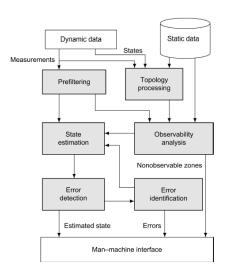
ECE 5314: Power System Operation & Control

Lecture 10: Power System State Estimation

Vassilis Kekatos

- R2 A. Gomez-Exposito, A. J. Conejo, C. Canizares, *Electric Energy Systems: Analysis and Operation*, Chapter 4.
- R1 A. J. Wood, B. F. Wollenberg, and G. B. Sheble, Power Generation, Operation, and Control, Wiley, 2014, Chapter 9.

Power system monitoring



Critical for

- situational awareness
- contingency analysis
- load forecasting
- economic operations
- billing

Power system state estimation (PSSE)

Problem: given meter readings and grid parameters, find actual state ${f v}$

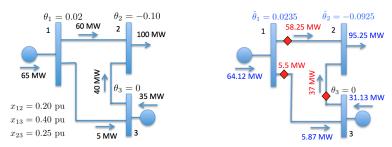


Figure: Left: actual state. Right: Measurements in red; use only $\{P_{12},P_{32}\}$ to find state $(\hat{\theta}_1,\hat{\theta}_2)$ via PF equations; estimated quantities in blue.

Why PSSE and not PF?

- 1. measurements are noisy
- 2. electric quantities are not measured at all grid locations

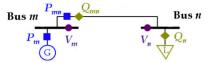
Lecture 11 V. Kekatos 3

Measurement model

ullet System state: vector of voltages ${f v}$ in polar or rectangular coordinates

$$z_m = h_m(\mathbf{v}) + \epsilon_m, \quad m = 1, \dots, M$$

• Function $h_m(\mathbf{v})$ can be linear or non-linear



- M: number of measurements
- ϵ_m noise of m-th measurement
- Zero-injection buses handled via constraints $p_n = q_n = 0$ or $\mathcal{I}_n = 0$
- If actual measurements are not enough, operators may use pseudo-measurements (scheduled generation, historical loads)

Introduction to estimation theory

• If ϵ_m is random and ${f v}$ deterministic, then $z_m=h_m({f v})+\epsilon_m$ is random

e.g., if
$$\epsilon_m \sim \mathcal{N}\left(0, \sigma_m^2\right) \implies z_m \sim \mathcal{N}\left(h_m(\mathbf{v}), \sigma_m^2\right)$$

$$p(z_m) = \frac{1}{\sqrt{2\pi\sigma_m^2}} \cdot \exp\left[-\frac{(z_m - h_m(\mathbf{v}))^2}{2\sigma_m^2}\right]$$

• If $\{\epsilon_m\}_{m=1}^M$ are independent, then $\{z_m\}_{m=1}^M$ are independent

$$p(\mathbf{z}; \mathbf{v}) = \prod_{m=1}^{M} p(z_m; \mathbf{v})$$

 ${\bf z}$ is measured; $h_m(\cdot)$'s and σ_m 's are known; and ${\bf v}$ is unknown

• For fixed v, the function p(z; v) is the **likelihood** of observing z

Maximum Likelihood Estimator (MLE)

Find the v that maximizes the **likelihood function** for the observed z

$$\max_{\mathbf{v}} \ p(\mathbf{z}; \mathbf{v}) = \prod_{m=1}^{M} p(z_m; \mathbf{v})$$

Equivalently [why?], minimize the negative log-likelihood function

$$\min_{\mathbf{v}} - \sum_{m=1}^{M} \log p(z_m; \mathbf{v})$$

For Gaussian noise $\epsilon_m \sim \mathcal{N}(0, \sigma_m^2)$, the MLE becomes

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \sum_{m=1}^{M} \frac{1}{2\sigma_m^2} (z_m - h_m(\mathbf{v}))^2$$

First-order optimality conditions

$$\nabla_{\mathbf{v}} \sum_{m=1}^{M} \log p(z_m; \mathbf{v}) = \mathbf{0}$$

Weighted least-squares (WLS)

