Power grids statistical modeling

Presented by: Anna Scaglione, UC Davis
joint work with Zhifang Wang and Robert J. Thomas

Motivation

- Power grids have grown organically over the past century (naturally random)
  - More balancing options: economic benefits + safety
- Our initial quest → a systematic approach to design the cyber support for the power grid
- Design and analysis of power grids has been based on reference samples and case studies
  - Does not help establishing macroscopic trends
- Can we capture in a model key features of the ensemble?

Background

What should we model?

- System of systems

Generators,Loads

Transmission Lines

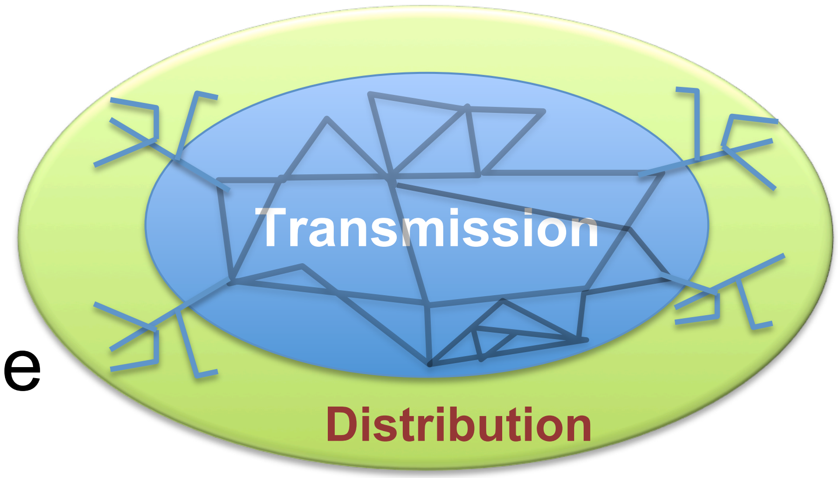
Power systems gear:
Switches,
Relays,Transformers...

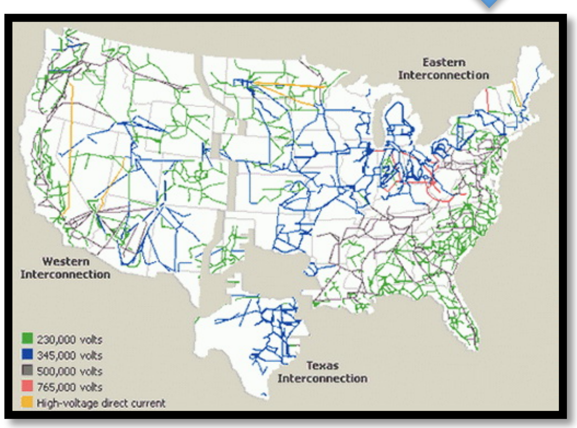
Computers and Sensors
(Substations, PLC,
Supervisory control)

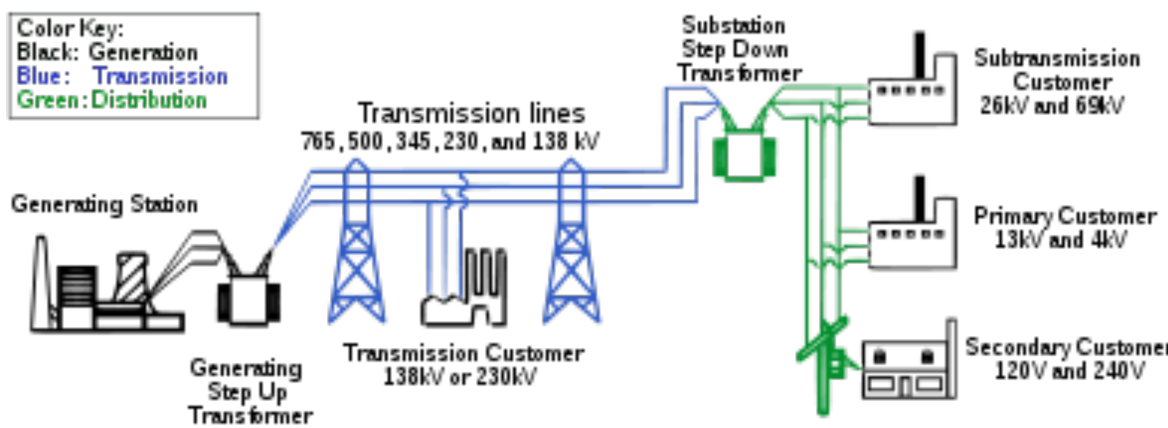
Market players (supply and
demand)

The grid transmission lines

- 3 sections:
  - High, Medium and Low voltage sections
- High and medium voltage networks wide areas

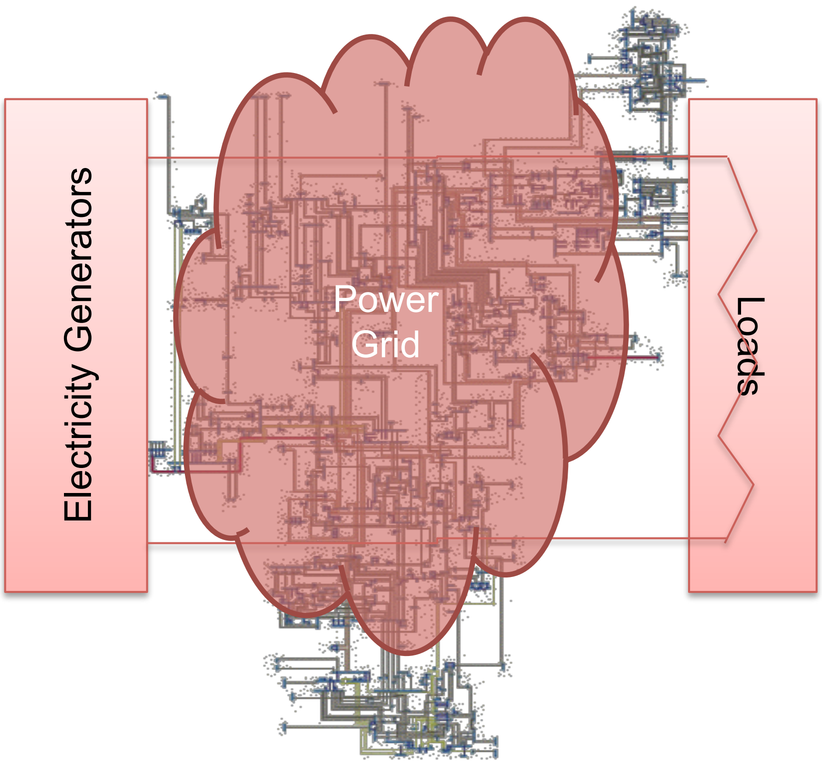






What we model

- Topological and electrical characteristics of the transmission grid
- As captured by the statistical properties of the grid admittance matrix
- Also one data point for Medium Voltage Distribution
- Leave out the distribution network



Admittance matrix and the graph topology

- Line-Node Incidence Matrix (M x N):

Line i connected to node $j \rightarrow A_{i,j} = 1,\ A_{j,i} = -1$

else $A_{i,j} = A_{j,i} = 0$

- Admittance matrix

$$Y = A^T diag(y_1, \ldots, y_M) A$$

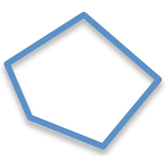
- Observation: Y is a weighted graph Laplacian
  - complex weights given by the admittances of the lines

The laws for the grid

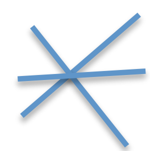
- Voltage, Currents, Powers → narrow spectrum

