Quantifying Cyber Security for Networked Control Systems

Henrik Sandberg
ACCESS Linnaeus Centre, KTH Royal Institute of Technology

Joint work with:
André Teixeira, György Dán, Karl H. Johansson (KTH)
Kin Cheong Sou (Chalmers)
Julien M. Hendrickx, Raphael M. Jungers (Louvain)

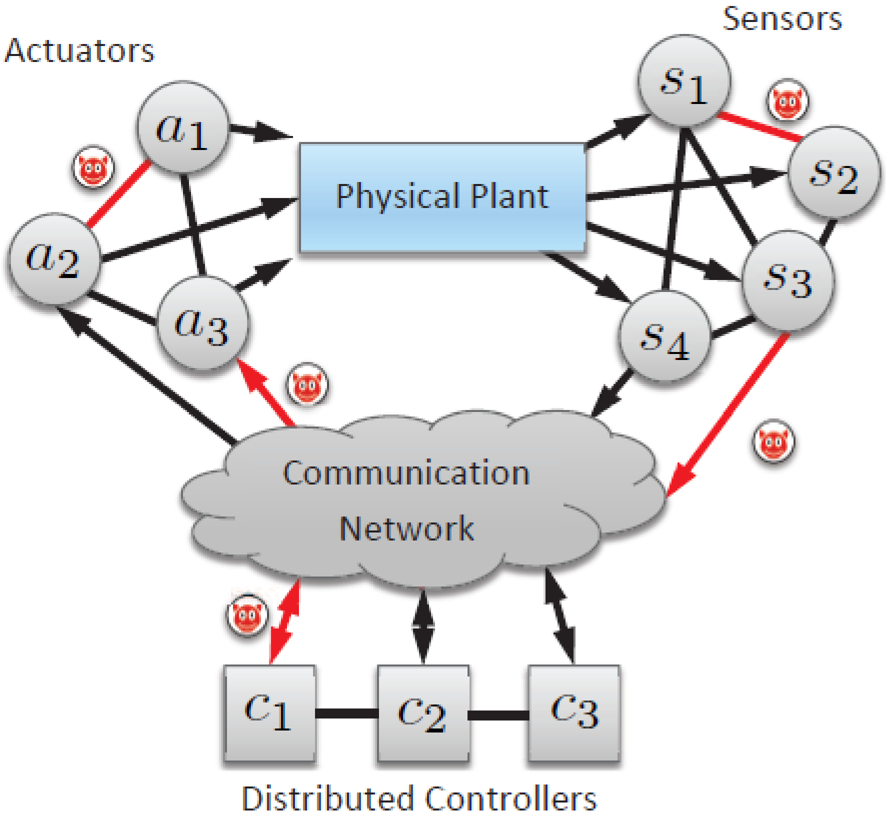
UC Berkeley
May 17th, 2013

Motivation

- Networked control systems are to a growing extent
  - based on commercial off-the-shelf components
  - integrated with data analytics environments etc.

- Leads to **increasing vulnerability** to **cyber-physical threats** with many potential points of attacks

- Need for tools and strategies to understand and mitigate attacks in networked control systems:
  - Which threats should we care about?
  - What impact can we expect from attacks?
  - Which resources should we protect (more)?



ROYAL INSTITUTE OF TECHNOLOGY

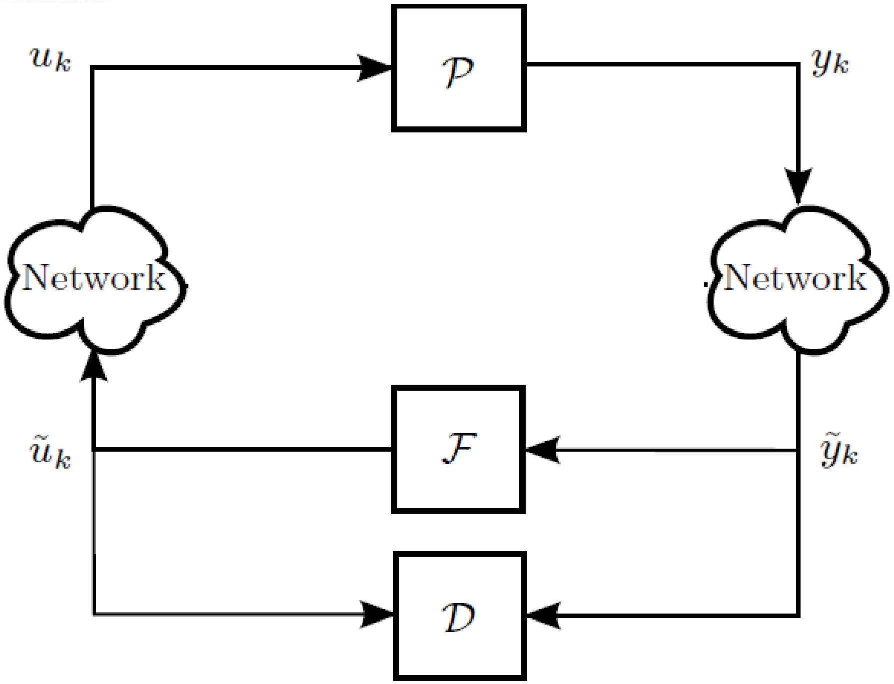
Contributions

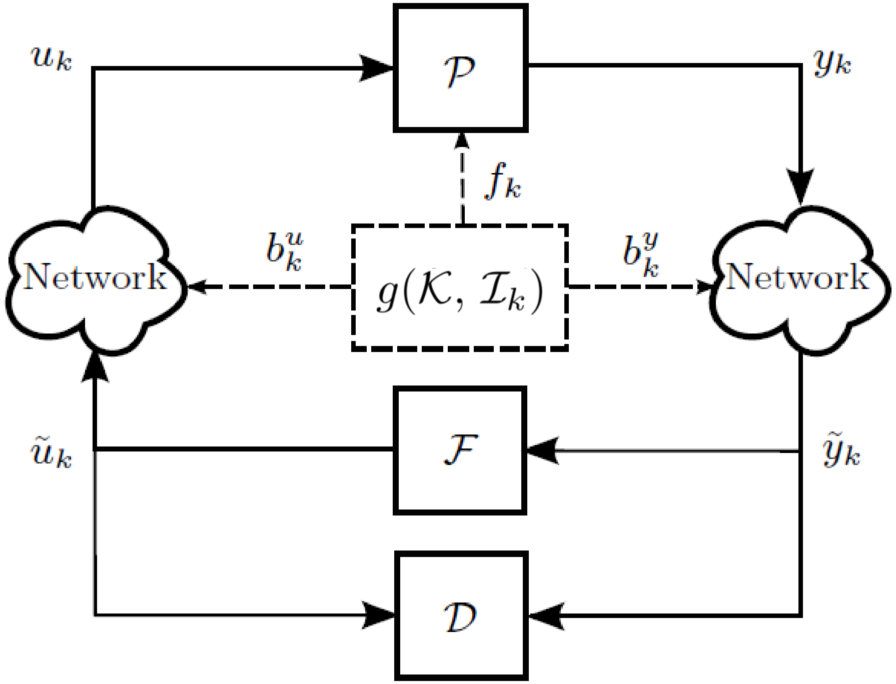
- Adversary models for networked control systems
- Optimization tools for quantization of cyber security
  - Trade-off between protection resources and level of security
  - Trade-off between attack resources and attack impact
- Security metric for power network state estimators. Efficient computation using graph Min Cut relaxations
- Security metric for wireless LQG-controlled quadruple tank. Computation using mixed integer linear programs

Outline

- **Adversary models for networked control systems**
- Application 1: Power network state estimation
- Application 2: Wireless LQG-controlled quadruple tank

Networked Control System under Attack





- Physical plant ($\mathcal{P}$)
- Feedback controller ($\mathcal{F}$)
- Anomaly detector ($\mathcal{D}$)

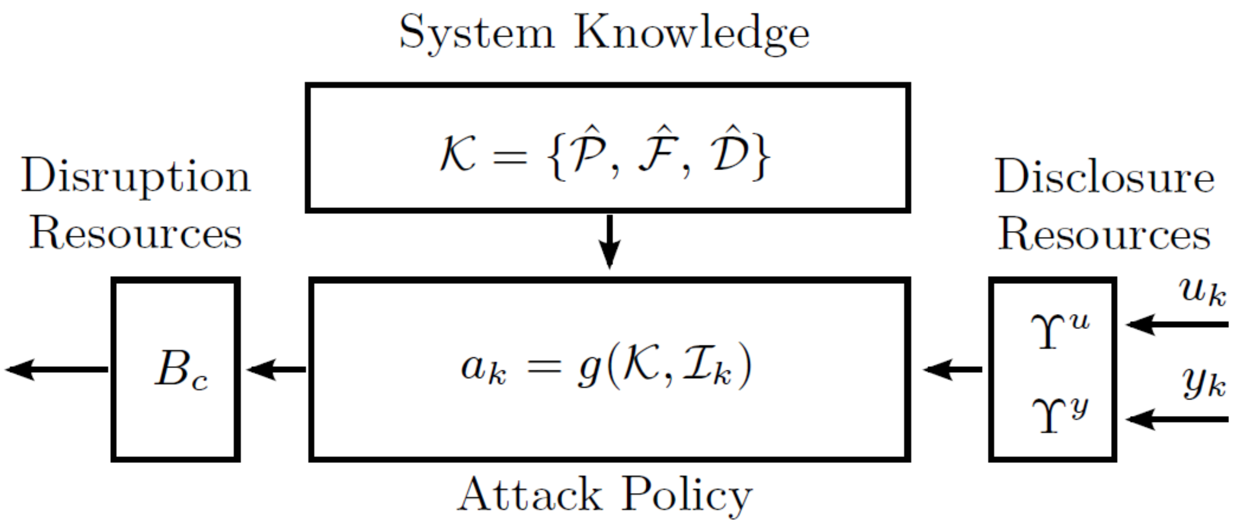
- Disclosure Attacks

- Physical Attacks f_k

- Deception Attacks

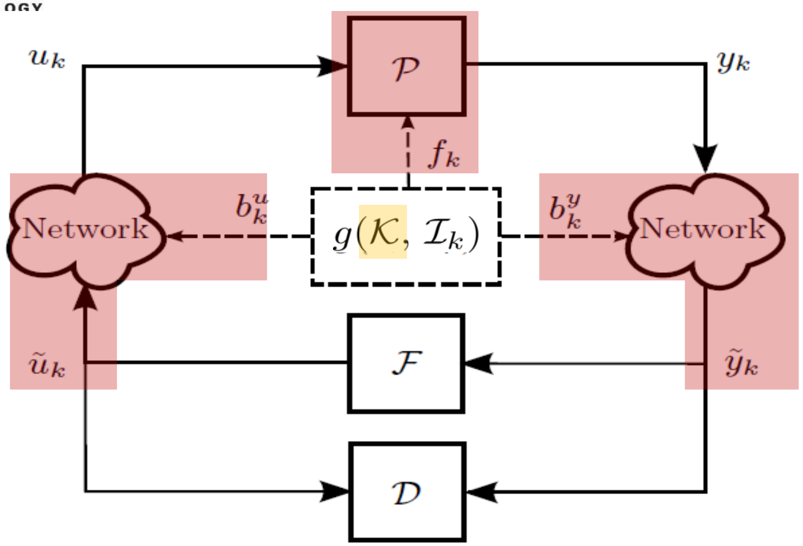
$$\tilde{u}_k = u_k + \Gamma^u b_k^u$$
$$\tilde{y}_k = y_k + \Gamma^y b_k^y$$

