A Study on the Sensitivity Matrix in Power System State Estimation by Using Sparse Principal Component Analysis

Adam Molin[1] Henrik Sandberg[1] Magnus Johansson[2]

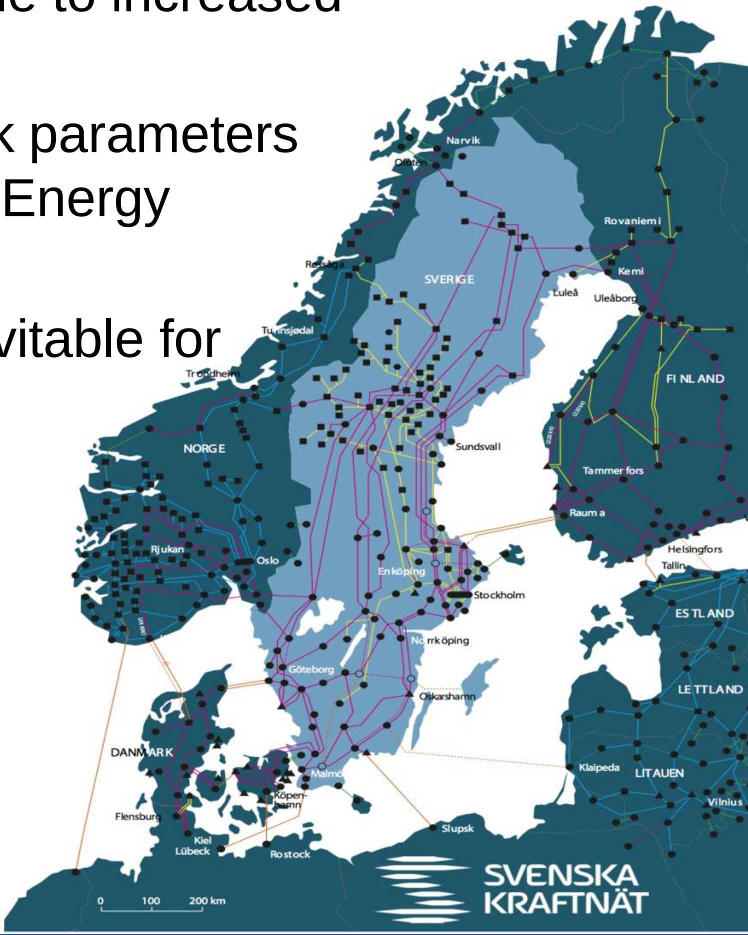
[1]KTH Royal Institute of Technology – Department of Automatic Control

[2]Svenska Kraftnät (Swedish National Grid)

CDC'16 Las Vegas Dec 12, 2016

Motivation

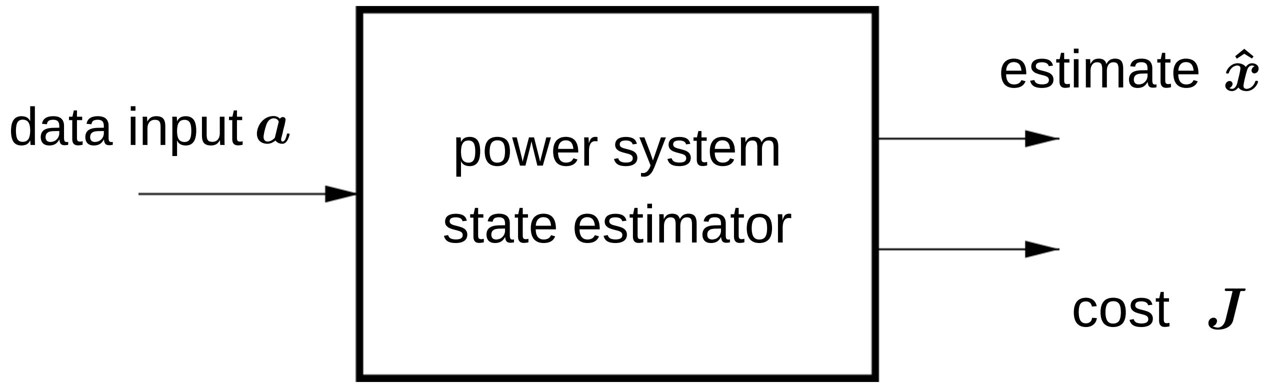
- Grid more frequently at operational limits due to increased power transfers
- Uncertainties in measurements and network parameters always present in the state estimator of the Energy Management System (EMS)
- Better understanding of the data quality inevitable for driving the power system closer to its limits