AC ~ 60-50 Hz

- Electrical transient dynamics → unimportant
  - Circuit laws replaced with algebraic equations (frequency) relating "*phasors*" (complex numbers whose phase and amplitude match the AC signal V and I)
- Kirchhoff's Voltage/Current laws (KVL-KCL)



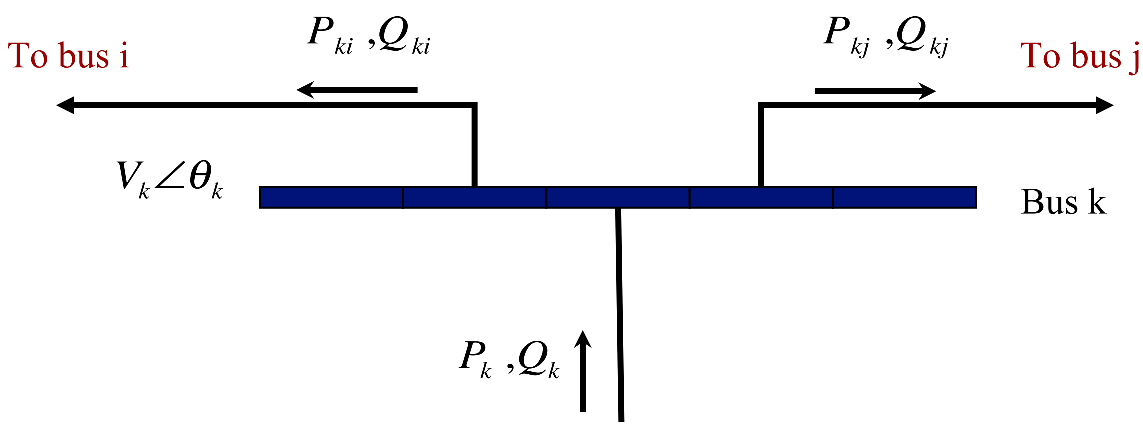
$$\sum_{i\in \text{Circuit}} V_i = 0, \qquad \sum_{i\in \text{Node}} I_i = 0$$



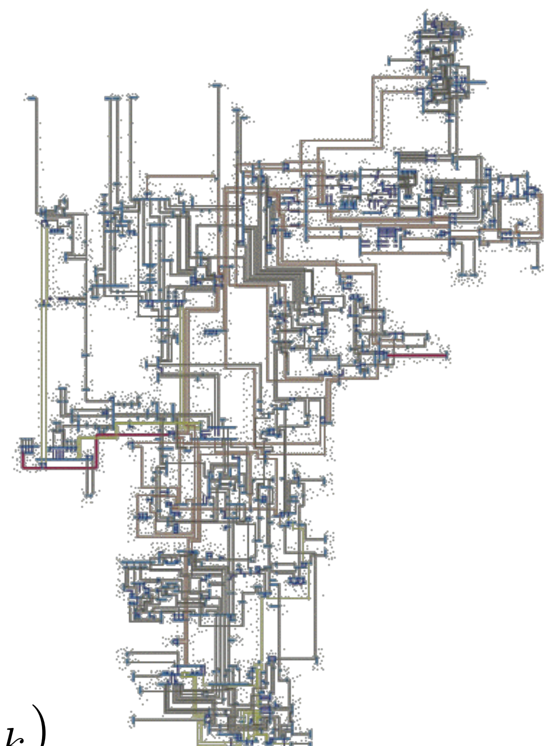
- Ohm's law

$$V_i = Z_i I_i$$

Relationship with power: The balance equations



Power Injection = Losses Admittance matrix



$$P_i + jQ_i = V_i I_i^* = V_i \sum_{j \in E} Y_{ij}^* V_j^*$$

AC Power Flow Model

$$P_i = \sum_{k=1}^{n} |V_i V_k| G_{ik} \cos(\theta_{ik}) + B_{ik} \sin(\theta_{ik})$$

$$Q_i = \sum_{k=1}^{n} |V_i V_k| G_{ik} \sin(\theta_{ik}) - B_{ik} \cos(\theta_{ik})$$

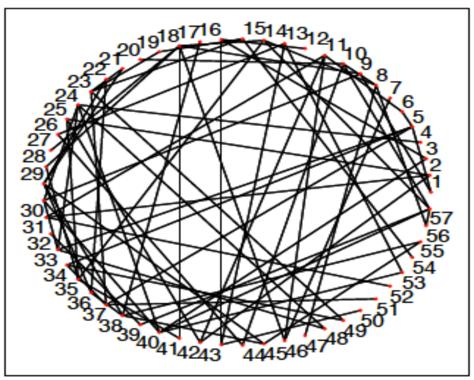
Precursors of our study

- Most of the literature has used real grids or reference models for testing
  - IEEE 30 57, 118 and 300 bus systems
  - Power systems test case archive
    - http://www.ee.washington.edu/research/pstca/
- Scalable models to grasp macroscopic trends
  - [Parashar and Thorp ’04] ring topology + “continuum model”
  - [Rosas-Casals, Valverde, Solé ’07] tree topology

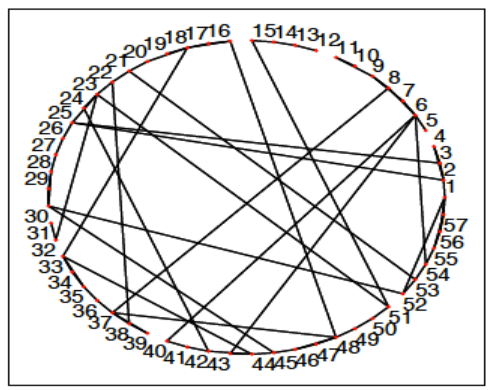
Random topology models

- '98 Watts and Strogatz, *Nature*
  - *Conjecture: Power Grids are small world networks*

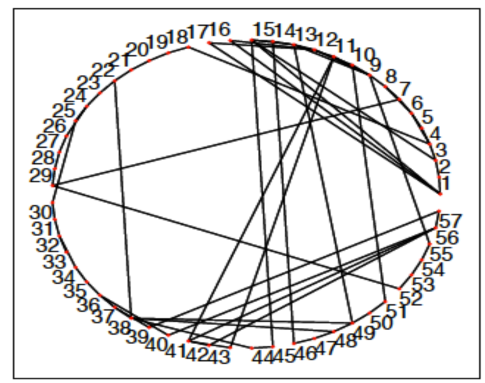
Erdos Reny



Small World



Power network



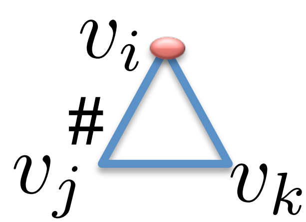
- Topological studies: [Newman '03], [Whitney & Alderson'06] [Wang, Rong,'09]
- Degree distribution: [Albert et al. '04],[Rosas-Casals et al. '07]

Small world → high clustering coefficient

- high average clustering coefficient of the sample power grid network examined

Definition of clustering coefficient

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i - 1)}$$



$$\forall \ (v_j, v_k) \in N_i, e_{jk} \in E$$

$(N_i = \text{ neighbors of } v_i)$

$(E = \text{edges})$

$(k_i = \text{degree of } i)$

Erdos Renyi

Grid	$C(G)$	$C(R)$
IEEE-30	0.2348	0.094253
IEEE-57	0.1222	0.048872
IEEE-118	0.1651	0.025931
IEEE-300	0.0856	0.009119
NYISO-2935	0.2134	0.001525
WSCC-4941	0.0801	0.000540

Our analysis

Degree distribution

- [Albert et al. ’04,Rosas-Casals’07] Geometric PDF
- Way to highlight:

  Probability Generating Function (PGF)

  - For a mixture model

$$G_k(z) = G_{k_1}(z) \ldots G_{k_p}(z)$$

Our analysis result

1. The degree distribution is a mixture of a truncated exponential and finite support random variable
2. The average degree vs. N is O(1)

Why the PGF?