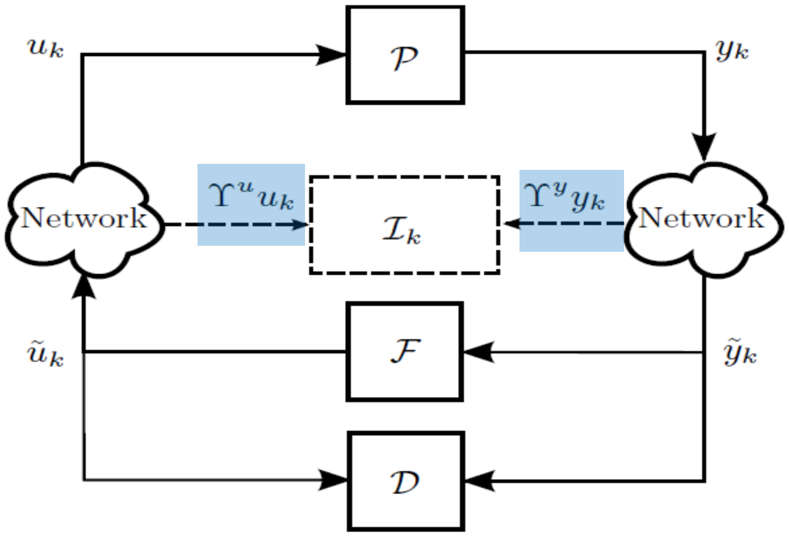
Adversary Model

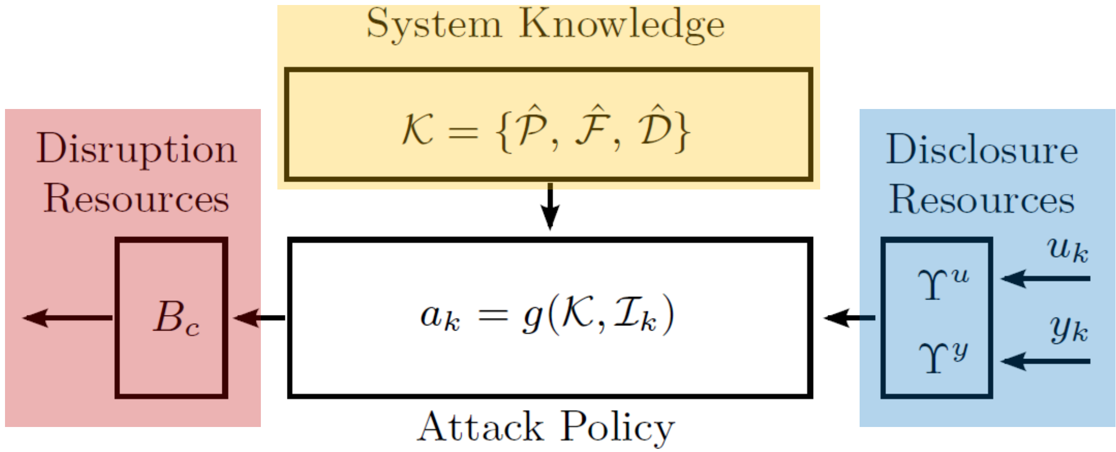


- Adversary's goal to force the process state into an unsafe region
- Attack should be stealthy, i.e., no alarms (at least until it is too late)
- Adversary constrained by limited resources

Networked Control System with Adversary Model

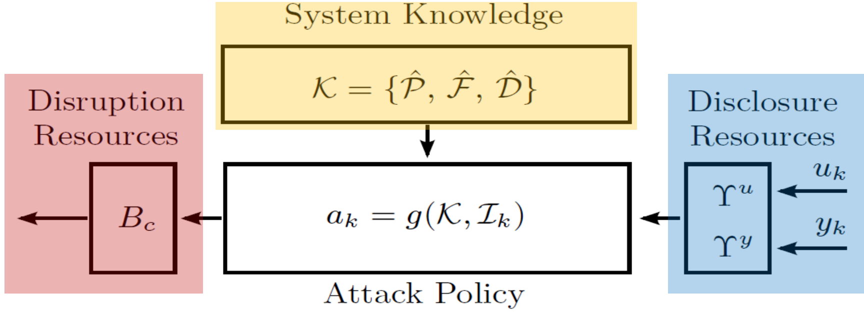


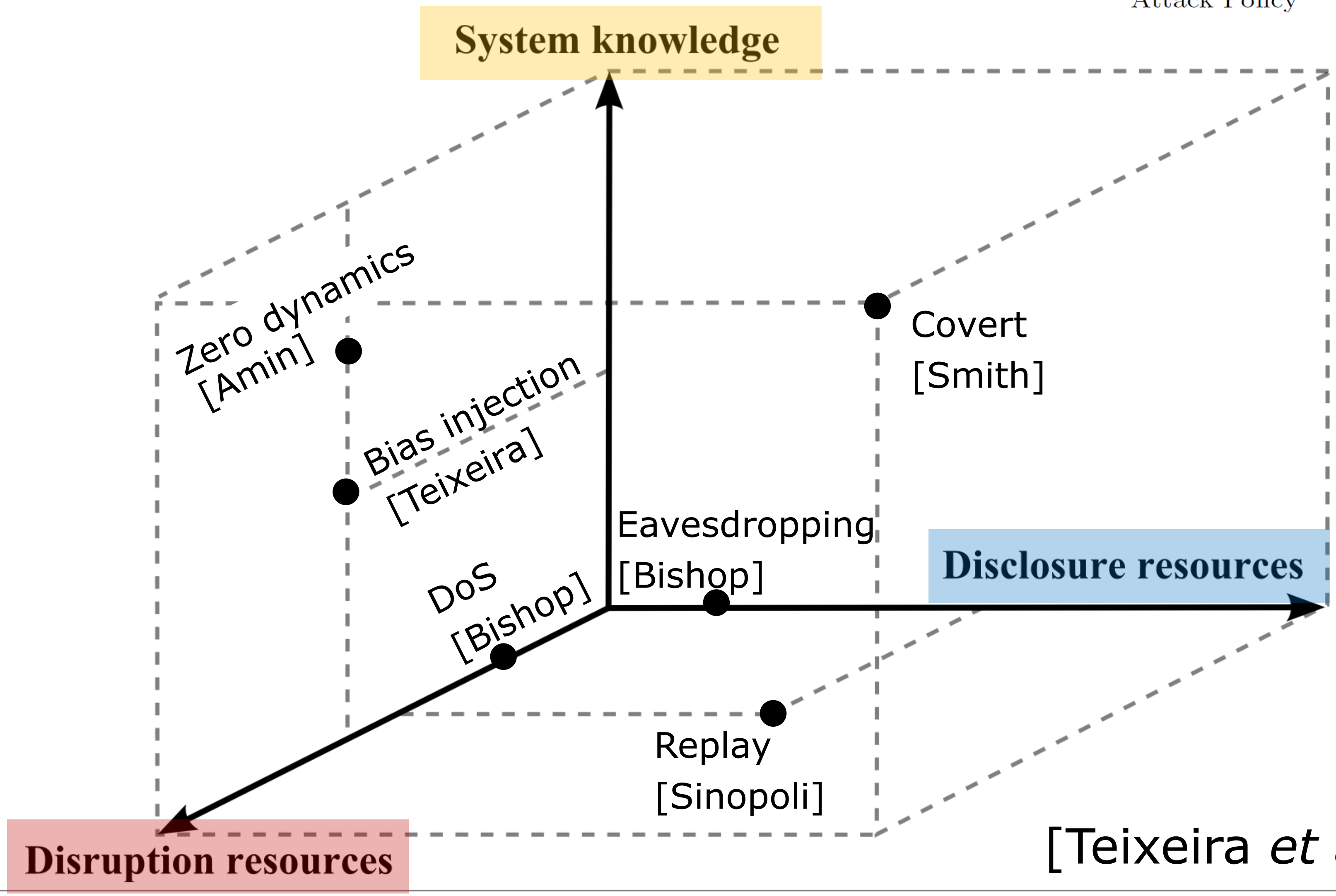




ROYAL INSTITUTE OF TECHNOLOGY

Attack Space





[Teixeira *et al.*, 2012]