Objective

Identify the most relevant data components (measurements & line parameters) for power system state estimation

Background



Data input a to power system state estimator:

$$\boldsymbol{a} = \{\text{measurements } \boldsymbol{z}, \text{branch parameter } \boldsymbol{Y}, \text{topology}\}$$

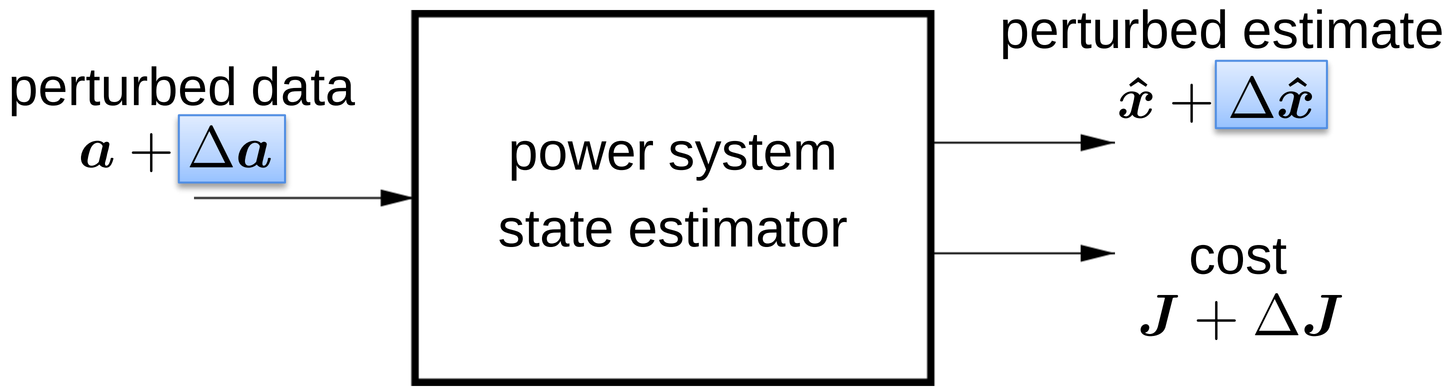
Network model (power flow equations):

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{Y}) + \boldsymbol{e} \qquad e \sim \mathcal{N}(0, \boldsymbol{R}_{\boldsymbol{z}})$$

Weighted least squares:

$$J(\hat{\boldsymbol{x}}, \boldsymbol{a}) = \frac{1}{2} \sum_{\ell=1}^{L} \frac{1}{\sigma_\ell^2} \left(z_\ell - h_\ell(\hat{\boldsymbol{x}}, \boldsymbol{Y})\right)^2$$

General approach



Data input $\boldsymbol{a} = \{\text{measurements } \boldsymbol{z}, \text{branch parameter } \boldsymbol{Y}, \cancel{\text{topology}}\}$

Sensitivity Analysis with Sparsity Constraint

Find the set D of data inputs with $|D| < t$, whose joint perturbations impact the state estimate most.

Methodology: Sparse PCA on (a variant of) sensitivity matrix

$$\boxed{\boldsymbol{S} = \frac{d\hat{\boldsymbol{x}}}{d\boldsymbol{a}}}$$

Computation of sensitivity matrix

Method of feasible directions [1]: *Compute the modifications (directions) of the optimal state estimate when there is a perturbation of the input data, such that the linearized necessary conditions are still valid.*

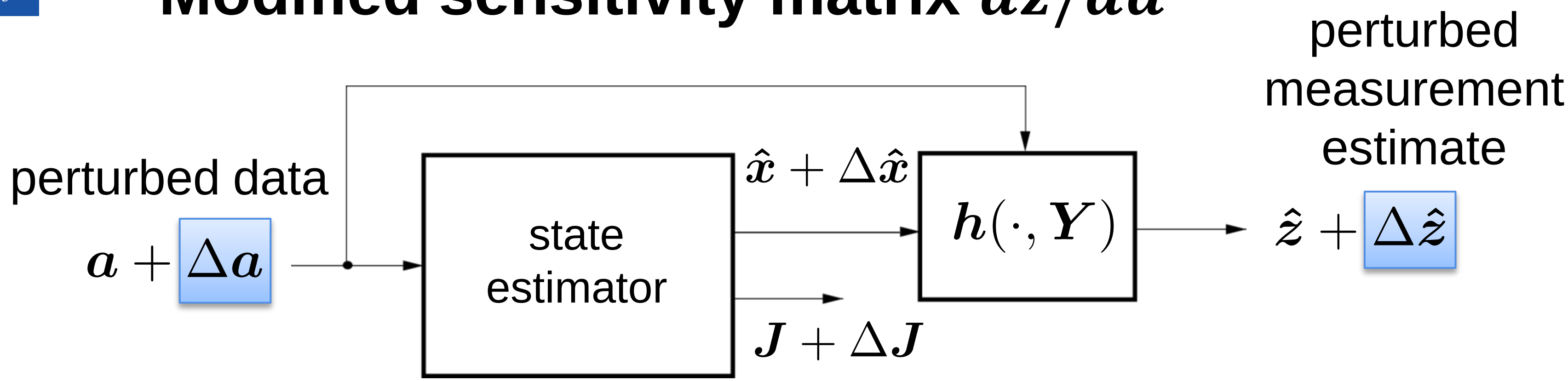
Necessary condition for optimality $\frac{\partial J}{\partial \hat{\boldsymbol{x}}}\Big|_{\hat{x}=\hat{x}^*} = \mathbf{0}$