- A finite support Probability Mass Function (PMF) is a finite order polynomial
  - We should see ‘zeros’ in the PGF

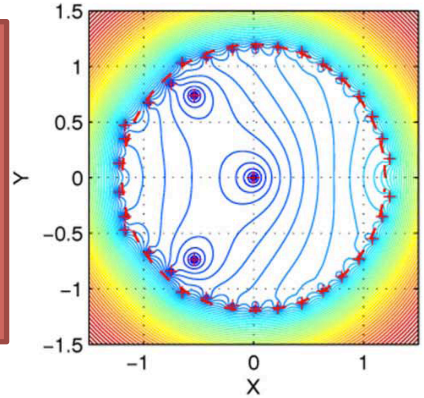
$$G_{\mathcal{D}}(z) = p_0 + p_1 z + \ldots + p_{k_t} z^{k_t}$$

- A purely geometric random variable is the reciprocal of a first order polynomial → ‘pole’
  - Impossible to observe, in practice a ‘clipped’ version

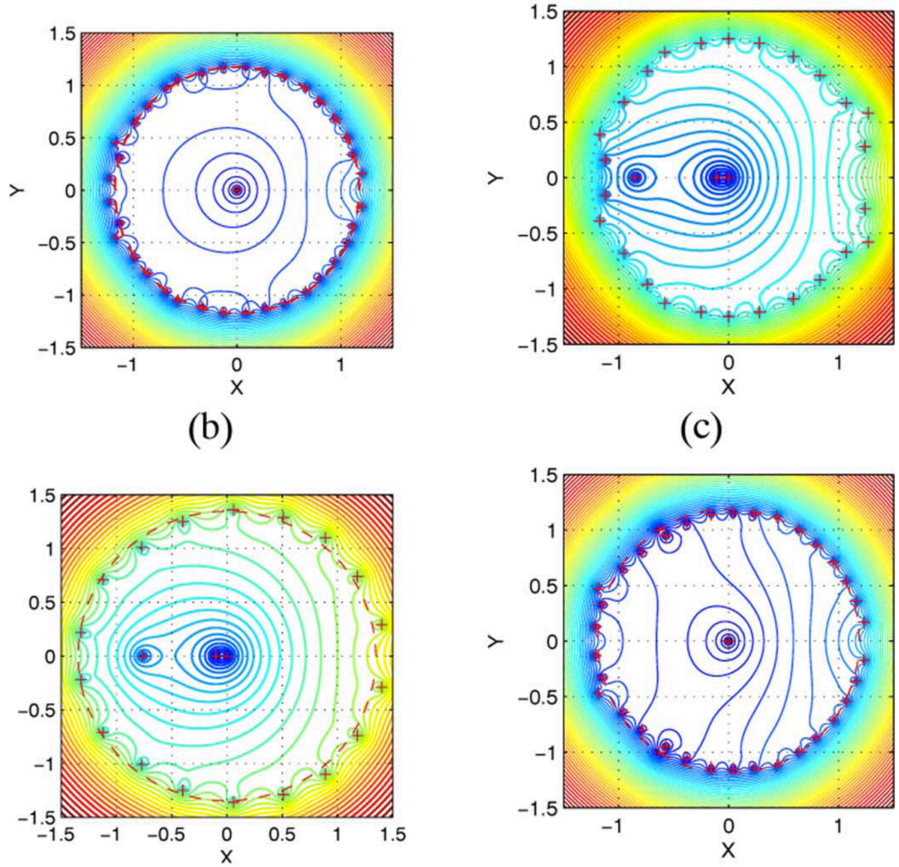
$$G_{\mathcal{G}}(z) \propto \frac{1 - [z(1-p)]^{k_{\max}+1}}{1-(1-p)z}$$

PGF NYSO data

(a) All buses
(b) Gen buses
(c) Load buses.
(d) Connection buses.
(e) Gen+Load buses.
The zeros are red '+'

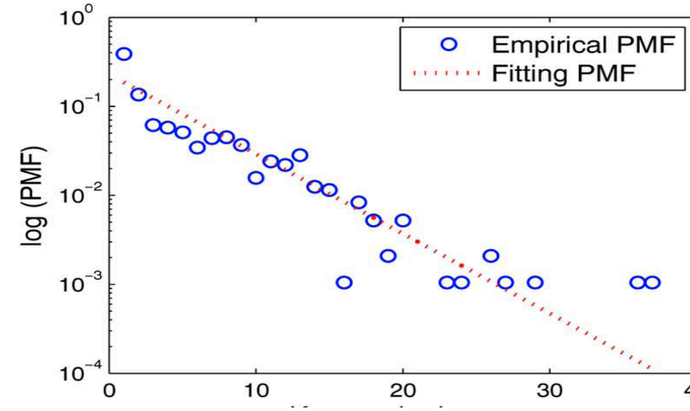


(a)

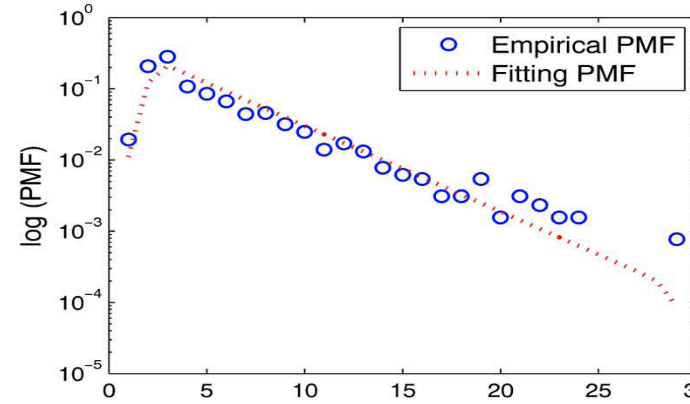


(b) (c)

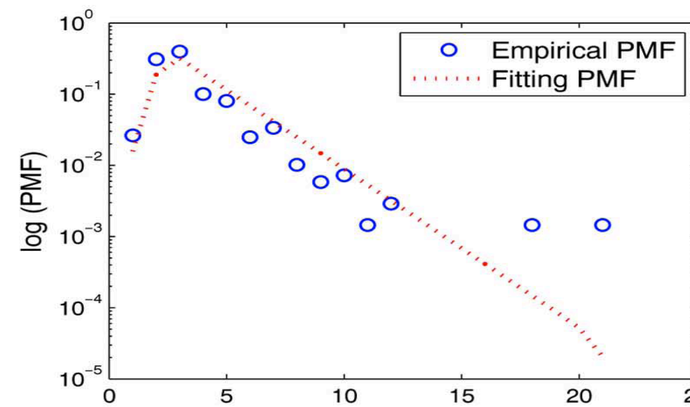
(d) (e)



Degree of Generator buses



Degree of Load buses



Degree of Connection buses

WSCC versus NYSO degree distribution

$$G_k(z) = G_{\mathcal{D}}(z) G_{\mathcal{G}}(z)$$

ESTIMATE COEFFICIENTS OF THE TRUNCATED GEOMETRIC AND THE IRREGULAR DISCRETE FOR THE NODE DEGREES IN THE NYISO AND WSCC SYSTEM

node groups	max($\underline{k}$)	p	k_{max}	k_t	$\{p_1, p_2, \cdots, p_{k_t}\}$
All	37	0.2269	34	3	0.4875, 0.2700, 0.2425
Gen	37	0.1863	36	1	1.000
Load	29	0.2423	26	3	0.0455, 0.4675, 0.4870
Conn	21	0.4006	18	3	0.0393, 0.4442, 0.5165
Gen+Load	37	0.2227	34	3	0.4645, 0.3385, 0.1970
All-WSCC	19	0.4084	16	3	0.3545, 0.4499, 0.1956