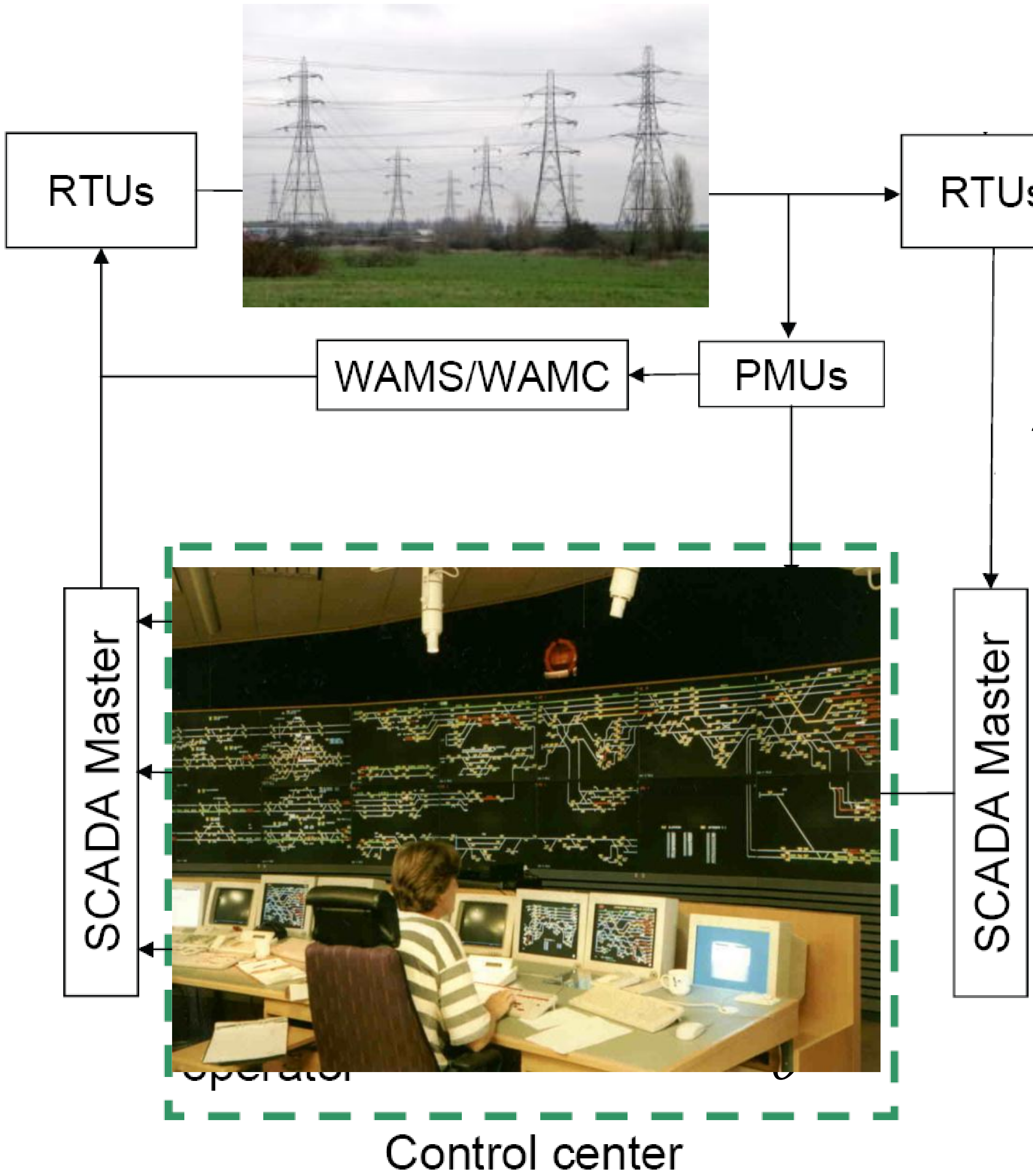
Outline

- Adversary models for networked control systems

- **Application 1: Power network state estimation**
  - **Security index definition**
  - **Computation with LASSO/graph Min Cut relaxations**

- Application 2: Wireless LQG-controlled quadruple tank

Power Network Control System



Remote
Terminal
Units
(in substations)

Z measurement

Supervisory
Control
And
Data
Acquisition

Model-Based State Estimation

Given redundant measurement z, find state estimate $\hat{\theta}$ based on **steady-state** model

state θ

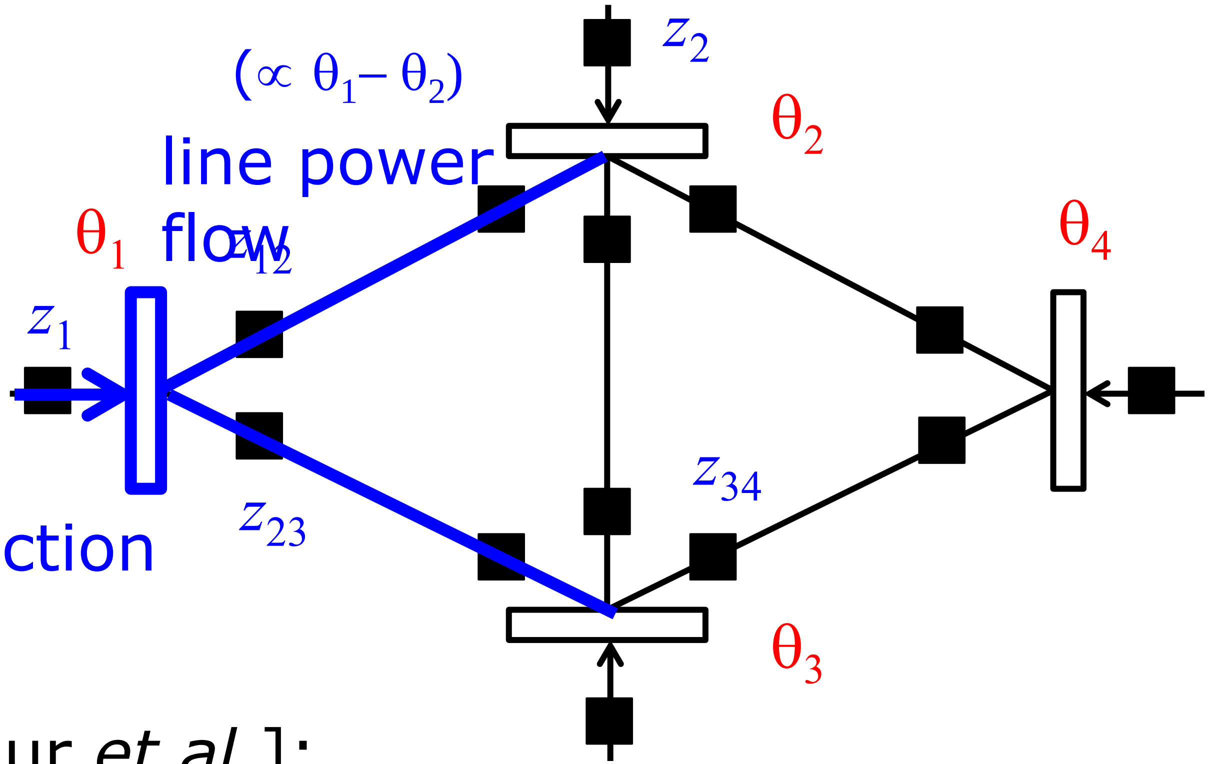
measurement z

$$z = h(\theta)$$

Power Network State Estimation Model

States (θ)
= bus voltage phase angles

(flow conservation)
bus injection

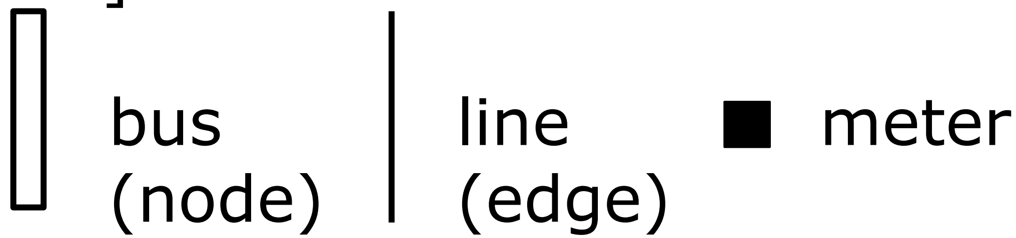


Measurements (z)
= line power flow & bus injection

“DC power flow model” [Abur *et al*.]:

$$z = H\,\theta$$

measurement matrix



State Estimation by Least Squares

State estimator (LS)

$$z = H\,\theta$$

$$\Rightarrow \hat{\theta} = \left(H^T H\right)^{-1} H^T z$$

wrong

wrong

Contingency analysis

wrong

OPF calculations

⋮

What if the measurements were **wrong**?

$$z \leftarrow z + \Delta z$$

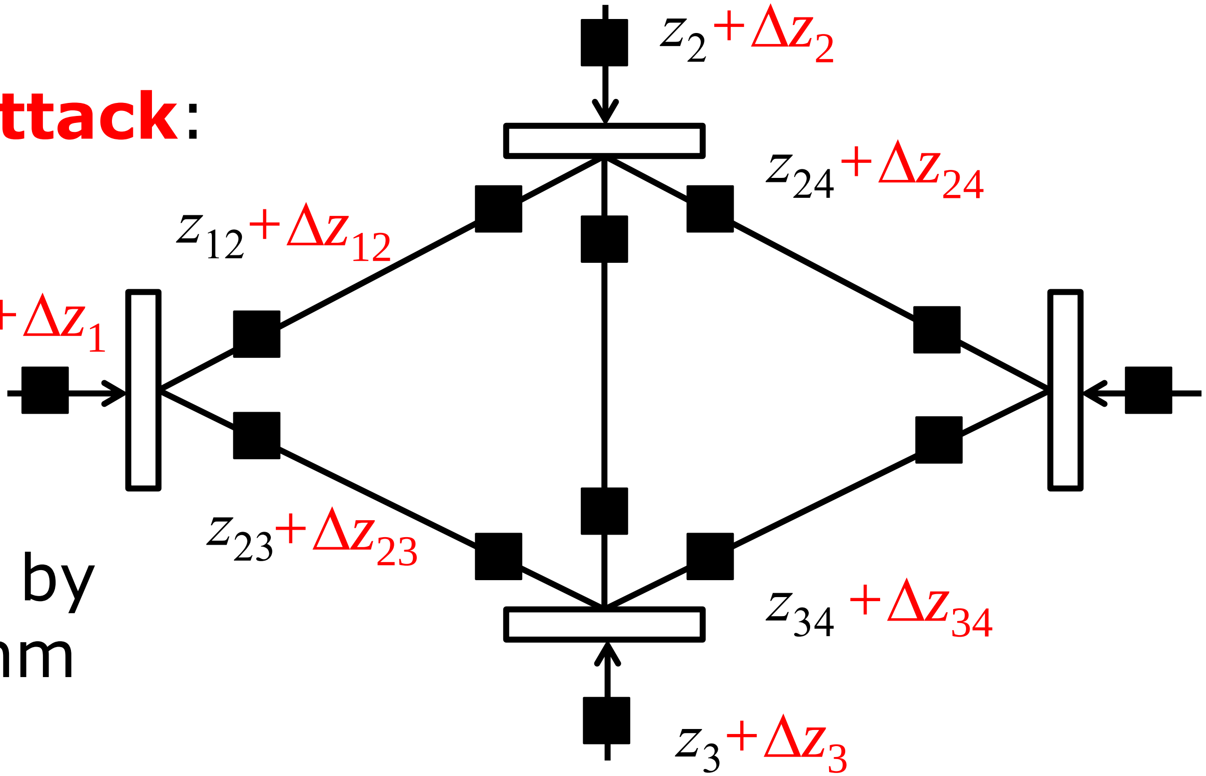
random measurement noise

intentional data attack

Stealth Additive False-Data Attack

Measurements subject to **attack**:

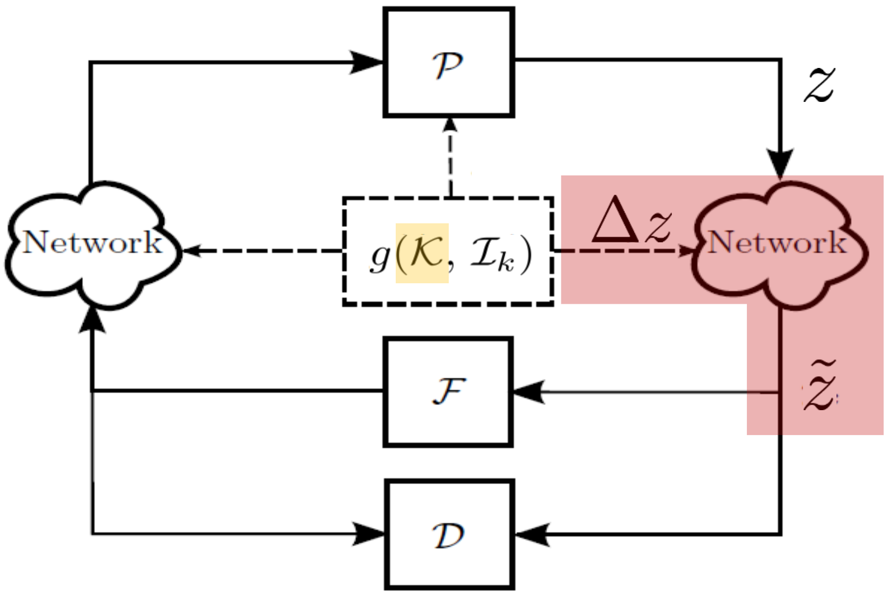
$$z \leftarrow z + \Delta z$$

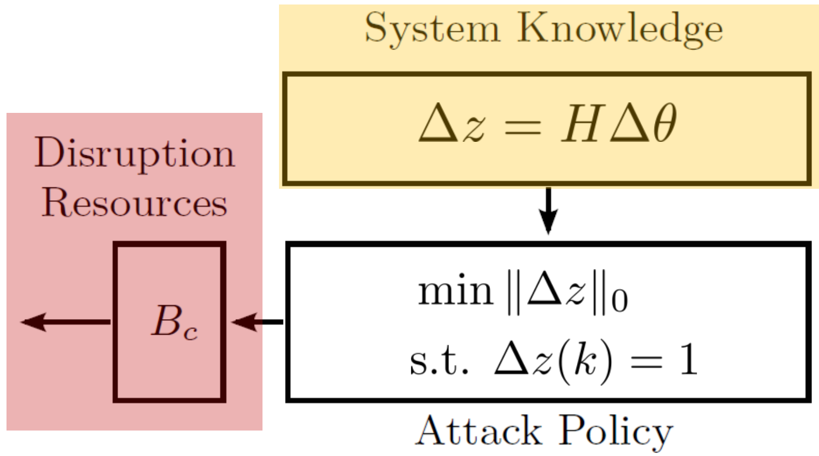


Attack is **constrained**; otherwise will be **detected** by Bad Data Detection algorithm

Stealth attack [Liu *et al*., Giani *et al*.]: $\Delta z = H\Delta\theta$

Adversary Model





- Adversary's goal to induce a bias in measurement channel k
- Attack should be stealthy, i.e., no alarms
- Adversary should use minimal disruption resources

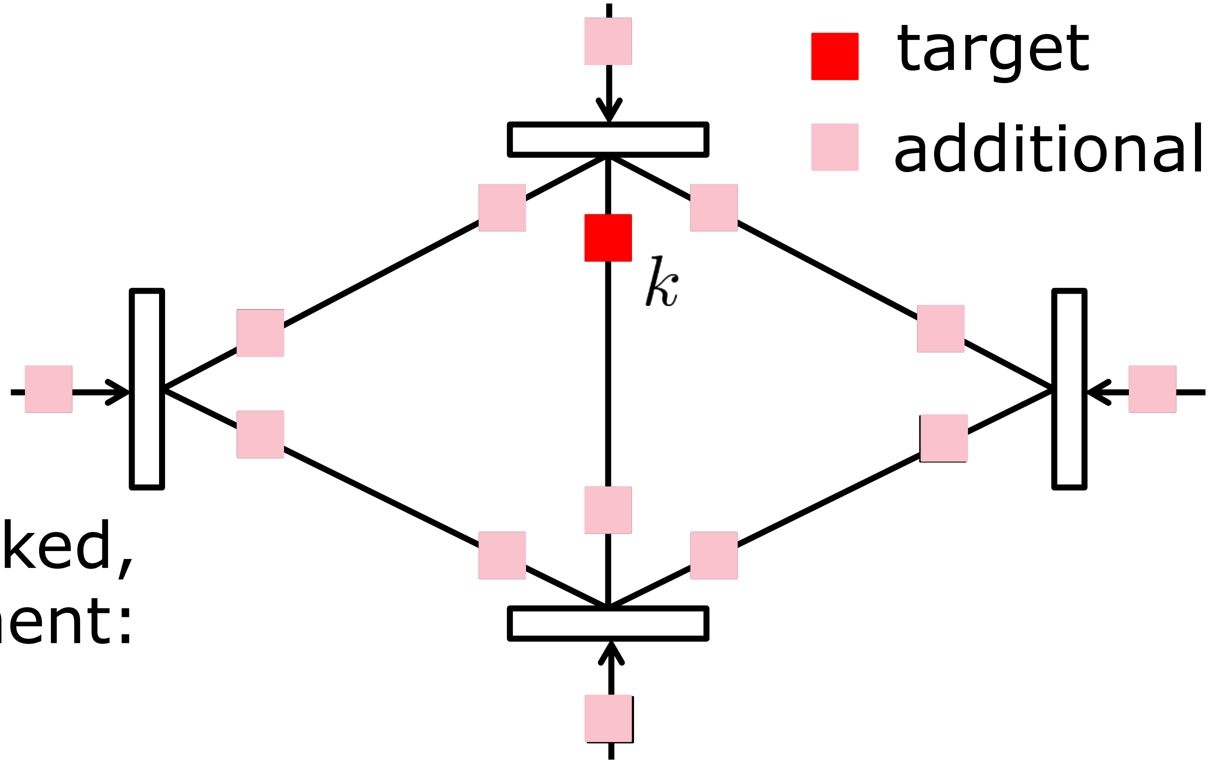
Security Index

Stealth attack $\Delta z = H\Delta\theta$

In general, $e_k \notin \operatorname{span}(H)$

Minimum # of meters attacked, targeting the k th measurement:

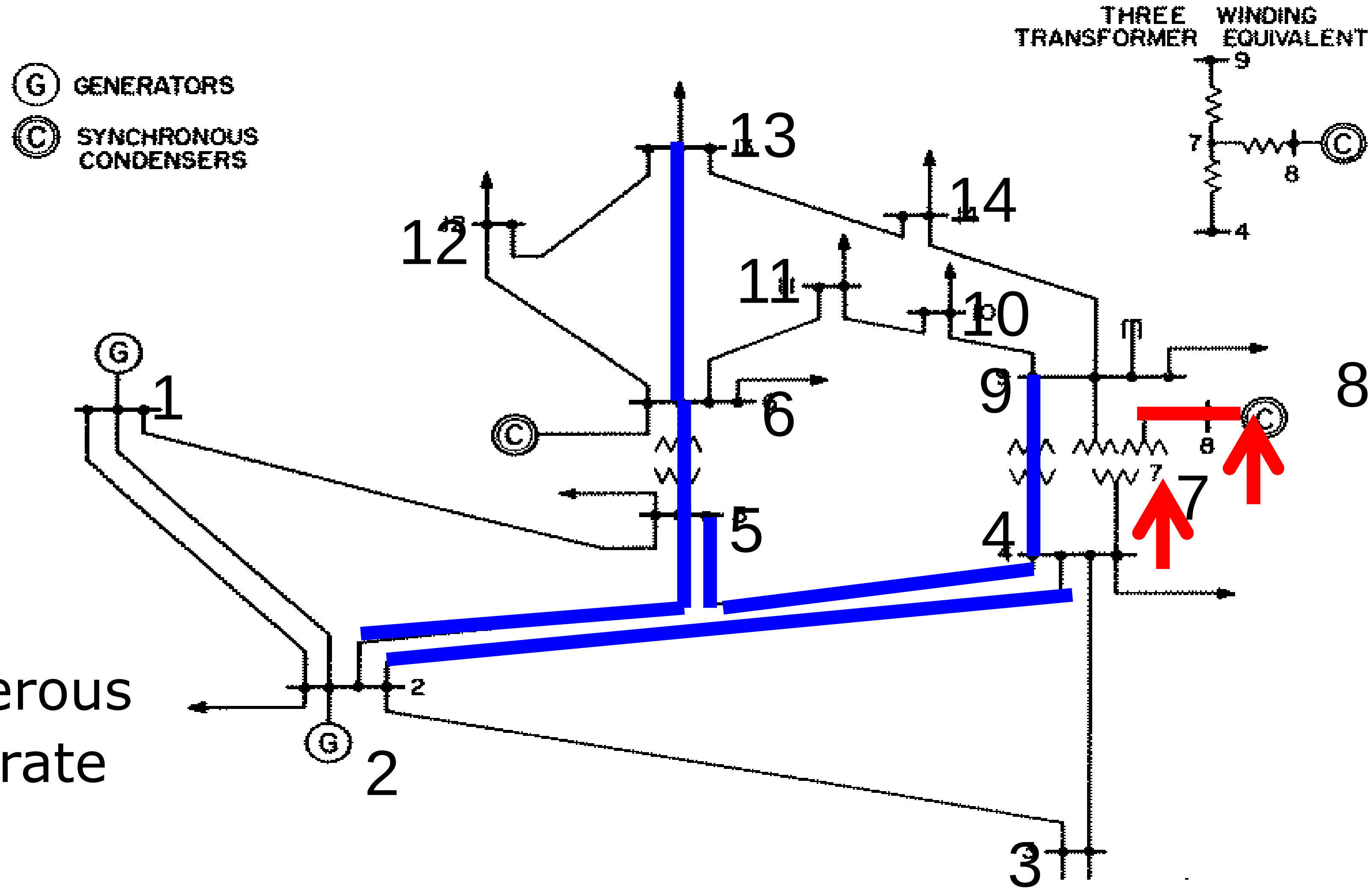
$$\min_{\Delta\theta} \quad \|H\Delta\theta\|_0$$
$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$



Minimum objective value = **security index** [Sandberg *et al*.]