Sensitivity equation $\frac{\partial^2 J(\hat{\boldsymbol{x}}, \boldsymbol{a})}{\partial \hat{\boldsymbol{x}}^2} d\boldsymbol{x} + \frac{\partial^2 J(\hat{\boldsymbol{x}}, \boldsymbol{a})}{\partial \hat{\boldsymbol{x}} \partial \boldsymbol{a}} d\boldsymbol{a} = 0$

Sensitivity matrix: $\boldsymbol{S}_{\hat{\boldsymbol{x}}} = \frac{d\hat{\boldsymbol{x}}}{d\boldsymbol{a}}\Big|_{\hat{\boldsymbol{x}}=\hat{x}^*} = -\left(\frac{\partial^2 J}{\partial \hat{\boldsymbol{x}}^2}\right)^{-1} \frac{\partial^2 J}{\partial \hat{\boldsymbol{x}} \partial \boldsymbol{a}}$

[1] R. Minguez, A.J. Conejo, “State estimation sensitivity analysis,” Power Systems, IEEE Trans. on, 2007

Modified sensitivity matrix $d\hat{\boldsymbol{z}}/d\boldsymbol{a}$



$$\boldsymbol{S} = \frac{d\hat{\boldsymbol{z}}}{d\boldsymbol{a}}\Big|^{\text{normalized}}_{\hat{\boldsymbol{x}}=\hat{x}^*} = \boldsymbol{R}_{\boldsymbol{z}}^{-\frac{1}{2}} \left(\frac{\partial \hat{\boldsymbol{z}}}{\partial \boldsymbol{x}} \boldsymbol{S}_{\hat{\boldsymbol{x}}} + \frac{\partial \hat{\boldsymbol{z}}}{\partial \boldsymbol{a}} \right) \boldsymbol{R}_{\boldsymbol{a}}^{\frac{1}{2}}$$

- Facilitates interpretability ($\frac{\partial \hat{\boldsymbol{z}}}{\partial \boldsymbol{z}}$ common in power sys.)
- Normalization with covariance $\boldsymbol{R}_{\boldsymbol{z}}^{-\frac{1}{2}}(\cdot)\boldsymbol{R}_{\boldsymbol{a}}^{\frac{1}{2}}$ enables the comparison of entries

Sparse Principal Component Analysis on $\boldsymbol{S}$

Conduct sparse PCA on sensitivity matrix $\boldsymbol{S}$ [2]:

Maximize impact on output

Normalized inputs

$$\max_{\boldsymbol{v}} \boldsymbol{v}^\top (\boldsymbol{S}^\top \boldsymbol{S}) \boldsymbol{v} \quad \text{s.t.} \begin{cases} \boldsymbol{v}^\top \boldsymbol{v} & = 1, \\ \|\boldsymbol{v}\|_0 & \leq t, \end{cases}$$

Constraint on number of inputs

Deflate matrix: $\boldsymbol{S} \leftarrow \boldsymbol{S} - \boldsymbol{S}\boldsymbol{v}\boldsymbol{v}^\top$

Rerun after deflation

Loading vector encodes relevant inputs

Non-zero entries in $\boldsymbol{v}$ are interpreted as crucial data inputs for state estimation

[2] A. d'Aspremont, et al., "A direct formulation for sparse PCA using semidef. program.," *SIAM review*, 2007

Algorithms for sparse PCA

1. SDP-SPCA:

Semi-definite programming relaxation technique [2]

2. Soft thresholding:

ℓ_1-regularized PCA inspired by LASSO [3]

[2] A. d'Aspremont, *et al*., "A direct formulation for sparse PCA using semidef. program.," *SIAM review*, 2007
[3] H. Zou, et al., "Sparse principal component analysis," *Journal of computational and graphical statistics*, 2006.

