradient, update solution
  - Converge to optimal solution with rate $O(\sqrt{1/t})$ for t iterations

ℓ^1 Regularization for Sparsity

- ℓ^0 constraints are combinatorial and non-convex: result in an NP-hard problem
- ℓ^1 norm is the tightest while being convex relaxation of ℓ^0 norm



- Sparsity encouraged by spiky ℓ^1 norm



Complete SparseLL Formulation

- Objective: regularized empirical risk function

$$\min_{L_{G'}} \frac{1}{N} \sum_{i=1}^{N} \|(L_G - L_{G'})v_i\|_2^2 + \lambda \|L'_G\|_1$$

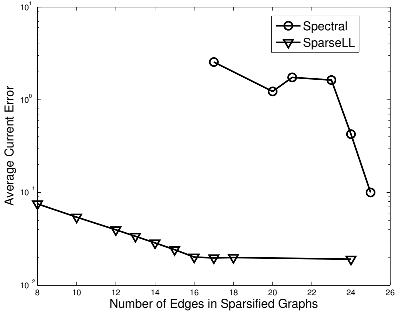
  - Parameter λ controls the degree of sparsity

- Constraints:

$$\begin{aligned} & L_{G'}(i,j) \leq 0, \text{ with } i \neq j, \\ & L_{G'} = L_{G'}^T, \\ & \sum_{j=1}^{n} L_{G'}(i,j) = 0, \forall i \in \{1, 2, \ldots, n\}. \end{aligned}$$

Experimental Results

- Optimizing a sample 17-node power grid



- A smaller error while significantly reducing # of edges

			Spectral [ZFZ14]		SparseLL		
	Nodes	Edges	Error	Edges	Error	Edges	Err Reduction
rand1	100	4000	5.40%	1031	0.18%	996	30X
rand2	500	100000	4.44%	8120	0.07%	8021	60X
rand3	1000	400000	4.80%	15114	0.03%	14213	160X
ibmpg1	5388	27000	3.80%	6703	0.01%	6570	380X

Summary

- Identified and specified a realistic range for currents
- Formulated power grid reduction as a convex optimization problem
  - With objective as average current error in that range
  - Use ℓ^1 norm to encourage sparsity
- Solved the problem using an efficient SGD algorithm