Security index identifies network vulnerabilities

The Goal: Quantify Security



security

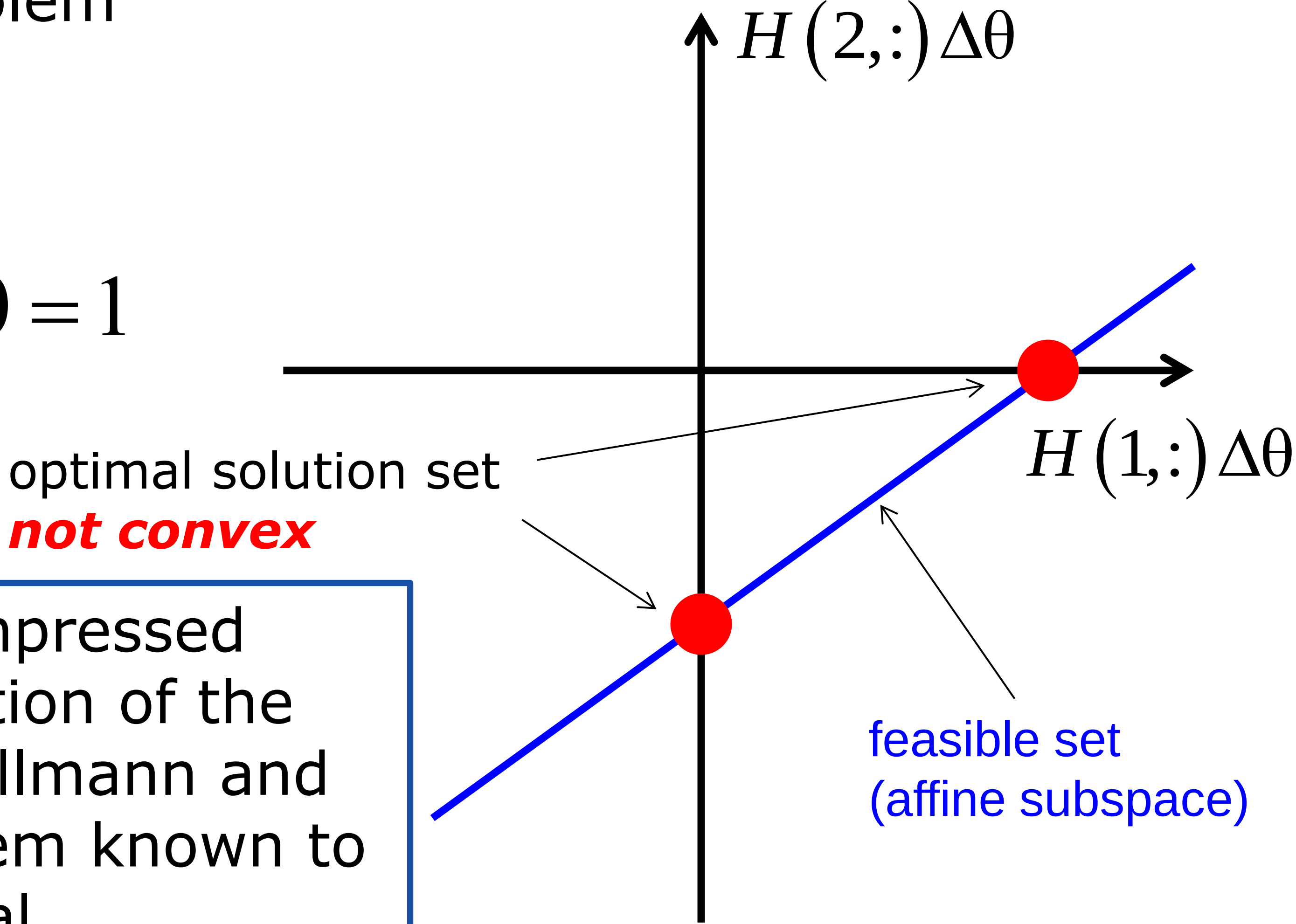
- dangerous
- moderate
- safe

Security Index – Cardinality Minimization

Security index problem

$$\min_{\Delta\theta} \quad \left\| H\Delta\theta \right\|_0$$

$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$



Closely related to compressed sensing and computation of the **cospark** of *H*, see [Tillmann and Pfetsch, 2005]. Problem known to be **NP-hard** in general.

ROYAL INSTITUTE OF TECHNOLOGY

Security index problem

$$\min_{\Delta\theta} \quad \left\| H \Delta\theta \right\|_0$$

$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

How to solve?

Security Index Computation – MILP

$$\min_{\Delta\theta} \quad \|H\Delta\theta\|_0$$
$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

$$\min_{\Delta\theta,y} \quad \sum_i y(i)$$
$$\text{s.t.} \quad -My \le H\Delta\theta \le My$$
$$H(k,:)\Delta\theta = 1$$
$$y(i) \in \{0,1\} \quad \forall i$$

MILP formulation

- Cardinality minimization problem
- Mixed integer linear program (MILP)
- **Exact** solution (solver: CPLEX)
- Solution algorithm **not scalable**

Security Index Computation – LASSO

$$\min_{\Delta\theta} \quad \left\| H\Delta\theta \right\|_1$$
$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

- Convex linear program (LP)
- Known as LASSO
- **Approximate** solution
- Less expensive to solve

$$\min_{\Delta\theta, y} \quad \sum_i y(i)$$
$$\text{s.t.} \quad -y \le H\Delta\theta \le y$$
$$H(k,:)\Delta\theta = 1$$
$$y(i) \in \Re \qquad \forall i$$

LP formulation

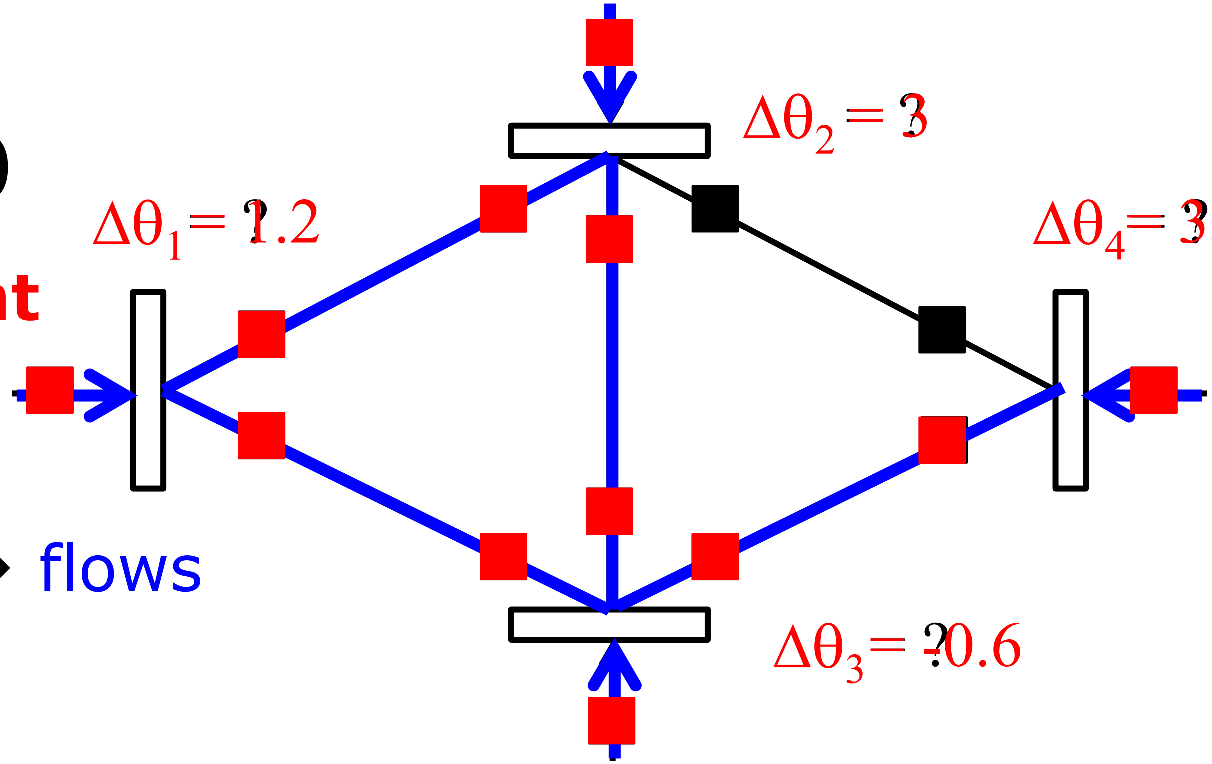
The Challenge

- Can we find solutions as **accurately** as MILP, and **faster** than LASSO?
- For general *H*, the answer is no (problem NP-hard)
- Let us exploit DC-power flow structure of *H* and make a **full measurement assumption**
- Specialize into graph problems with accurate and efficient algorithms

Graph Interpretation of Stealth Attack

Stealth attack $\Delta z = H\Delta\theta$

= **phase angle assignment**



Phase angle differences → flows

attack cost $\|H\Delta\theta\|_0$ = **# of meters with nonzero flows**

Binary Phases Assignment is Optimal

No phase angle difference → No flows ⇒ Attack cost = 0

⇒ No attack…

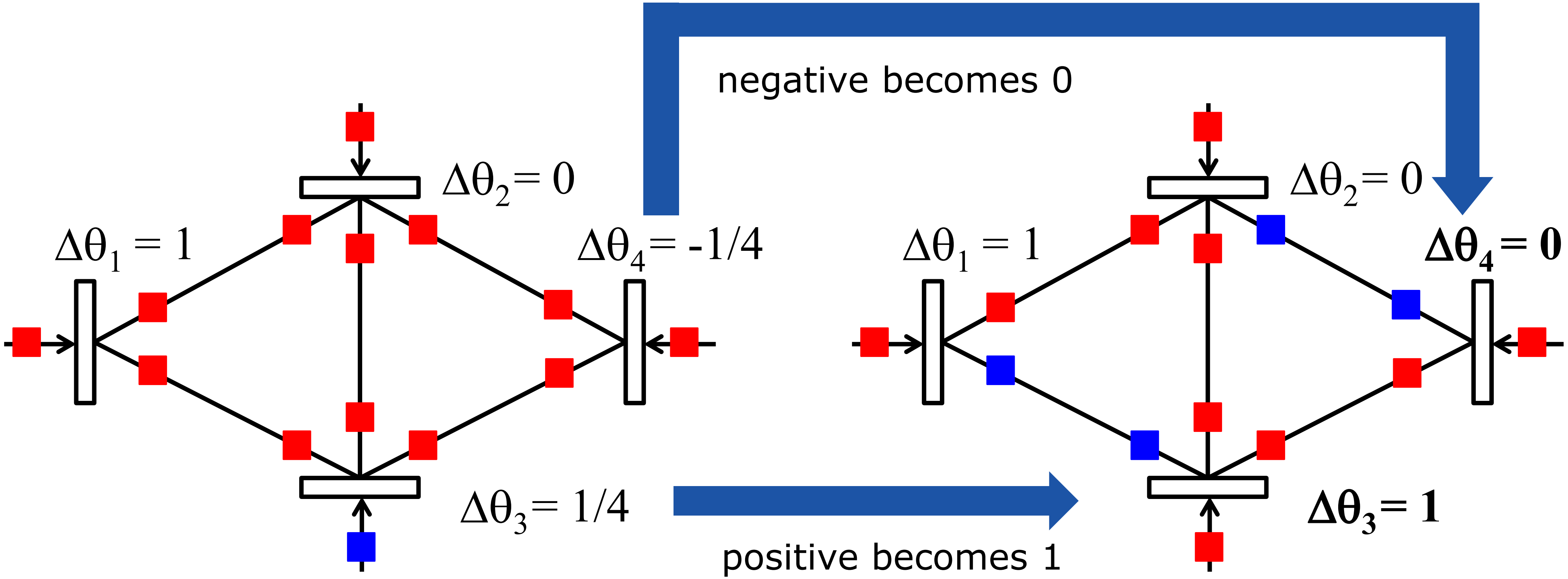
Next guess: (0,1) phase angle assignment?