- Statistical formulation of PSSE by seminal work of Schweppe
- Linear or non-linear weighted least-squares problem

$$\hat{\mathbf{v}}_{\text{WLS}} = \arg\min_{\mathbf{v}} \sum_{m=1}^{M} \frac{1}{2\sigma_m^2} (z_m - h_m(\mathbf{v}))^2$$

Non-convex problem unless measurement model is linear

$$\hat{\mathbf{v}} := \arg\min_{\mathbf{v}} \|\mathbf{z} - \mathbf{h}(\mathbf{v})\|^2$$

with
$$\mathbf{h}(\mathbf{v}) := [h_1(\mathbf{v}) \ \dots \ h_M(\mathbf{v})]^{\top}$$

- Let us ignore weights wlog; replace $z_m o rac{z_m}{\sigma_m}$ and $h_m({f v}) o rac{h_m({f v})}{\sigma_m}$
- F. C. Schweppe et al, "Power System Static-State Estimation, Part I-III," *IEEE Trans. Power Apparatus and Systems*, Vol. PAS-89, No. 1, Jan. 1970.

7

Phasor measurement units (PMUs)

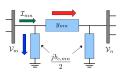
PMUs are *linear* functions of the state \mathbf{v} (if state in rectangular coordinates)!

Measurement of voltage at bus n

$$z_k = \mathcal{V}_n + \epsilon_k = \mathbf{e}_n^\top \mathbf{v} + \epsilon_k$$

Measurement of line current

$$\mathcal{I}_{mn} = \mathbf{y}_{mn} (\mathcal{V}_m - \mathcal{V}_n) + j \frac{b_{mn}^c}{2} \mathcal{V}_m$$
$$z_k = \mathbf{h}_k^{\mathsf{T}} \mathbf{v} + \epsilon_k$$



• Measurement of injection current $\mathcal{I}_m = \sum_{n \sim m} \mathcal{I}_{mn}$

$$z_k = \mathbf{h}_k^{\top} \mathbf{v} + \epsilon_k$$

• Vectors \mathbf{h}_k 's are known

A. Phadke - J. Thorp, 'Synchronized phasor measurements and their applications.' Springer, 2008.

Linear WLS estimator

• Collect measurement vectors \mathbf{h}_k 's as rows of matrix \mathbf{H}

$$z = Hv + \epsilon$$

complex- or real-valued model \mathbf{H} is $\mathbb{C}^{M \times N}$ or $\mathbb{R}^{2M \times 2N}$

• MLE is the (Weighted) Least-Squares estimator found by solving

$$\min_{\mathbf{v}} \|\mathbf{z} - \mathbf{H}\mathbf{v}\|_2^2$$

(unconstrained) convex quadratic program

• Minimizer is unique if H is full column-rank

$$\hat{\mathbf{v}}_{\mathsf{WLS}} = (\mathbf{H}^{\top} \mathbf{H})^{-1} \mathbf{H} \mathbf{z}$$

need at least $M \ge N$; typically $M \sim 1.7 - 2.3N$

SCADA measurements

Supervisory Control And Data Acquisition (SCADA) system measurements:

- line power flows $\{P_{mn}, Q_{mn}\}$
- line current magnitudes $\{|\mathcal{I}_{mn}|\}$
- bus voltage magnitudes $\{V_n\}$
- bus power injections $\{p_n, q_n\}$



• Functions h_m 's for SCADA are *nonlinear*; quadratic if ${\bf v}$ in Cartesian

• MLE on SCADA data yields a non-linear WLS fit (non-convex problem)

$$\hat{\mathbf{v}}_{\mathsf{WLS}} := \arg\min_{\mathbf{v}} \|\mathbf{z} - \mathbf{h}(\mathbf{v})\|_2^2$$

Gauss-Newton method

- Applies to non-linear LS; do not confuse with Newton-Raphson's method
- Approximate $\mathbf{h}(\mathbf{v})$ with its first-order Taylor's series expansion at \mathbf{v}_k

$$\mathbf{h}(\mathbf{v}) \simeq \hat{\mathbf{h}}_k(\mathbf{v}) = \mathbf{h}(\mathbf{v}_k) + \mathbf{J}_k(\mathbf{v} - \mathbf{v}_k), \qquad \mathbf{J}_k$$
 : Jacobian matrix