Vulnerability studies

- Fraction of nodes removal before breakdown
  - [Rosas-Casals et. al '07] $r = 1$ [Wang et al. '09]

$$f_c^{rand} = \left(1 - \ln \frac{f_c^{sel}}{r}\right) \quad f_c^{sel} = 1 - \frac{1}{<k> - 1}$$

The Theoretical versus the Empirical Critical Breakdown Thresholds

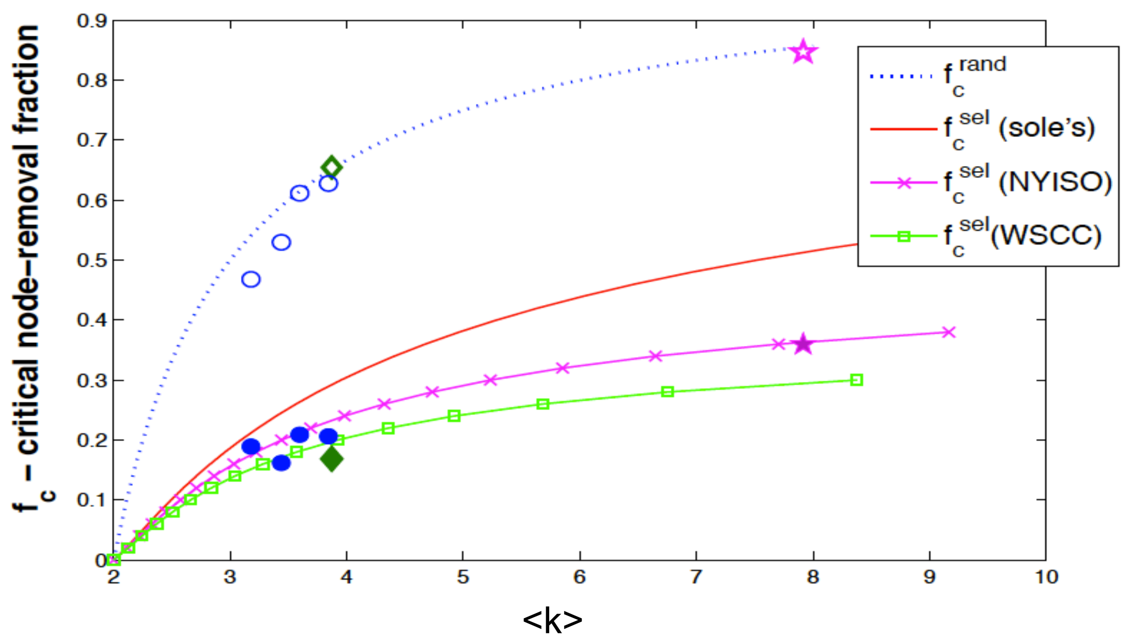
$r_{NYISO} = 1.4074$

$r_{WSCC} = 1.9690$

IEEE (circles), WSCC **(diamond), NYISO (star)**

Hollow - f_c^{rand}

Filled - f_c^{sel}



Small World conjecture

- Some evidence contradicting it
  - For a SW network with N nodes, to guarantee with high probability a connected network (no isolated component) the scaling laws for the average degree *<k> >> log N*
  - The average degree in power grids is ~ constant (3-4)

N: Number of nodes
m: number of lines
-
<k> Average Degree
<l> Average shortest path length
ρ Pearson Coefficient
r{k>k} Ratio of nodes with largest nodal degree

TOPOLOGICAL CHARACTERISTICS OF REAL-WORLD POWER NETWORKS

	(N, m)	$\langle l \rangle$	$\langle k \rangle$	ρ	$r\{k > \bar{k}\}$
IEEE-30	(30,41)	3.31	2.73	-0.0868	0.2333
IEEE-57	(57,78)	4.95	2.74	0.2432	0.2105
IEEE-118	(118,179)	6.31	3.03	-0.1526	0.3051
IEEE-300	(300, 409)	9.94	2.73	-0.2206	0.2367
NYISO	(2935,6567)	16.43	4.47	0.4593	0.1428
WSCC	(4941, 6594)	18.70	2.67	0.0035	0.2022

Average shortest path

- Observation: $< l > \approx 3 \log_{10}(N)$

N: Number of nodes
m: number of lines
-
<k> Average Degree
<l> Average shortest path length
ρ Pearson Coefficient
r{k>k} Ratio of nodes with largest nodal degree

TOPOLOGICAL CHARACTERISTICS OF REAL-WORLD POWER NETWORKS

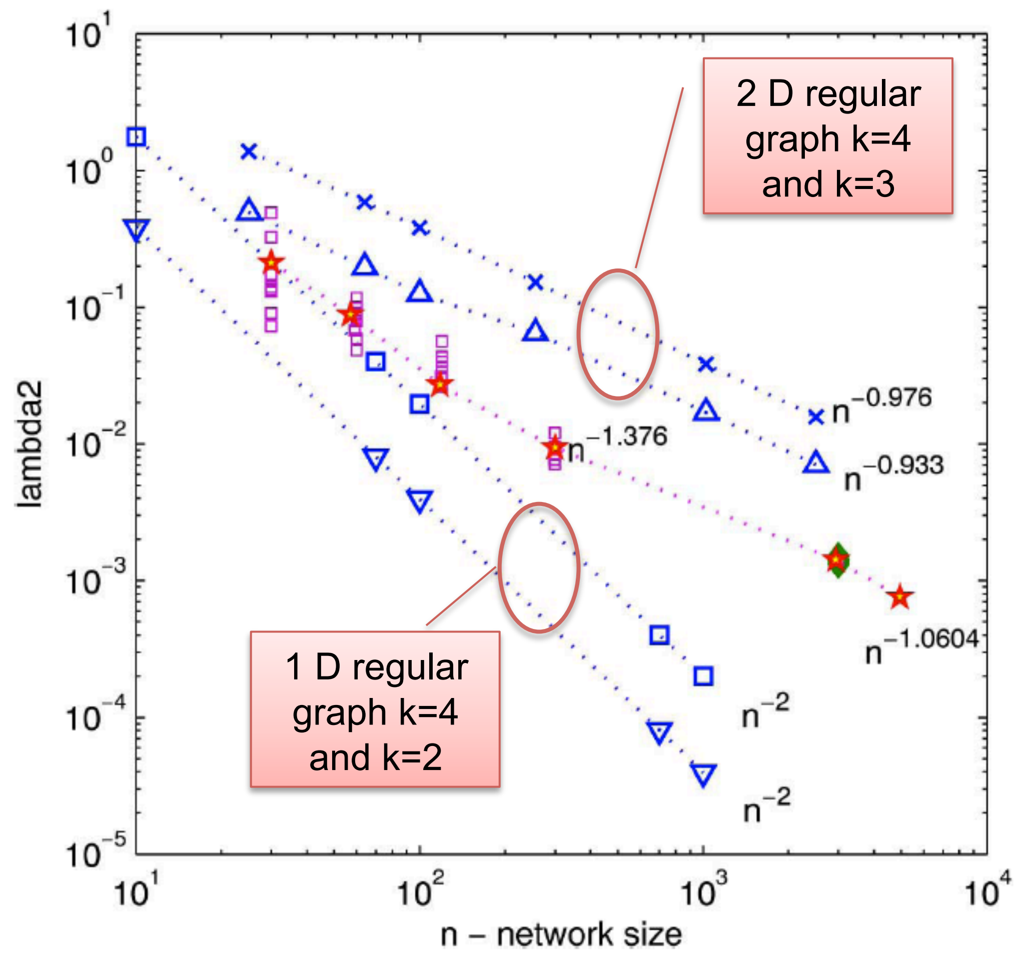
	(N, m)	$\langle l \rangle$	$\langle k \rangle$	ρ	$r\{k > \bar{k}\}$
IEEE-30	(30,41)	3.31	2.73	-0.0868	0.2333
IEEE-57	(57,78)	4.95	2.74	0.2432	0.2105
IEEE-118	(118,179)	6.31	3.03	-0.1526	0.3051
IEEE-300	(300, 409)	9.94	2.73	-0.2206	0.2367
NYISO	(2935,6567)	16.43	4.47	0.4593	0.1428
WSCC	(4941, 6594)	18.70	2.67	0.0035	0.2022