Theorem: Optimal $\Delta\theta_i$ can be restricted to 0 or 1, for all i

$$\begin{aligned} \min_{\Delta\theta} \quad & \left\| H\Delta\theta \right\|_0 \\ \text{s.t.} \quad & H(k,:)\Delta\theta = 1 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \min_{\Delta\theta} \quad & \left\| H\Delta\theta \right\|_0 \\ \text{s.t.} \quad & H(k,:)\Delta\theta = 1 \\ & \Delta\theta_i \in \{0,1\} \end{aligned}$$

Binary Optimal Solution Justification

Can always find (0,1) feasible solution with no worst cost



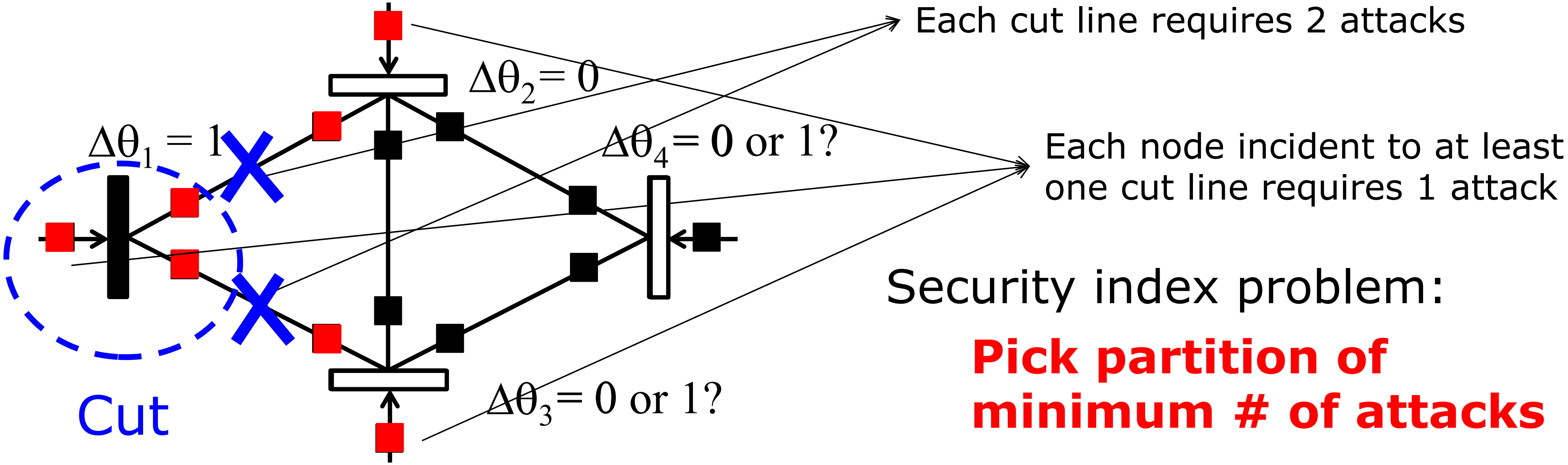
Cost = **13** attacks

Cost = **10** attacks

Reformulation as Node Partitioning

Optimal $\Delta\theta_i$ can be restricted to 0 or 1, for all i

Phase angle assignment becomes **node partitioning**



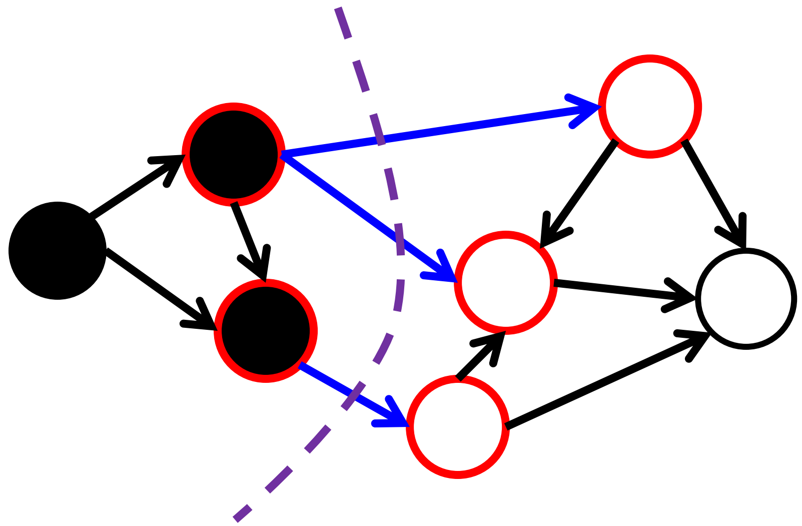
Security index problem:

Pick partition of minimum # of attacks

Security index problem

$$\min_{\Delta\theta} \quad \|H\Delta\theta\|_0$$
$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

Generalized Min Cut problem



How to solve generalized Min Cut?

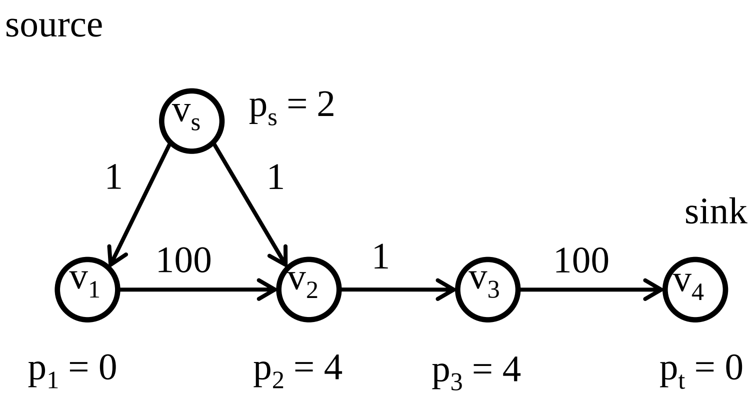
Standard Min Cut on Appended Graph

Generalized Min Cut = Standard Min Cut on **appended** graph

generalized min cut ⟺ standard min cut appended graph

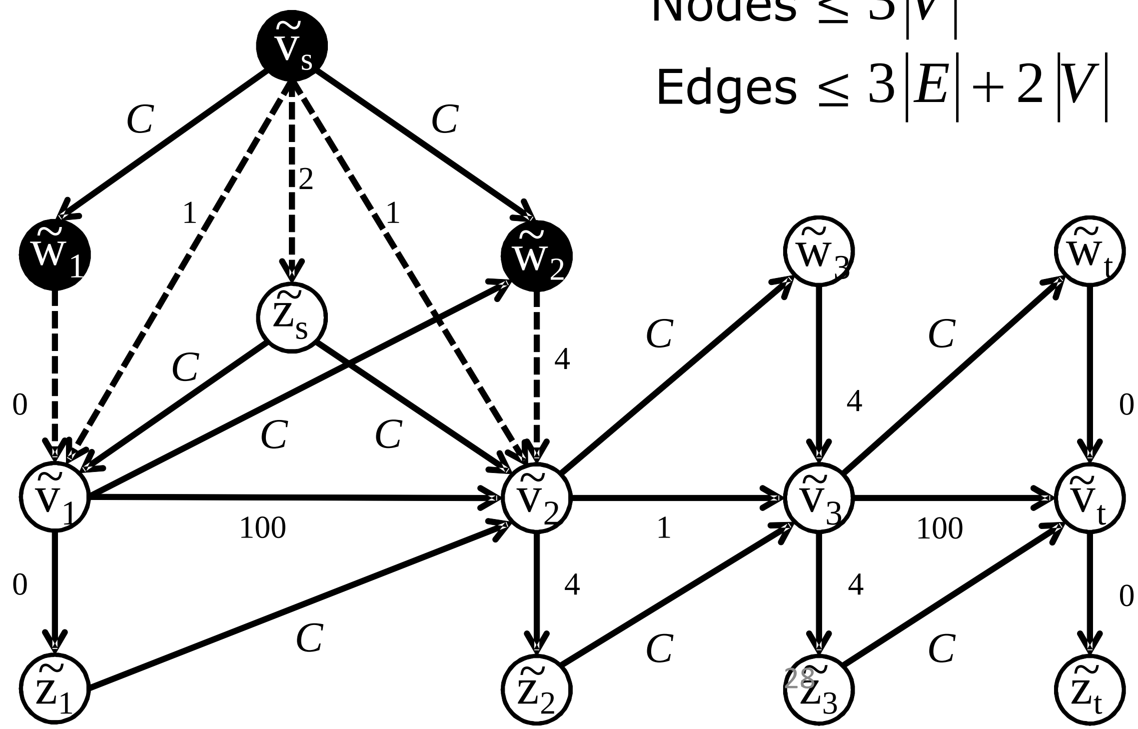
$|V|$ nodes

$|E|$ edges



Nodes $\leq 3|V|$

Edges $\leq 3|E| + 2|V|$



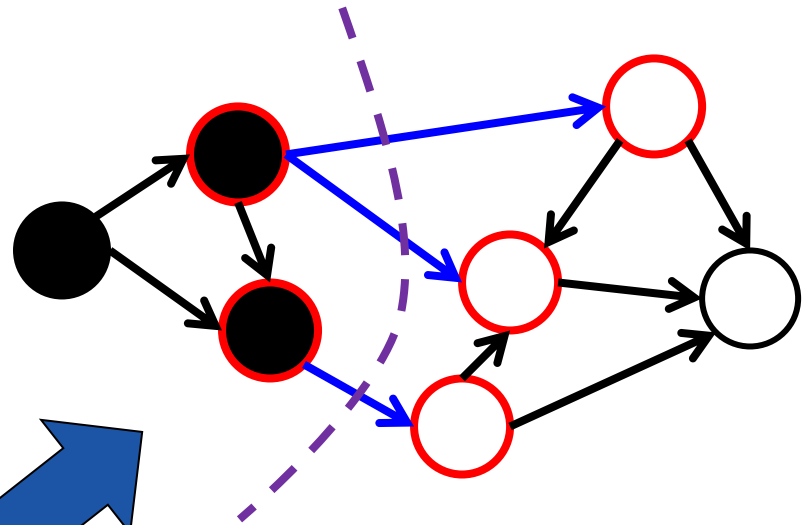
Security Index Problem – Summary

Security index problem

$$\min_{\Delta\theta} \quad \|H\Delta\theta\|_0$$

$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

Generalized Min Cut problem



Practical implications?