Two Extreme Cases

Take $\|v\|_0 = 1$ **and replace objective with** $\|Sv\|_\infty$

➔returns entry with maximum absolute value.

– Only shows impact of single input/single output

Or

Omit cardinality constraint $\|v\|_0 \leq t$

➔ returns singular value (SVD)

– upper bound

– non-structured

Structure of spectral properties of $\boldsymbol{S}$

Calculation of $\boldsymbol{S}$ reveals the form:

$$\boxed{\boldsymbol{S} = \begin{bmatrix} \Pi & (\boldsymbol{I}_L - \Pi)\boldsymbol{L}_{\boldsymbol{Y}} \end{bmatrix}}$$

with

$$\Pi = \tilde{\boldsymbol{H}} \left(\tilde{\boldsymbol{H}}^\top \tilde{\boldsymbol{H}} \right)^{-1} \tilde{\boldsymbol{H}}^\top \quad \text{(projection matrix)}$$

$$\tilde{\boldsymbol{H}} = \boldsymbol{R}_{\boldsymbol{z}}^{-\frac{1}{2}} \boldsymbol{H}$$

$$\boldsymbol{L}_{\boldsymbol{Y}} = \boldsymbol{R}_{\boldsymbol{z}}^{-\frac{1}{2}} \boldsymbol{H}_{\boldsymbol{Y}} \boldsymbol{R}_{\boldsymbol{Y}}^{\frac{1}{2}}$$

and the Jacobians

$$\boldsymbol{H} = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\Big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}^*} \qquad \boldsymbol{H}_{\boldsymbol{Y}} = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{Y}}\Big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}^*}$$

Structure of spectral properties of $\boldsymbol{S}$

From the structure of $\boldsymbol{S}$, we have

Proposition 1:
Let $\boldsymbol{S} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top$ be the SVD with $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L]$.
Assume that $(\boldsymbol{I}_L - \Pi)\boldsymbol{L}_{\boldsymbol{Y}}$ has no singular value at 1.
Then, for any positive singular value, we have $\boldsymbol{v}_\ell \in \mathcal{V}^0$.

Same holds for

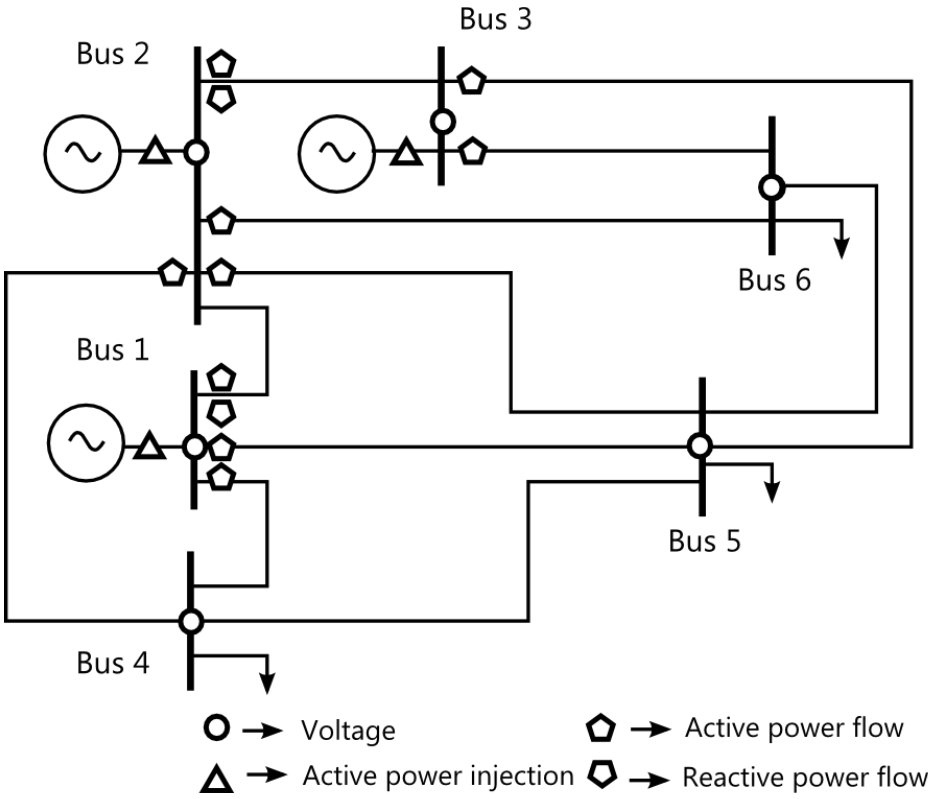
- Cardinality constrained PCA
- Soft thresholding algorithm