- Not bad to overlay communications with the lines – relatively short distance

Algebraic connectivity

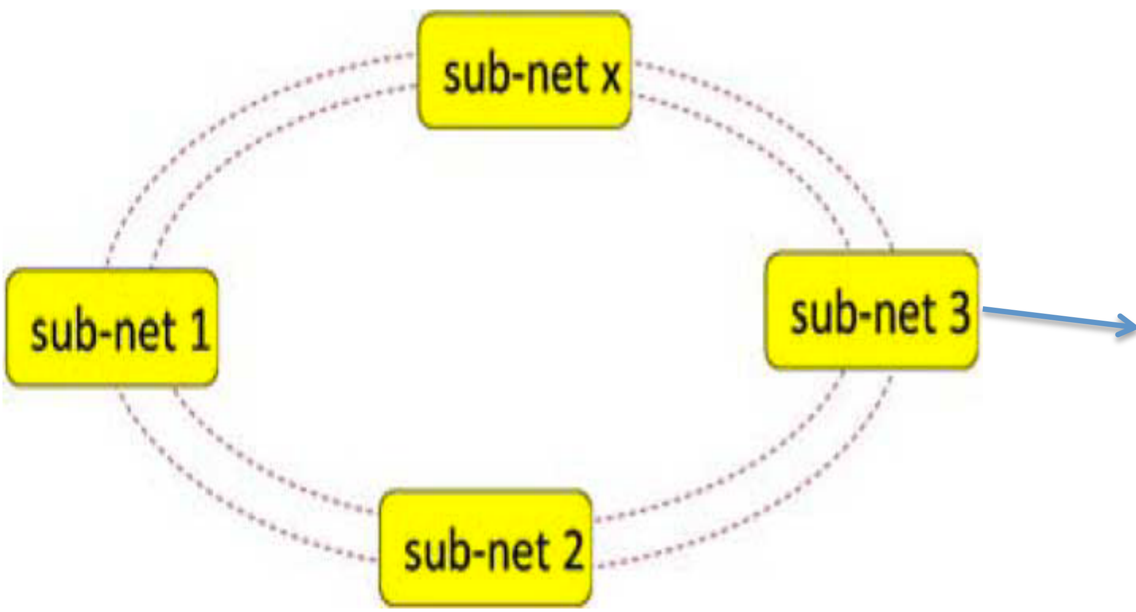
- Graph Laplacian second smallest eigenvalue
- Values shown in ☆

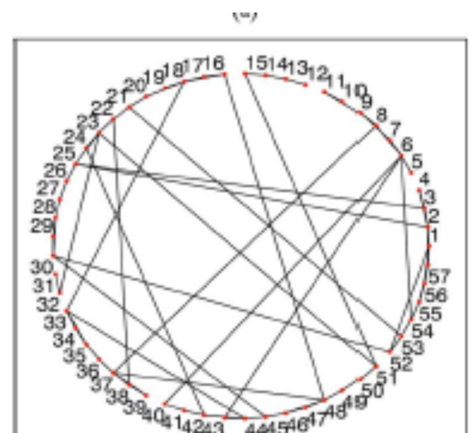
	$\lambda_2(L)$
IEEE-30	0.21213
IEEE-57	0.088223
IEEE-118	0.027132
IEEE-300	0.0093838
NYISO-2935	0.0014215
WSCC-4941	0.00075921



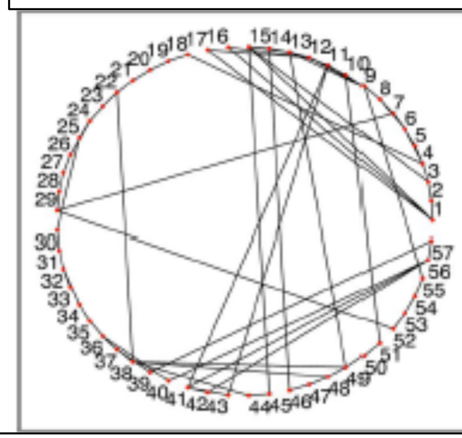
Plausible topology

- The model that matches this trend is what we call Nested-Small-world graph
  - IEEE → SW subnet 30; NYSO & WSCC → SW sub-net 300





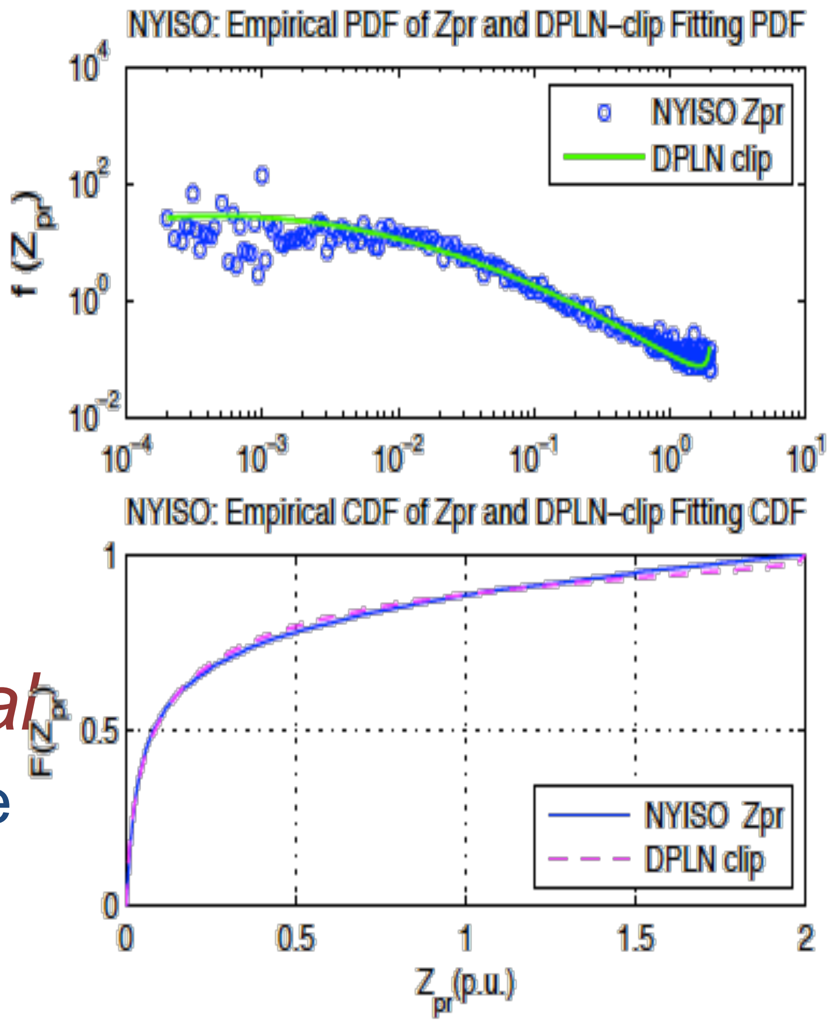
SW: independent rewiring



IEEE 300: Correlated rewiring

Impedance distribution

- Absolute values of the impedances
  $Z_{pr} = R + jX \approx jX$
- Prevailingly <u>*heavy tailed distributions*</u>
- NYSO best fit → *clipped Double Pareto Log-normal*
  - Did not pass KS test but was the closest to pass it