Standard Min Cut problem on an **appended** graph

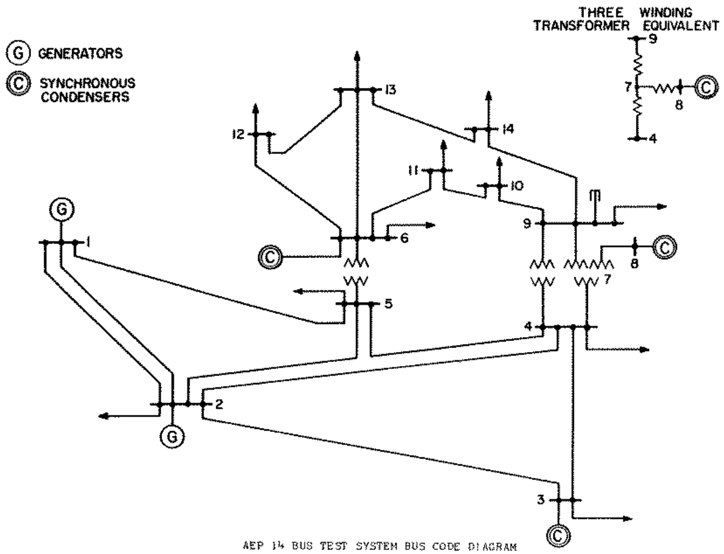
[Sou *et al.*, 2011]

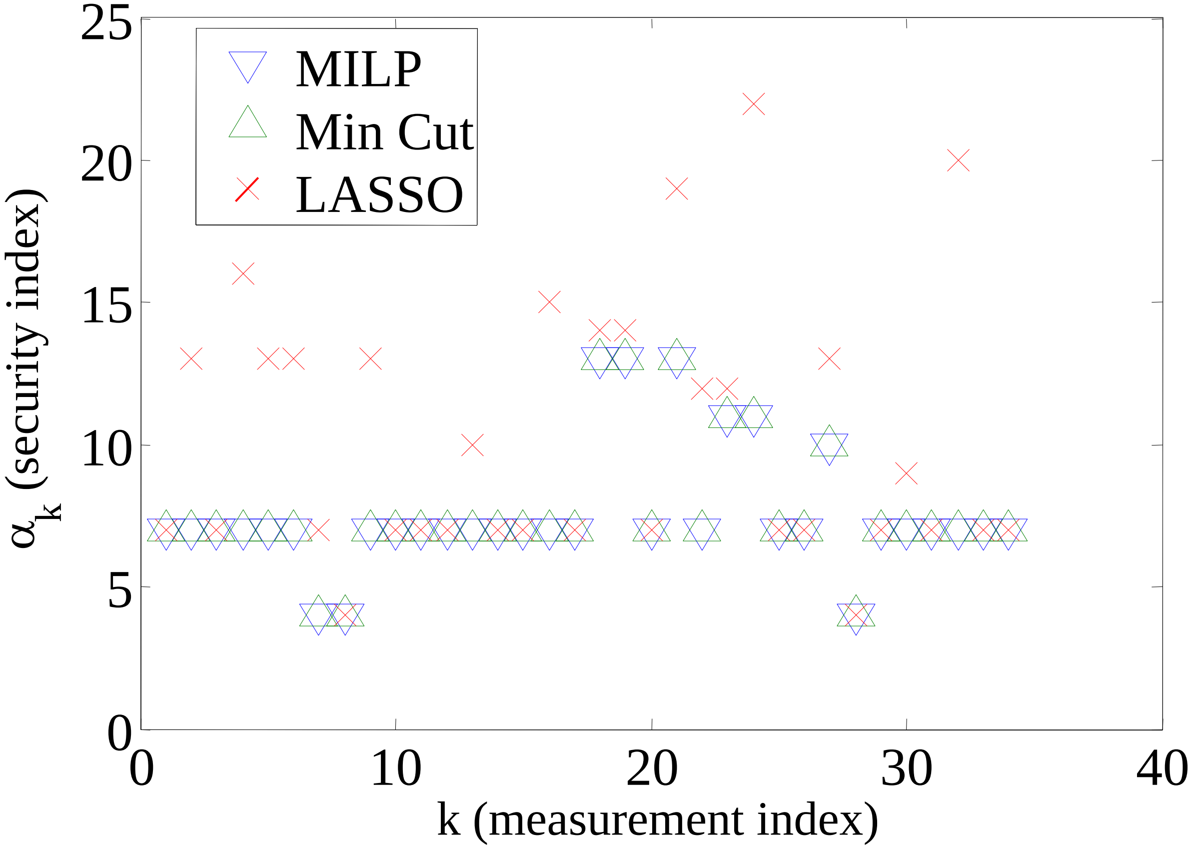
[Hendrickx *et al.*, 2013]

>> [maxflow,mincut] = max_flow(A,source,sink);

IEEE 14 Bus Benchmark Test Result

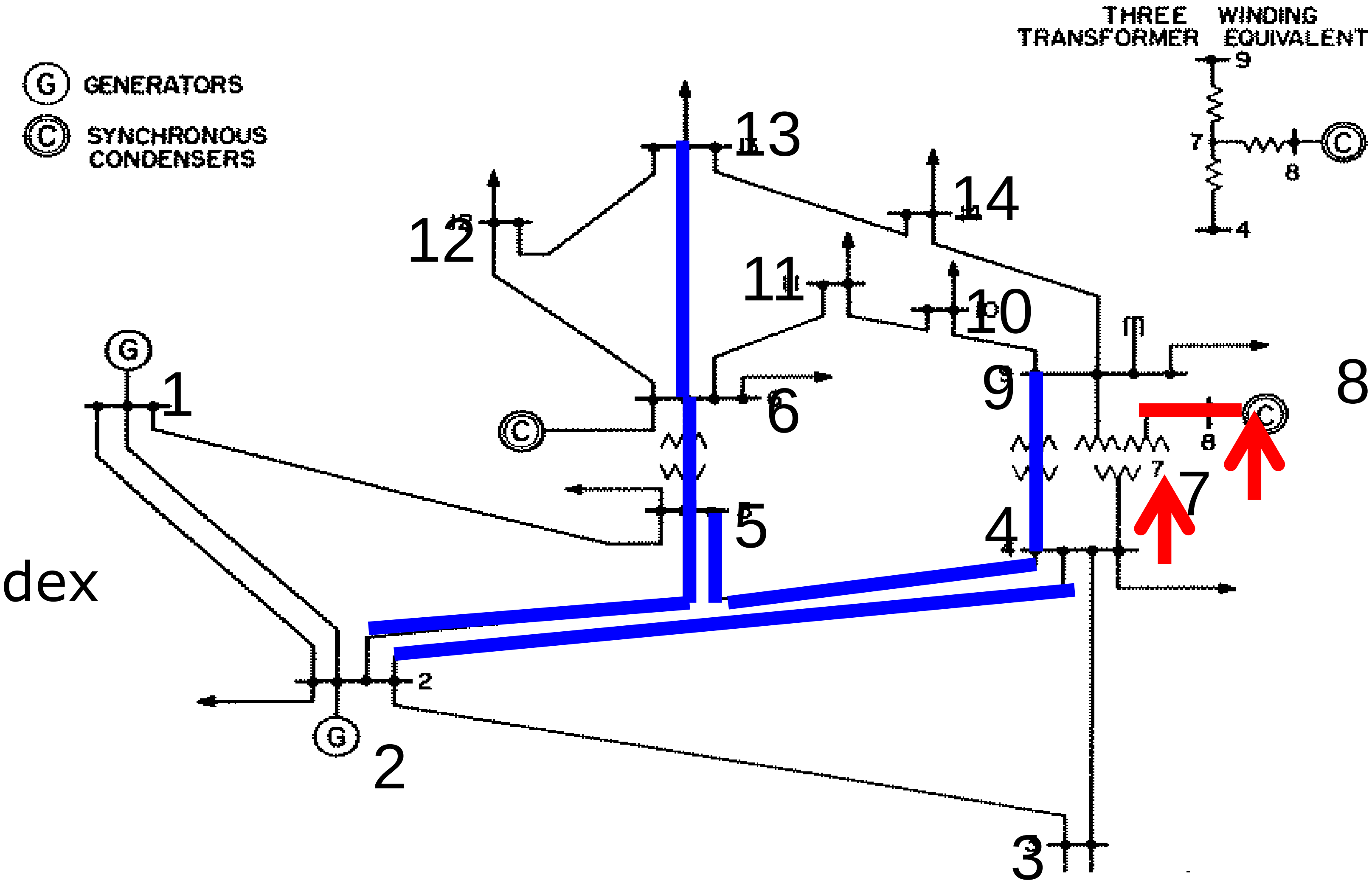
Security indices for all measurements





Solve time: MILP 1.1s; LASSO 0.6s; Min Cut 0.02s

IEEE 14 Bus Vulnerable Measurements



security index

4

7

≥ 10

ROYAL INSTITUTE OF TECHNOLOGY

IEEE 118, 300, 2383 Bus Benchmarks

Min Cut solution is **exact**

Solve time comparison:

Method/Case	118 bus	300 bus	2383 bus
MILP	763 sec	6708 sec	About 5.7 days
Min Cut	0.3 sec	1 sec	31 sec

What about LASSO (1-Norm Relaxation)?

$$\min_{\Delta\theta} \quad \|H\Delta\theta\|_1$$

$$\text{s.t.} \quad H(k,:)\Delta\theta = 1$$

We have seen LASSO relaxation in general yields non-optimal solution

Will LASSO ever work?