- Gauss-Newton updates: $\mathbf{v}_{k+1} = \arg\min_{\mathbf{v}} \|\mathbf{z} \hat{\mathbf{h}}_k(\mathbf{v})\|_2^2$
- Since $\hat{\mathbf{h}}_k(\mathbf{v})$ is linear, GN solves a sequence of *linear* LS problems

$$\mathbf{v}_{k+1} = \mathbf{v}_k + (\mathbf{J}_k^{\mathsf{T}} \mathbf{J}_k)^{-1} \mathbf{J}_k^{\mathsf{T}} (\mathbf{z} - \mathbf{h}(\mathbf{v}_k))$$

• GN method converges to a local minimum

Solvers using QR decomposition*

Inverting $\mathbf{J}_k^{ op} \mathbf{J}_k$ is sensitive to its condition number

Solvers based on the QR decomposition are numerically more robust

- QR decomposition: $\mathbf{J}_k = \mathbf{Q}_k \tilde{\mathbf{R}}_k = \mathbf{Q}_k \begin{bmatrix} \mathbf{R}_k \\ \mathbf{0} \end{bmatrix}$ for orthogonal $\mathbf{Q}_k^{\top} \mathbf{Q}_k = \mathbf{I}_M$ and upper triangular \mathbf{R}_k
- · Find the minimizer of

$$\mathbf{v}^{k+1} = \arg\min_{\mathbf{v}} \ \|\mathbf{z}_k - \mathbf{J}_k \mathbf{v}\|_2^2 = \|\mathbf{Q}_k^{\top} \mathbf{z}_k - \tilde{\mathbf{R}}_k \mathbf{v}\|_2^2$$

by solving a linear system with upper triangular form

- Upper part of error vector can be made zero
- Same complexity (due to QR), yet better numerical properties

Semidefinite program relaxation

- As in PF and OPF, introduce $V = vv^H$
- Measurements are quadratic functions of v; linear functions of V

$$h_m(\mathbf{V}) = \text{Tr}(\mathbf{M}_m \mathbf{V}) + \epsilon_m, \qquad m = 1, \dots, M$$

Express PSSE using SDP variable

$$\min_{\mathbf{V}} \quad \sum_{m=1}^{M} (z_m - \text{Tr}(\mathbf{M}_m \mathbf{V}))^2$$
s.to $\mathbf{V} \succeq \mathbf{0}, \text{ rank}(\mathbf{V}) = 1$

Drop rank constraint to get a convex problem

H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. of Sel. Topics in Signal Processing*, Vol. 8, No. 6, Dec. 2014.

Lecture 11 V. Kekatos 13

More on semidefinite program relaxation

Problem transformed to SDP via epigraph trick and Schur's complement

$$\min_{\mathbf{V}, \mathbf{t}} \sum_{m=1}^{M} t_m$$
s.to
$$\begin{bmatrix} t_m & z_m - \text{Tr}(\mathbf{M}_m \mathbf{V}) \\ z_m - \text{Tr}(\mathbf{M}_m \mathbf{V}) & 1 \end{bmatrix} \succeq \mathbf{0}, \ m = 1, \dots, M$$

$$\mathbf{V} \succeq \mathbf{0}$$

- · Relaxation is NOT exact for noisy measurements
- Heuristics based on minimizer $\hat{\mathbf{V}}$ can give good state estimates
 - Select $\hat{\mathbf{v}}$ as eigenvector corresponding to largest eigenvalue of $\hat{\mathbf{V}}$
 - Draw random states $\hat{\mathbf{v}}\sim\mathcal{CN}(\mathbf{0},\hat{\mathbf{V}})$ and keep the one with the smallest fit

R. Madani, A. Ashraphijuo, J. Lavaei, and R. Baldick, "Power System State Estimation with a Limited Number of Measurements," in Proc. *IEEE Conf. on Decision and Control*, Dec. 2016.