Distributions comparison

DISTRIBUTION FITTING FOR LINE IMPEDANCES

System	Fitting Distribution	ML Parameter Estimates (*alph*=0.05)
IEEE-30	$\Gamma(x\|a,b)$	$a = 2.14687$ $b = 0.10191$
IEEE-118	$\Gamma(x\|a,b)$	$a = 1.88734$ $b = 0.05856$
IEEE-57	$gp(x\|k,\sigma,\theta)$	$k = 0.33941$ $\sigma = 0.16963$ $\theta = 0.16963,$
IEEE-300	$gp(x\|k,\sigma,\theta)$	$k = 0.45019$ $\sigma = 0.07486$ $\theta = 0.00046,$
NYISO-2935	$logn_{clip}(x\|\mu,\sigma,Z_{\max})$	$\mu = -2.37419$ $\sigma = 2.08285$ $Z_{max} = 1.9977$
	$dPlN_{clip}(x\|\alpha,\beta,\mu,\sigma,Z_{\max})$	$\alpha = 44.25000$ $\beta = 44.30000$ $\mu = -2.37420$ $\sigma = 2.082600$ $Z_{max} = 1.9977$

Gamma:

$$\Gamma(x \mid a, b) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{x/b},$$

Generalized Pareto (GP):

$$gp(x \mid k, \sigma, \theta) = \left(\frac{1}{\sigma}\right)\left(1 + k\frac{(x-\theta)}{\sigma}\right)^{-1-(1/k)}.$$

Lognormal:

$$\log n(x \mid \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(\log x - u)^2}{2\sigma^2}}.$$

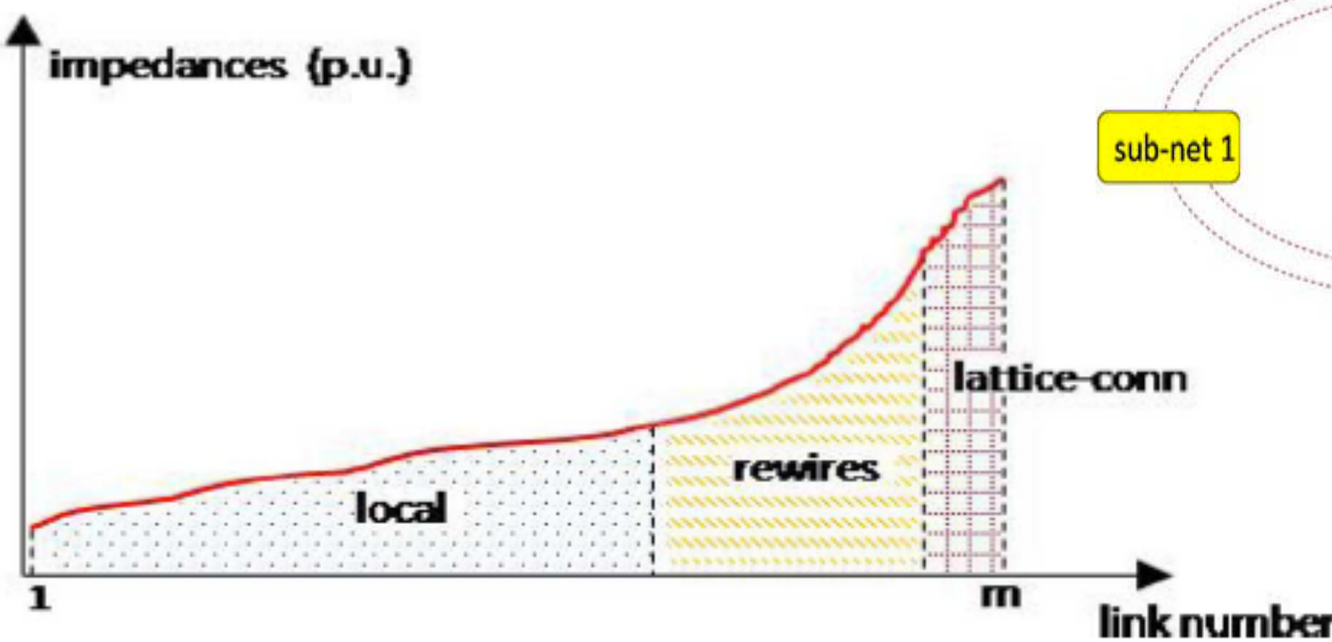
DPLN:

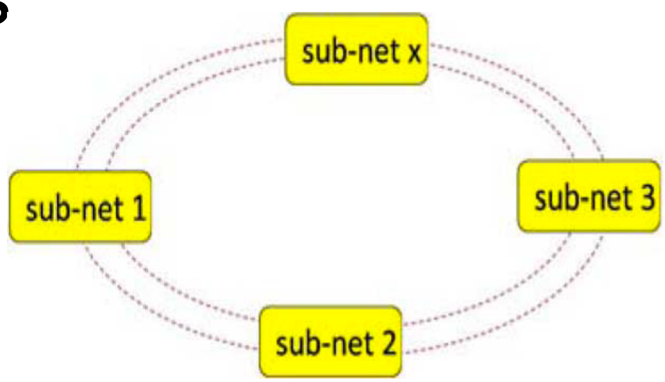
$$\mathrm{DPLN}(x|\alpha,\beta,\mu,\sigma) = \frac{\alpha\beta}{\alpha+\beta}\Bigg[A(\alpha,\mu,\sigma)x^{-\alpha-1}\Phi\left(\frac{\log x - \mu - \alpha\sigma^2}{\sigma}\right) + A(-\beta,\mu,\sigma)x^{\beta-1}\Phi^c\left(\frac{\log x - \mu + \beta\sigma^2}{\sigma}\right)\Bigg].$$

where $A(\theta, \mu, \sigma) = e^{(\theta\mu + \theta^2\sigma^2/2)}$.

Impedance attribution

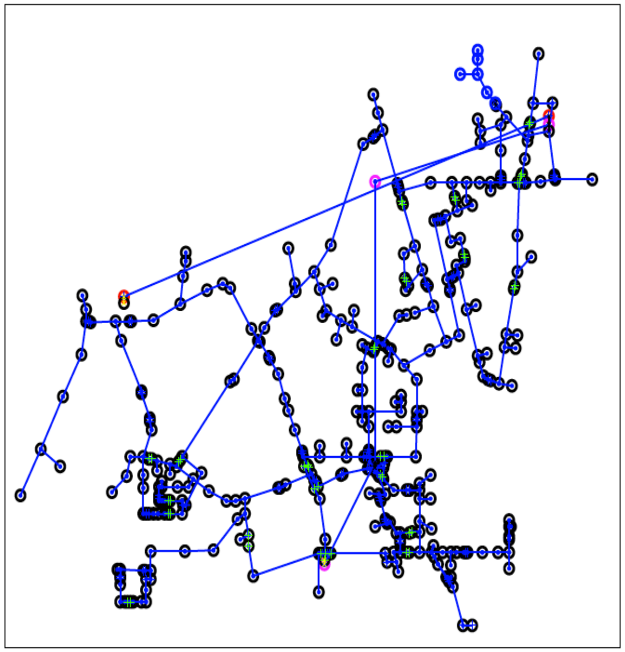
- Impedance grows with distance
- Conjecture: local → short; rewires → medium; lattice connections → long lines





396-node Medium Voltage distribution network

- US distribution utility
  - The power supply from the 115 kV-34.5 kV step-down substation.
  - Most nodes or buses in the network are 12.47 kV (>95%), and only a small number of them are 34.5 kV or 4.8 kV.