Target set for loading vectors: (either belong to measurement inputs or to line parameters)

$$\mathcal{V}^0 = \{\boldsymbol{v} \in \mathbb{R}^M \,|\boldsymbol{v} \in \begin{bmatrix} \boldsymbol{v}_{\boldsymbol{z}} \\ \boldsymbol{0} \end{bmatrix}, \boldsymbol{v}_{\boldsymbol{z}} \in \mathbb{R}^L \vee \boldsymbol{v} \in \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{v}_{\boldsymbol{Y}} \end{bmatrix}, \boldsymbol{v}_{\boldsymbol{Y}} \in \mathbb{R}^{M-L}\}.$$

Numerical illustration of approach

6-bus system



- Input data:

  20 measurements

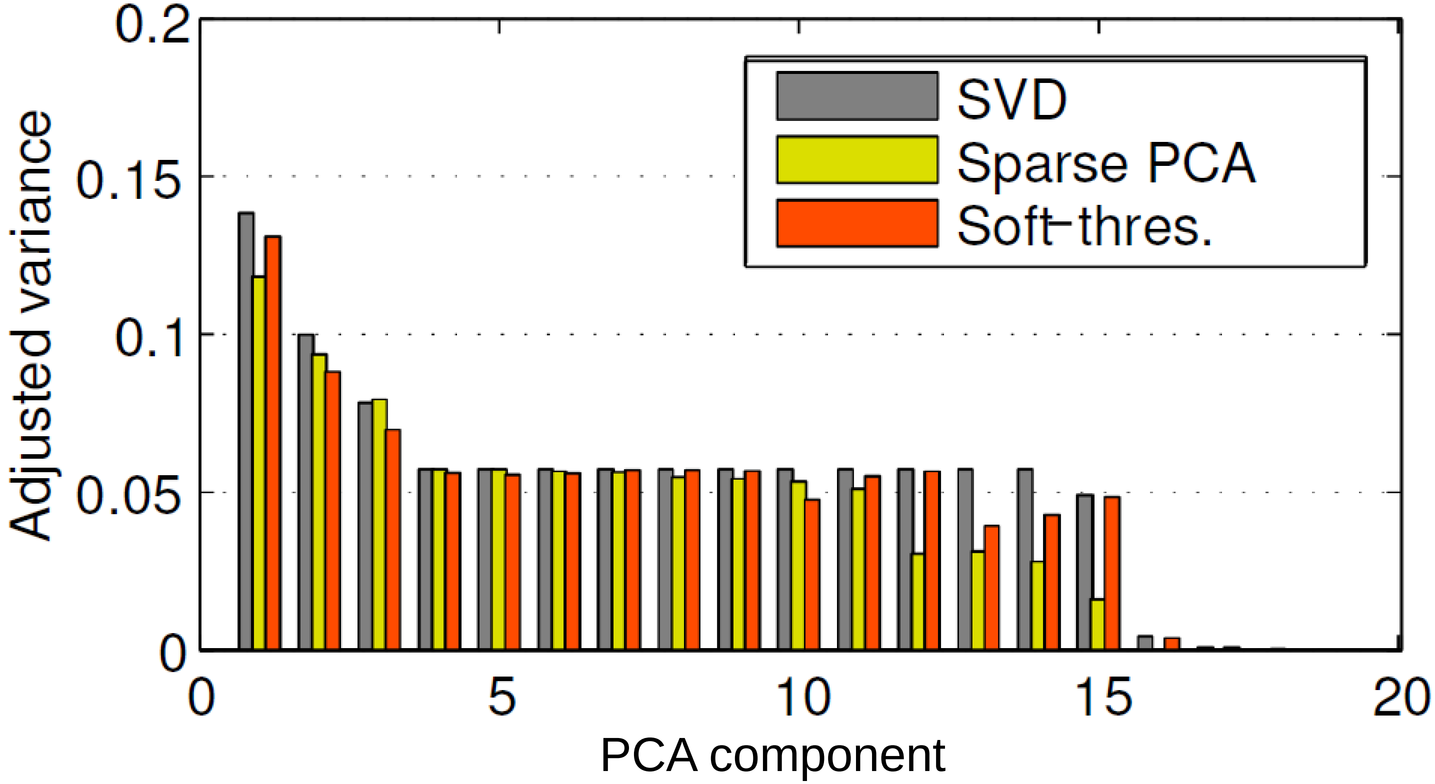
  2 x 11 branch parameters

  (resistance & reactance)