Fast decoupled solver

To avoid evaluating \mathbf{J}_k and inverting $\mathbf{J}_k^{\top}\mathbf{J}_k$

- resort to similar tricks as in power flow solvers
- decoupling between $P-\theta$ and Q-V
- ullet approximate $(\mathbf{J}_k^{ op}\mathbf{J}_k)^{-1}$ at flat voltage profile $\mathbf{v}=\mathbf{1}+j\mathbf{0}$

A. Abur and A. Gomez-Exposito, "Power system state estimation: theory and implementation," *CRC Press*, 2004.

Lecture 11 V. Kekatos 15

SCADA vs. PMU measurements

SCADA:

- readings every 4 secs
- no grid-wide angles due to lack of timing $\{|V_m|, P_m, Q_m, P_{mn}, Q_{mn}, I_{mn}\}$
- ullet non-linear model: $\mathbf{z} = \mathbf{h}(\mathbf{v}) + oldsymbol{\epsilon}$
- MLE via non-convex problem: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{z} \mathbf{h}(\mathbf{v})\|_2^2$

PMUs:

- 30 readings per second
- time-synchronized via GPS signaling: complex currents and voltages
- linear model (v in rectangular coordinates): $\mathbf{z} = \mathbf{H}\mathbf{v} + \boldsymbol{\epsilon}$
- MLE via convex problem: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{z} \mathbf{H}\mathbf{v}\|_2^2 = (\mathbf{H}^{\top}\mathbf{H})^{-1}\mathbf{H}\mathbf{z}$

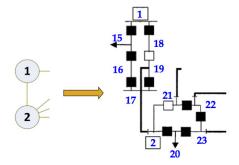
Circuit breakers

Devices used for protection and grid reconfiguration

CBs are zero-impedance elements

• closed CB_{15,16}: $V_{15} = V_{16}$

• open $CB_{18,19}$: $I_{18,19} = 0$



Bus-branch model (left); bus section-switch model (right)

Generalized state estimation

- Network topology processor collects circuit breaker (CB) statuses
- Topology errors easily detected, but hardly identified by PSSE
- Generalized state estimation (GSE) seeks state and grid topology

$$\begin{split} \hat{\mathbf{x}} &= \arg\min_{\mathbf{x}} \quad \frac{1}{2} \|\mathbf{z} - \mathbf{H}\mathbf{x}\|_2^2 \\ &\text{s.to} \quad \mathbf{C}\mathbf{x} = \mathbf{0} \end{split}$$

- State augmented to include section voltages and CB currents
- CB statutes effect linear equality constraints to ensure identifiability
- Solution via linear system

$$\begin{bmatrix} \mathbf{H}^{\top}\mathbf{H} & \mathbf{C}^{\top} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{\top}\mathbf{z} \\ \mathbf{0} \end{bmatrix}$$

Observability (identifiability) analysis

Given measurement set and grid parameters, assess state identifiability. If non-identifiable, find observable islands (maximally connected subgrids).

Q: Why bother? A: To select locations for pseudo-measurements

Observable: if different states $\mathbf{v}_1 \neq \mathbf{v}_2$ yield different outputs $\mathbf{h}(\mathbf{v}_1) \neq \mathbf{h}(\mathbf{v}_2)$

- runs online to cope with meter failures, meter delays, and grid changes
- uses the DC power flow model (pairs of active-reactive meters)

$$P_m = \sum_{n \neq m} b_{mn} (\theta_m - \theta_n)$$

numerical and topological observability

K. A. Clements, P. R. Krumpholz, and G. W. Davis, "PSSE with measurement deficiency: An observability/measurement placement algorithm," *IEEE Trans. Power App. Syst.*, July 1983. A. Monticelli, "Electric power system state estimation," *Proc. of the IEEE*, Feb. 2000.