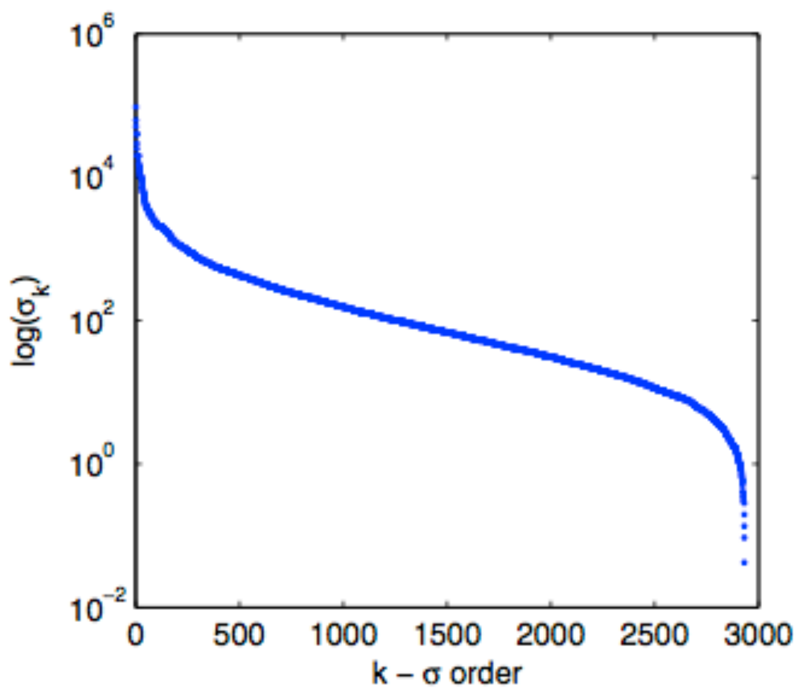
	(N, m)	$\langle k \rangle$	$\langle l \rangle$	ρ	$\lambda_2(L)$	$C(G)$
IEEE-300	(300, 409)	2.73	9.94	-0.2206	0.0094	0.0856
WSCC	(4941, 6594)	2.67	18.70	0.0035	0.00076	0.0801
396-node MV-Distr	(396, 420)	2.12	21.10	-0.2257	0.00030	0

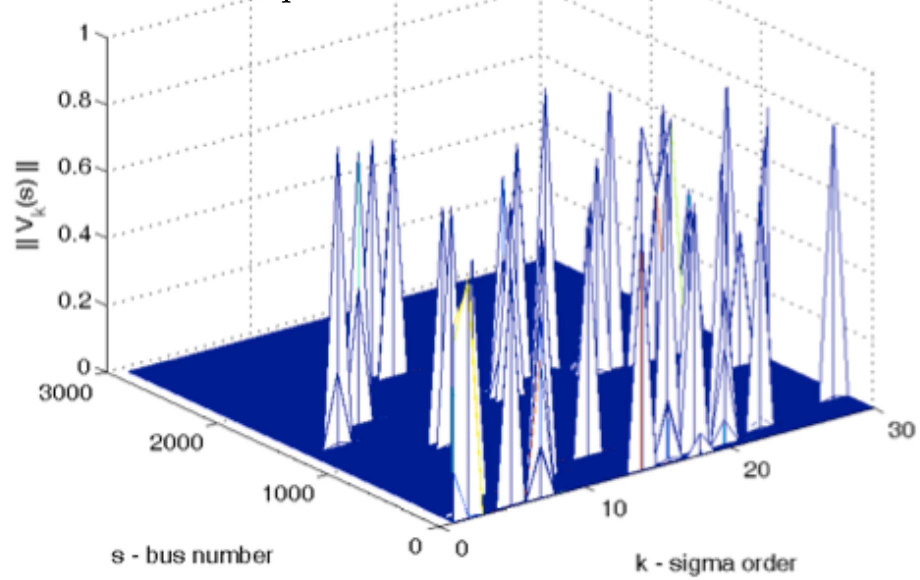
Interesting facts relating to CPS

Sparse principal eigenvectors

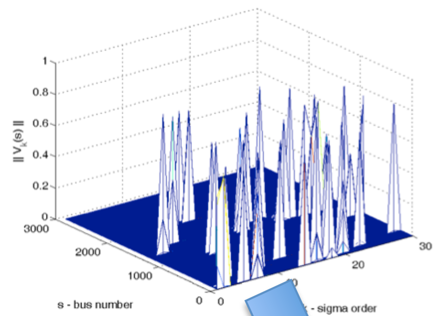
- We have found that the $\mathbf{Y} = \mathbf{U \Lambda U^H}$ has sparse eigenvalues with sparse principal components

$$\mathbf{U} = [\underbrace{\mathbf{u}_1, \cdots, \mathbf{u}_K}_{\triangleq \text{Principal}}, \underbrace{\mathbf{u}_{K+1}, \cdots, \mathbf{u}_N}_{\triangleq \text{Minor}}]$$

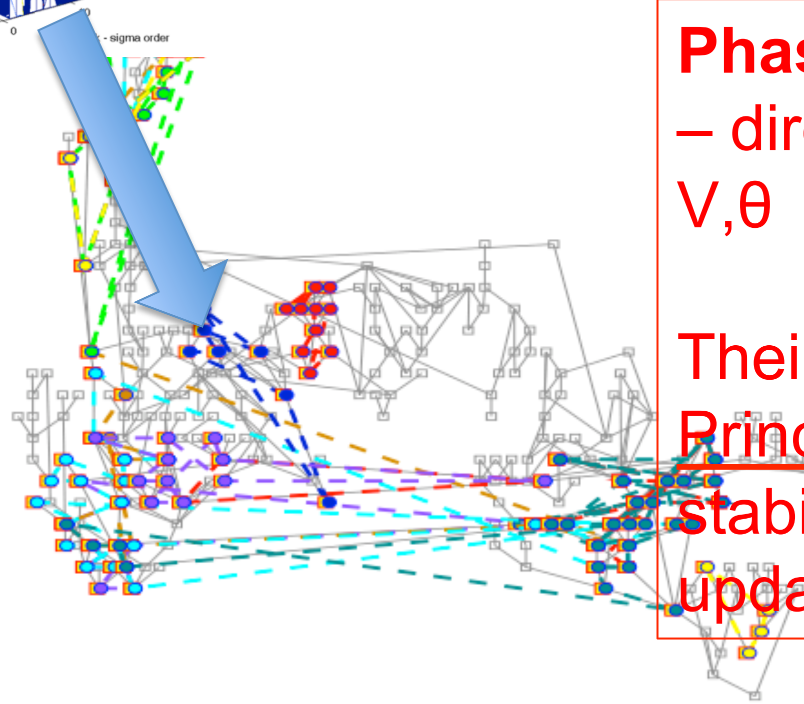




Sensor placement for robust state estimation



$$\text{Clique}_k(\gamma) = \left\{ i : [\mathbf{u}_k] \geq \frac{\gamma}{\sqrt{N}} \right\}$$



Phasor Measurement Units – directly measure the state V,θ

Their placement on the K Principal Cliques best stabilizes the Gauss-Newton update for State Estimation