Sensitivity matrix: 20 x 42 – matrix

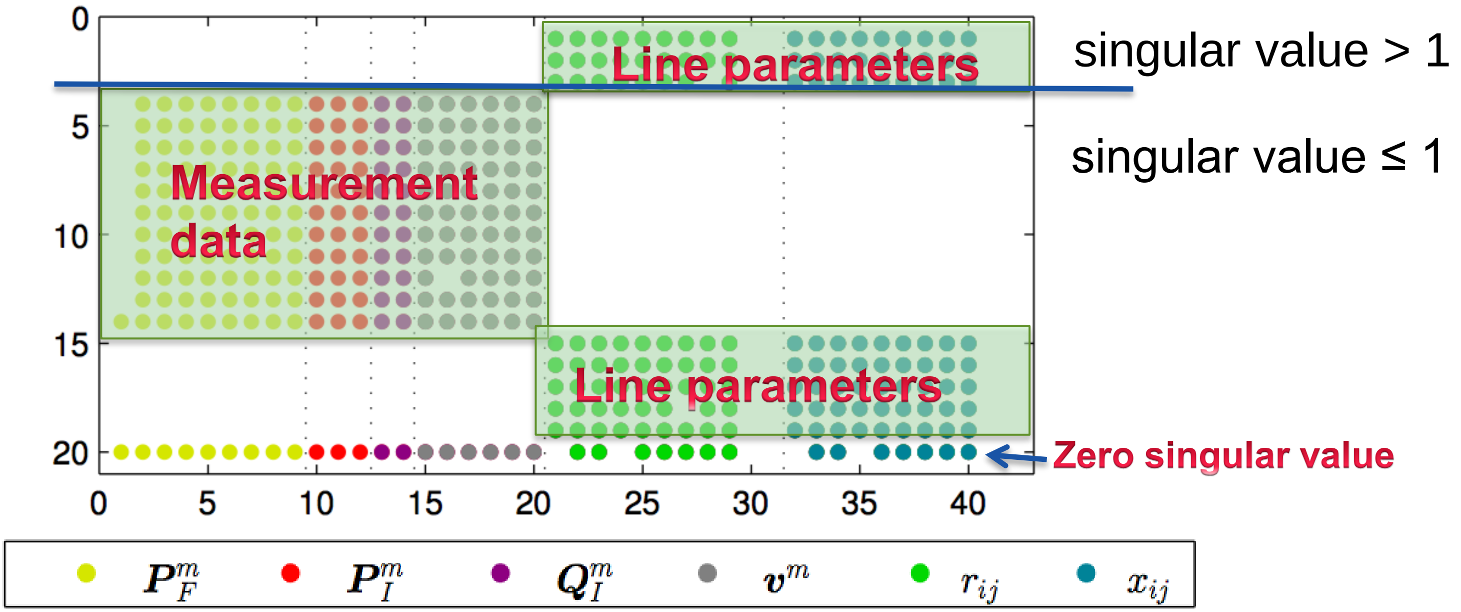
Spectral profile of sparse PCA

Adjusted variance (squared singular values, add up to 1)