Numerical observability

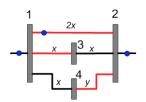
• Recall from DC grid model (state is θ):

$$\mathbf{p} = \mathbf{A}^{\top} \mathbf{X}^{-1} \mathbf{A} \boldsymbol{\theta} \quad \text{and} \quad \mathbf{f} = \mathbf{X}^{-1} \mathbf{A} \boldsymbol{\theta}$$

- If measurements $\mathbf{z} = \mathbf{H} \boldsymbol{\theta}$ are only some p_n 's and f_ℓ 's, then $\mathbf{H} = \mathbf{S} \mathbf{A}$ for a wide matrix $\mathbf{S} \in \mathbb{R}^{M \times L}$
- ullet In general, system ${f z}={f H}{m heta}$ is identifiable iff ${f H}$ is full column-rank
- But, in PSSE there is an angle shift ambiguity: $\mathbf{A} oldsymbol{ heta} = \mathbf{A} (oldsymbol{ heta} + c \mathbf{1})$
- Power system is observable iff $null(\mathbf{H}) = null(\mathbf{A}) = \{c\mathbf{1}, c \in \mathbb{R}\}$
- If not, for $u \in \operatorname{null}(H)$, the non-zeros of Au define unobservable lines
- · Systematic removal of unobservable branches reveals observable islands

Topological observability

- Graph-theoretic approach that studies the sparsity pattern of H rather than the values of its entries
- Builds spanning tree (forest) by selecting branches (lines)
 - a flow measurement is assigned to its branch
 - an injection measurement can be assigned to an incident branch
 - at the end, reject injection meas. if they have a non-selected incident branch that does not form a loop with existing branches

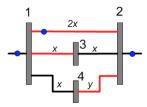


• Different algorithms implement the above

[R2] Gomez-Exposito, Conejo, Canizares, Electric Energy Systems, Chapter 4.6

Topological vs. numerical observability

- If a system is numerically observable, it is also topologically observable
- · Converse does not hold



- Counter-example: shown system with x = y
 - · topologically is observable
 - but numerically, it is not

A. Monticelli, "Electric power system state estimation," Proc. of the IEEE, Feb. 2000.

Bad data

Sources: time-skews, uncalibrated meters, reverse wiring

- Preprocessing (polarity and range tests)
- Assume linear model (DC, linearized AC, or PMU measurements)

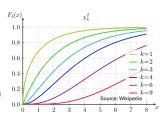
$$\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}, \quad \mathbf{H} \in \mathbb{R}^{M \times N}$$

- Least-squares estimator (LSE) solves $\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{z} \mathbf{H}\mathbf{x}\|_2^2$
- LSE: $\hat{\mathbf{x}} = (\mathbf{H}^{\top} \mathbf{H})^{-1} \mathbf{H}^{\top} \mathbf{z}$
- LSE residual: $\mathbf{r} := \mathbf{z} \mathbf{H}\hat{\mathbf{x}} = \mathbf{P}_{\mathbf{H}}^{\perp}\mathbf{z} = \mathbf{P}_{\mathbf{H}}^{\perp}\boldsymbol{\epsilon}$ where $\mathbf{P}_{\mathbf{H}} := \mathbf{H}(\mathbf{H}^{\top}\mathbf{H})^{-1}\mathbf{H}^{\top}$ is a projection matrix onto $\mathrm{range}(\mathbf{H})$ and $\mathbf{P}_{\mathbf{H}}^{\perp} := \mathbf{I}_{M} - \mathbf{P}_{\mathbf{H}}$ is a projection matrix onto $\mathrm{range}(\mathbf{H})^{\perp}$
- For $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, it holds that $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\mathbf{H}}^{\perp})$, $\mathrm{rank}(\mathbf{P}_{\mathbf{H}}^{\perp}) = M N$

Revealing outliers

Chi-squared test (detection test)

- Because $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\mathbf{H}}^{\perp}),$ then $\|\mathbf{r}\|_2^2 \sim \chi_{M-N}^2$
- Find α for which $\Pr(\|\mathbf{r}\|_2^2 \le \alpha) \ge 0.95$
- If $\|\mathbf{r}\|_2^2 > \alpha = F_{\chi^2_{M-N}}^{-1}(0.95)$, declare bad data



Largest normalized residual test (LNRT) (identification test)

- $r_i/\sqrt{P_{ii}}\sim \mathcal{N}(0,1),\ P_{ii}$ is the *i*-th diagonal entry of $\mathbf{P}_{\mathbf{H}}^{\perp}$
- If $\max_i \frac{|r_i|}{\sqrt{P_{ii}}} > t$, measurement z_i for maximizing i is deemed bad
- Remove bad datum and re-compute LSE
- LSE can be computed efficiently using the matrix inversion lemma