Geometrical insights from AC to DC Power flow

- Admittance matrix $\mathbf{Y} = \mathbf{G} + \mathrm{i}\mathbf{B}'$

conductance | susceptance

- Susceptance >> Conductance
- Small angle difference $|\theta_i - \theta_j|, V_i \approx 1$
- DC Power Flow Model approximaion

$$\mathbf{P} = \mathbf{B}\boldsymbol{\theta}, \quad \mathbf{B} = -\mathbf{B}' \text{ with shunt removed}$$

Power injection | Phase angle

Impact on Power Injections

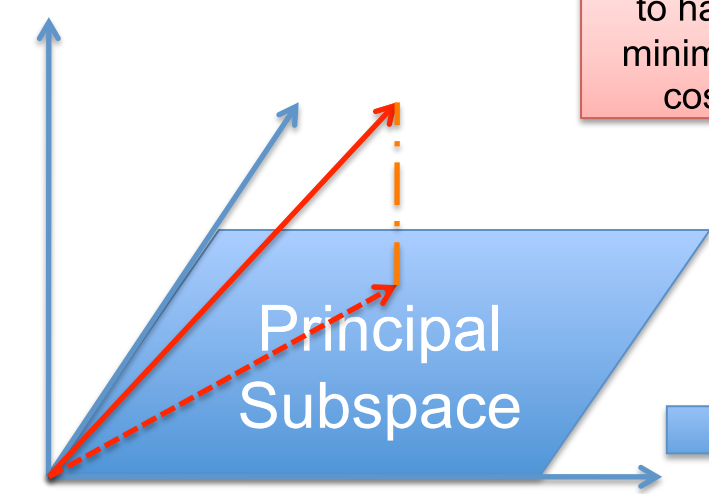
- Low rank approximation $\mathbf{B} = \sum_{k=1}^{K} \lambda_k \mathbf{u}_k \mathbf{u}_k^T + \mathcal{O}(\lambda_{K+1})$

$$\mathbf{B}\boldsymbol{\theta} = \mathbf{P}_G - \mathbf{P}_L$$

$$\sum_{k=1}^{K} \rho_k \lambda_k \mathbf{u}_k \approx \mathbf{P}_G - \mathbf{P}_L, \quad \rho_k = \mathbf{u}_k^T \boldsymbol{\theta}.$$

Dispatched to have minimum cost

The balance constraint in the Optimal Power Flow Economic dispatch will tend to line up the injection with the principal subspace

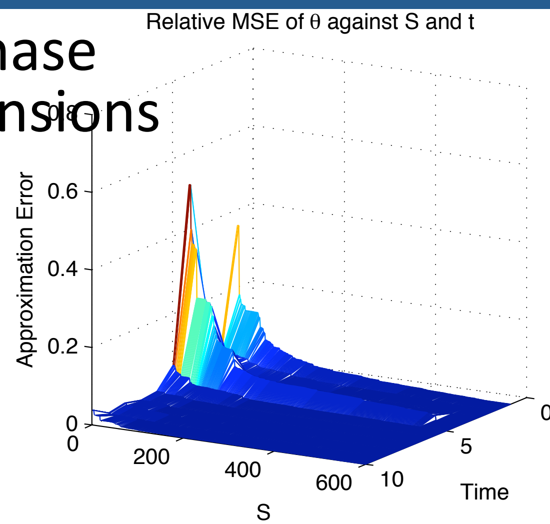


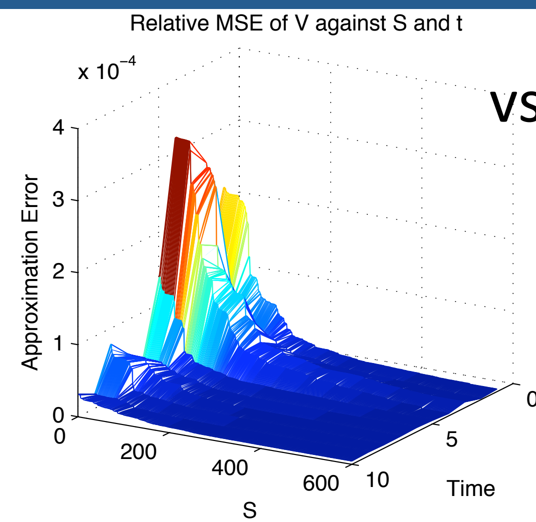
Load fluctuation → OPF generation adjustment

The sensitivity analysis suggests that greatest variations are in the least significant subspace component

Low Dimensional Representation

MSE of phase
vs # of dimensions

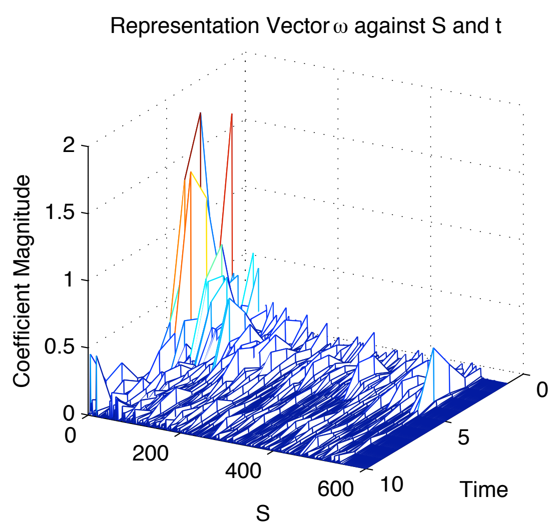


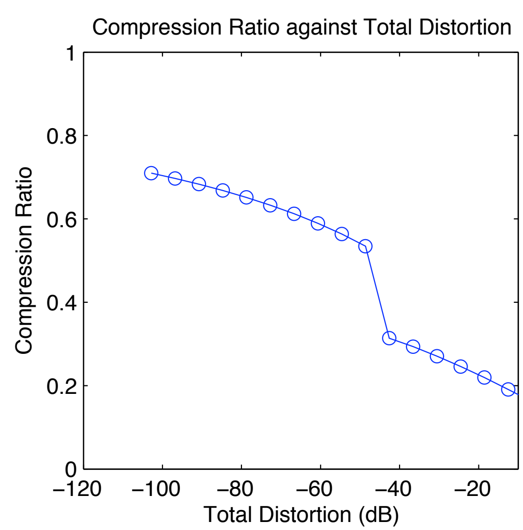


MSE of voltage
vs # of dimensions

IEEE-300 bus system

10 snapshots





Conclusions

- The admittance matrix of power grids has peculiar features that follow clear statistical trends
- The analysis can help grasping macroscopic phenomena
- It is critical to understand how to design it better
- It is critical incorporate flexibility and adaptability via monitoring and control

References

1) *Zhifang Wang, Anna Scaglione, and Robert J. Thomas* "Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks", IEEE Transactions on Smart Grid, Vol. 1, No. 1. (June 2010), pp. 28-39.

2) *Zhifang Wang, Anna Scaglione, and Robert J. Thomas*, "The Node Degree Distribution in Power Grid and Its Topology Robustness under Random and Selective Node Removals" IEEE International Workshop on Smart Grid Communications, Cape Town, South Africa, May 2010.

3) Zhifang Wang; Scaglione, A.; Thomas, R.J.; , "Compressing Electrical Power Grids," *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on* , vol., no., pp.13-18, 4-6 Oct. 2010

4) Zhifang Wang; Scaglione, A.; Thomas, R.J.; , "Electrical centrality measures for electric power grid vulnerability analysis," *Decision and Control (CDC), 2010 49th IEEE Conference on* , vol., no., pp.5792-5797, 15-17 Dec. 2010

5) Galli, S.; Scaglione, A.; Zhifang Wang; , "For the Grid and Through the Grid: The Role of Power Line Communications in the Smart Grid," *Proceedings of the IEEE* , vol.99, no.6, pp.998-1027, June 2011