Numerical illustration of approach

Results from SVD: Sparsity pattern of V^T

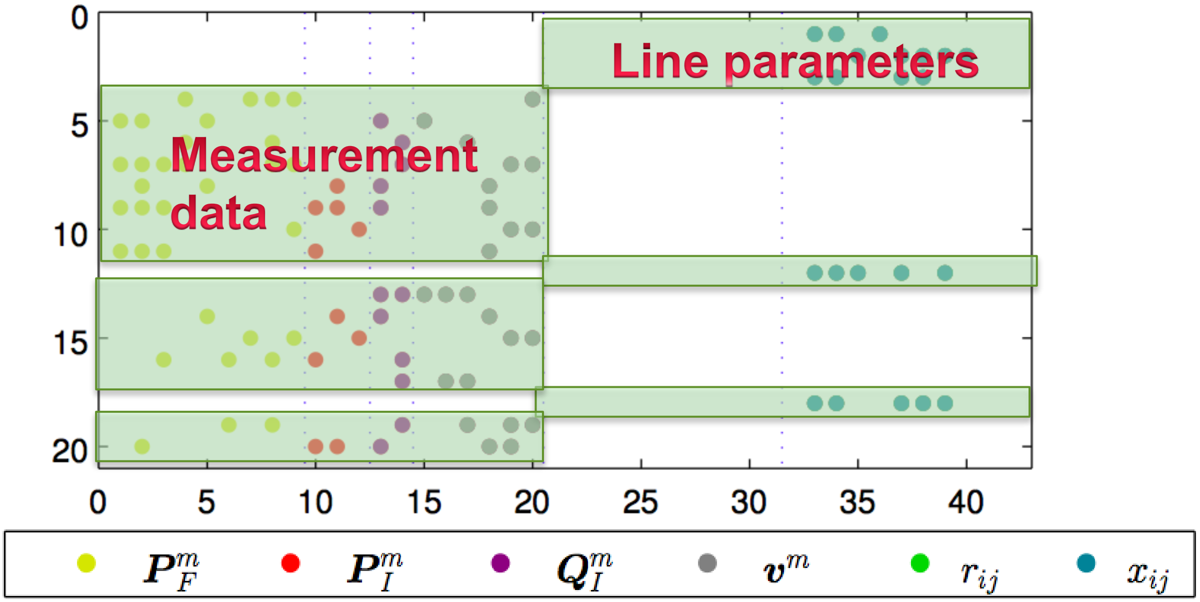


➔ Separation into measurement & line parameter inputs (Structure as theoretically proven)

➔ Otherwise dense structure

Numerical illustration of approach

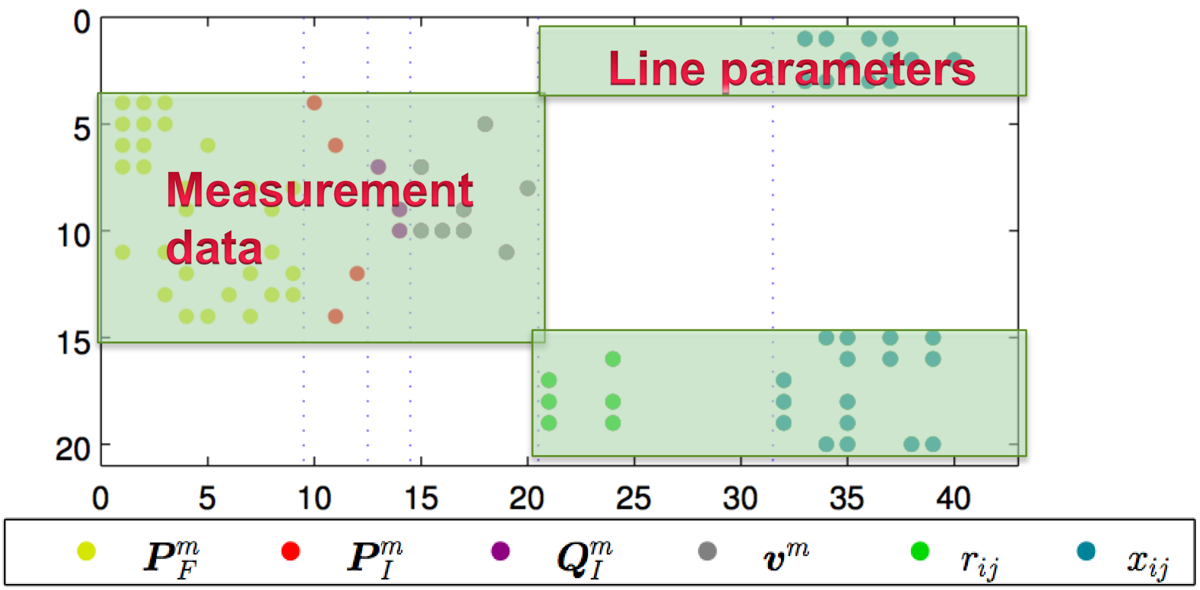
Results from SDP-SPCA: Sparsity pattern of V^T