Least Absolute Deviations (LAV) for robustness

- LSE is sensitive to outliers: single bad datum can deteriorate estimate
- Least-median of squares (LMS) is an NP-hard problem

$$\hat{\mathbf{x}}_{\mathrm{LMS}} := \arg\min_{\mathbf{x}} \max_{i} (z_i - \mathbf{h}_i^{\top} \mathbf{x})^2$$

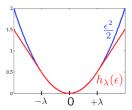
Least-absolute deviations (LAV) can be expressed as a linear program

$$\begin{split} \hat{\mathbf{x}}_{\mathrm{LAV}} := \arg\min_{\mathbf{x}} \|\mathbf{z} - \mathbf{H}\mathbf{x}\|_1 &\iff &\min_{\mathbf{x}, \mathbf{t}} \ \mathbf{t}^\top \mathbf{1} \\ &\text{s.to} \ -\mathbf{t} \leq \mathbf{z} - \mathbf{H}\mathbf{x} \leq \mathbf{t} \end{split}$$

LAV with SDP relaxation for robust nonlinear PSSE

$$\min_{\mathbf{V} \succeq \mathbf{0}, \mathbf{t}} \quad \mathbf{t}^{\top} \mathbf{1}$$
s.to $-t_m \le z_m - \text{Tr}(\mathbf{M}_m \mathbf{V}) \le t_m, \quad m = 1, \dots, M$

Huber estimator



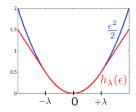
Huber's and quadratic functions

Huber estimator: combination of LS and LAV

$$\hat{\mathbf{x}}_{\mathrm{H}} := \arg\min_{\mathbf{x}} \sum_{i=1}^{M} h_{\lambda}(z_i - \mathbf{h}_i^{\top} \mathbf{x})$$

L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics of power systems," *IEEE Trans. on Power Systems*, Vol. 11, No. 2, May 1996.

Reformulating Huber's estimator



Convex problem that can be reformulated as

$$\min_{\mathbf{x}, \mathbf{o}} \ \frac{1}{2} \|\mathbf{z} - \mathbf{H}\mathbf{x} - \mathbf{o}\|_2^2 + \lambda \|\mathbf{o}\|_1$$

• Interpretation via compressed sensing and the outlier model

$$z = Hx + \epsilon + o$$

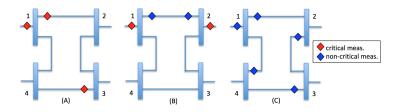
27

V. Kekatos and G. B. Giannakis, "Distributed Robust Power System State Estimation," *IEEE Trans. on Power Syst.*, Vol. 28, No. 2, May 2013.

Critical measurements

Lecture 11

Measurement i is critical if once removed from measurement set, the power system becomes unobservable



- Claim: column i of $\mathbf{P}_{\mathbf{H}}^{\perp}$ is zero; hence, $r_i=0$ $(\mathbf{r}=\mathbf{P}_{\mathbf{H}}^{\perp}\boldsymbol{\epsilon})$
- For critical measurement: undefined $\frac{|r_i|}{\sqrt{P_{ii}}}$, impossible cross-validation
- Bad data processing vulnerable to critical measurements
- Multiple corrupted readings: communication link failures, cyber-attacks

O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. on Smart Grid*, Vol. 2, No. 4, Dec. 2011.

V. Kekatos

28

Cyber attacks

- Numbers of cyber-incidents on power SCADA: 3 (2009), 25 (2011)
 - Challenged by increased sensing and networking
 - GPS spoofing, generator controls, CB tripping
- Cyber-attacks on PSSE (situational awareness, billing, trading)

$$z = Hx + \epsilon + a$$

compromised meters at non-zero entries of a

- ullet Stealth attacks can arbitrarily mislead PSSE by ${f a}={f H}{f v}$
- Deleting related rows of H deems the system unobservable another perspective of multiple critical measurements

Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," ACM Trans. Info. and System Security, May 2011.

Topics not covered

- Dynamic state estimation
- Decentralized solvers
- Attacks to energy markets
- Phasor measurement units (internal estimation algorithms)
- Measurement placement
- Estimating line and transformer parameters