➔ Separation into measurement & line parameter inputs
➔ Sparse structure

Numerical illustration of approach

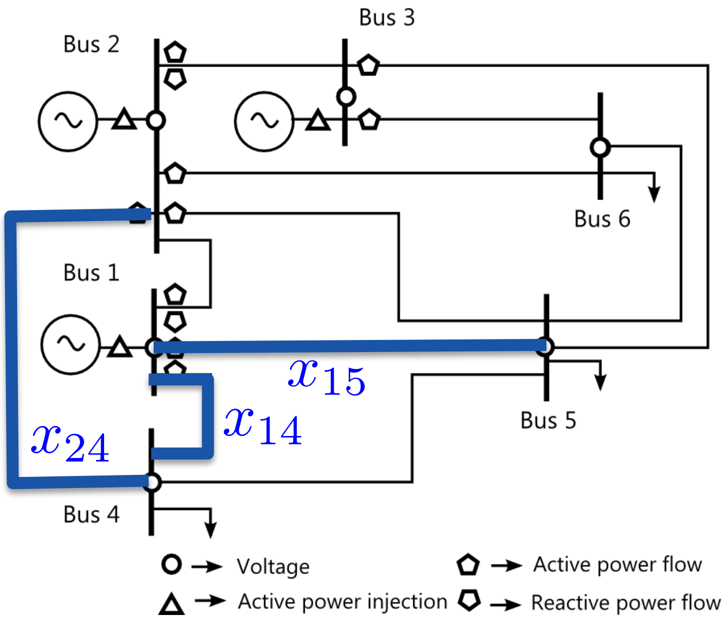
Results from Soft-Thresholding: Sparsity pattern of V^T

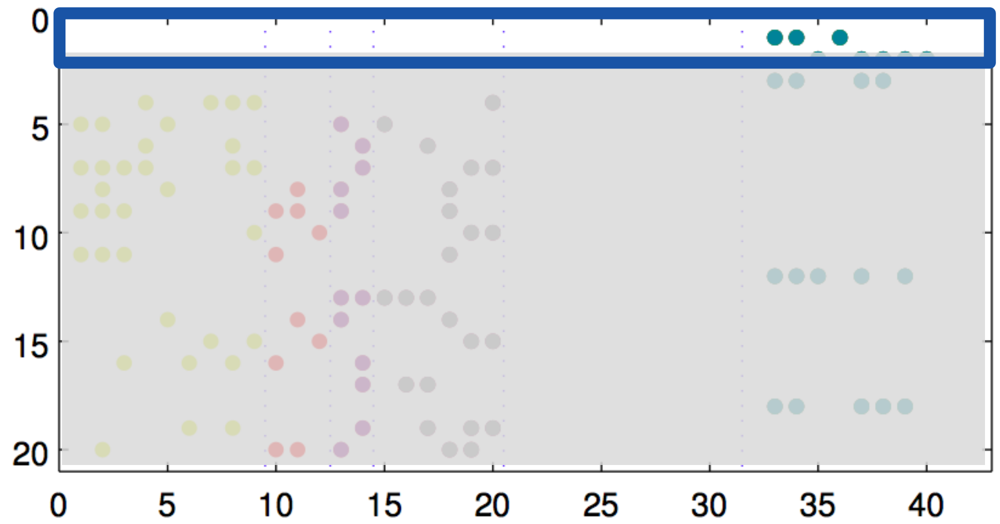


➔ Separation into measurement & line parameter inputs (Structure as theoretically proven)

➔ Sparse structure

Interpretation of sparsity pattern





➔ Perturbation of these 3 reactances may lead to 1.44 amplification of estimation error.

Conclusion

Analysis tool for identifying relevant components for power system state estimation – Valuable for model & sensor calibration

---

Approach concatenates sensitivity analysis & sparse PCA

Reveals relevant additional information for the Energy Management System (EMS)

Future work:

- Sparsity constraint on principal component to reflect concentration of perturbation impact
- Numerical tractability for large-scale power networks

SDP-SPCA

Define $\boldsymbol{X} = \boldsymbol{v}\boldsymbol{v}^\top$

$$\max_{\boldsymbol{X}} \operatorname{tr}\left[\boldsymbol{S}^\top \boldsymbol{S}\boldsymbol{X}\right] \quad \text{s.t.} \quad \begin{cases} \operatorname{tr}\left[\boldsymbol{X}\right] & = 1 \\ \mathbf{1}^\top |\boldsymbol{X}| \mathbf{1} & \leq t \\ \boldsymbol{X} \succeq 0 & \\ \cancel{\operatorname{rank}[\boldsymbol{X}] = 1} & \end{cases}$$

SDP relaxation:

- Omit rank constraint
- ℓ_1-relaxation of cardinality constraint

Comparison with simulated sensitivities

Predicted (from analytical model) vs simulated (Monte Carlo)

	SDP–SPCA		Soft-thres.	
# PC	predicted	simulated	predicted	simulated
1	1.44	1.42	1.51	1.48
2	1.28	1.26	1.25	1.20
3	1.22	1.18	1.16	1.10
4	1.00	0.97	0.99	1.00
5	1.00	0.99	0.99	0.99
6	0.99	0.97	0.99	0.98
7	0.99	0.99	1.00	0.98
8	0.98	0.97	1.00	1.03