Yes, when H is **totally unimodular!** [Sou *et al*., 2013]

Totally Unimodular Matrices

A matrix is totally unimodular

= determinant of all square sub-matrices are -1,0,1

network incidence matrix

$$\begin{bmatrix} 1 & 1 & 0 & 0 & -1 \\ -1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 1 \end{bmatrix}$$

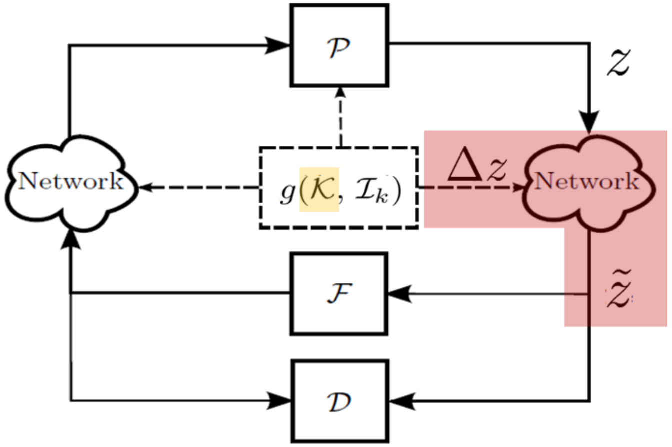
consecutive one matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Corresponds to full flow measurements
(no bus injection measurements)

Summary – Power Network State Estimation

- Adversary model
  - Induce measurement bias undetected
  - DC-power flow model known
  - Minimum disruption resources desired



- Security index problem yields lower bounds on required disruption resources. Suggests protection strategy [Vukovic *et al*., 2012]

- Security index computation in general NP-hard. Under appropriate assumptions graph Min Cut relaxation works very well

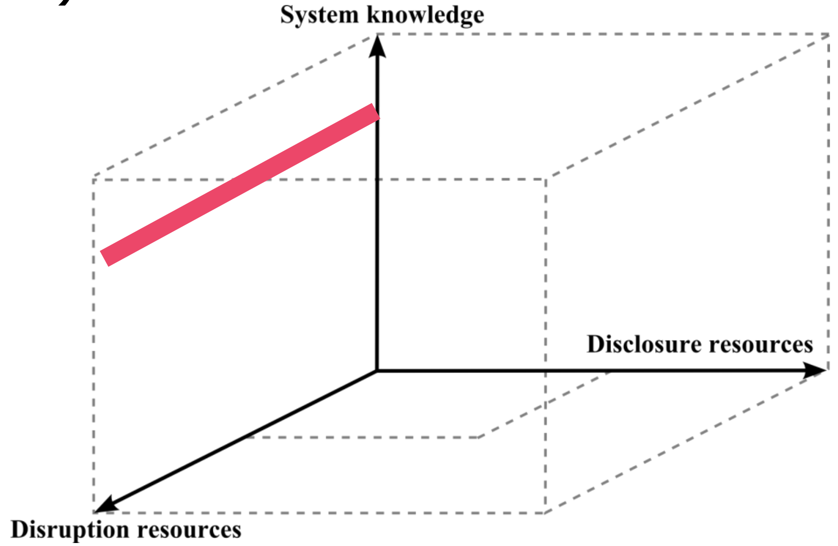
Outline

- Adversary models for networked control systems
- Application 1: Power network state estimation
- **Application 2: Wireless LQG-controlled quadruple tank**
  - **Max-impact/min-resource attacks**

Extension to Dynamical Systems

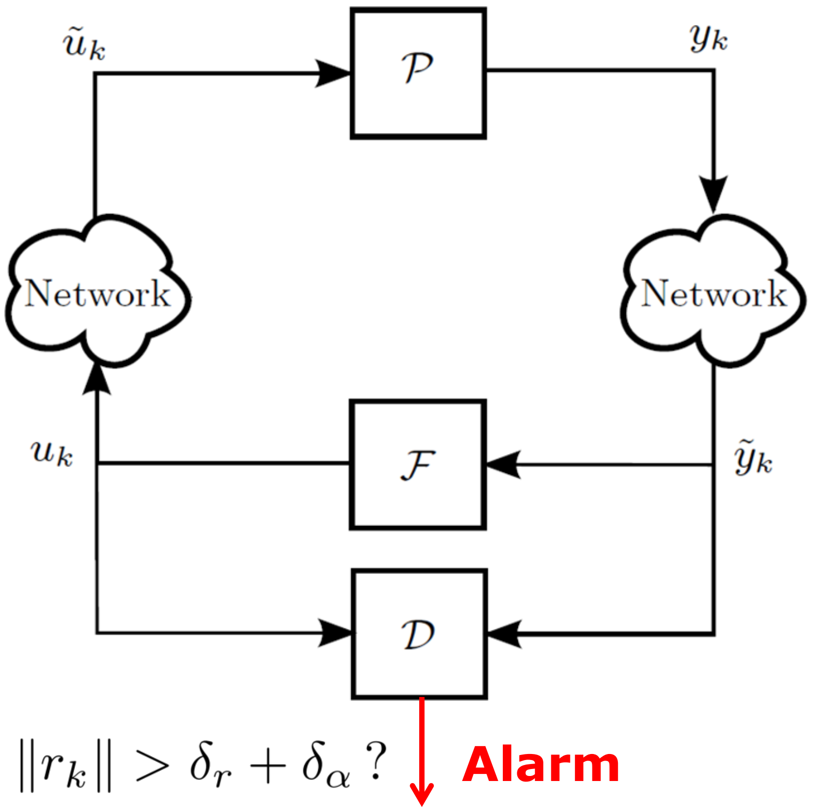
- Attacker needs to satisfy constraints not only across channels (*spatial dimension*) but also constraints across time (*temporal dimension*)

- Cases considered:
  1. Minimum resource attacks
  2. Maximum impact attacks
  3. **Maximum impact bounded resource attacks**



[Teixeira *et al.*, 2013]

Dynamical Networked Control System



- Physical Plant

$$\mathcal{P} : \begin{cases} x_{k+1} = Ax_k + B\tilde{u}_k + Gw_k \\ y_k = Cx_k + v_k \end{cases}$$

- Feedback Controller

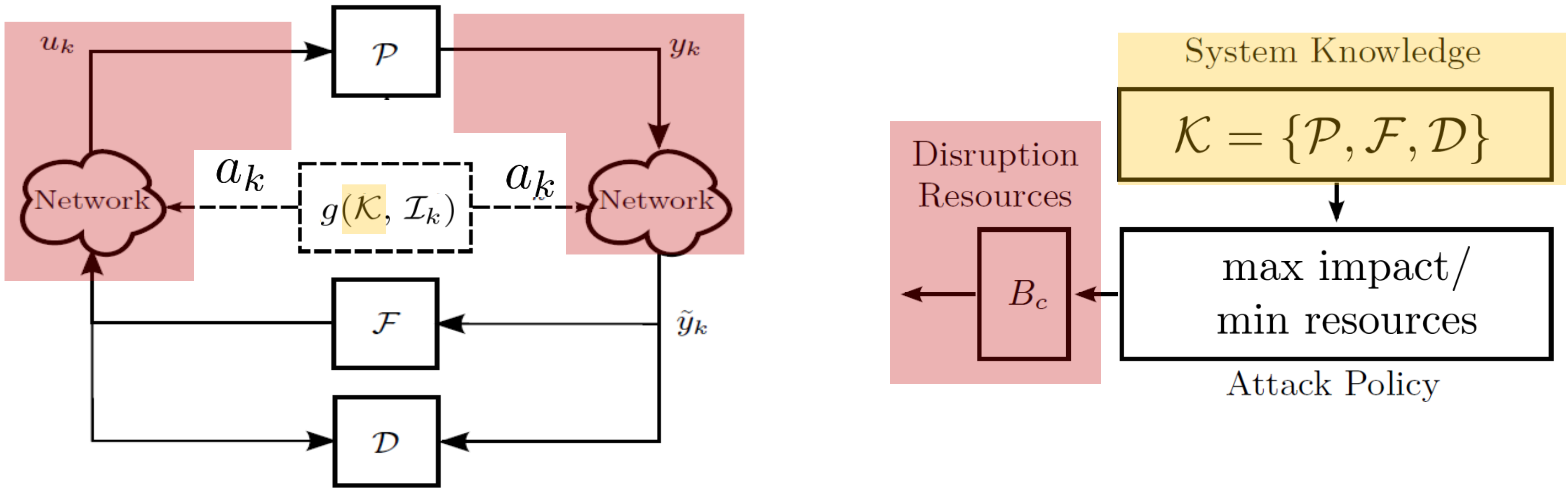
$$\mathcal{F} : \begin{cases} z_{k+1} = A_c z_k + B_c \tilde{y}_k \\ u_k = C_c z_k + D_c \tilde{y}_k \end{cases}$$

- Anomaly Detector

$$\mathcal{D} : \begin{cases} \hat{x}_{k|k} = A\hat{x}_{k-1|k-1} + Bu_{k-1} + K(\tilde{y}_k - \hat{y}_{k|k-1}) \\ r_k = V(\tilde{y}_k - \hat{y}_{k|k}) \end{cases}$$

  - Alarm triggered if $\|r_k\| > \delta_r + \delta_\alpha$

Adversary Model



- Adversary's goal is to force the process state into an unsafe region
- Attack should be stealthy, i.e., no alarms
- Adversary constrained by limited resources

The Dynamical Systems Case (1)

Dynamical anomaly detector for closed-loop system:

$$\begin{aligned} \xi_{k+1} &= \mathbf{A_e}\xi_k + \mathbf{B_e}a_k + \mathbf{G_e}w_k \\ r_k &= \mathbf{C_e}\xi_k + \mathbf{D_e}a_k + \mathbf{H_e}v_k \end{aligned}$$

Lift to time interval $[0, N]$
with zero-initial conditions and no noise:

$$\underbrace{\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}}_{\mathbf{r}} = \underbrace{\begin{bmatrix} \mathbf{D_e} & 0 & \dots & 0 \\ \mathbf{C_eB_e} & \mathbf{D_e} & \dots & 0 \\ \mathbf{C_eA_eB_e} & \mathbf{C_eB_e} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \mathbf{C_eA_e}^{N-1}\mathbf{B_e} & \mathbf{C_eA_e}^{N-2}\mathbf{B_e} & \dots & \mathbf{D_e} \end{bmatrix}}_{\mathcal{T}_r} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}}_{\mathbf{a}}$$

ROYAL INSTITUTE OF TECHNOLOGY

The Dynamical Systems Case (2)

Dynamics of plant and controller:

$$\begin{aligned} \eta_{k+1} &= \mathbf{A}\eta_k + \mathbf{B}a_k + \mathbf{G}w_k \\ x_k &= \mathbf{C}\eta_k + \mathbf{D}a_k + \mathbf{H}v_k \end{aligned}$$

Lift to time interval $[0, N]$
with zero-initial conditions and no noise:

$$\underbrace{\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} \mathbf{D} & 0 & \dots & 0 \\ \mathbf{CB} & \mathbf{D} & \dots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \mathbf{CA}^{N-1}\mathbf{B} & \mathbf{CA}^{N-2}\mathbf{B} & \dots & \mathbf{D} \end{bmatrix}}_{\mathcal{T}_x} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}}_{\mathbf{a}}$$

Max Impact/Bounded Resource Attack

$$\max_{\mathbf{a}} \|\mathcal{T}_x \mathbf{a}\|_\infty$$ (physical impact)

s.t.

$$\|\mathbf{r}\|_\infty = \|\mathcal{T}_r \mathbf{a}\|_\infty \leq \delta_\alpha$$ (residual in detector)

$$\|h_p(\mathbf{a})\|_0 \leq \epsilon$$ (# channels attacked)

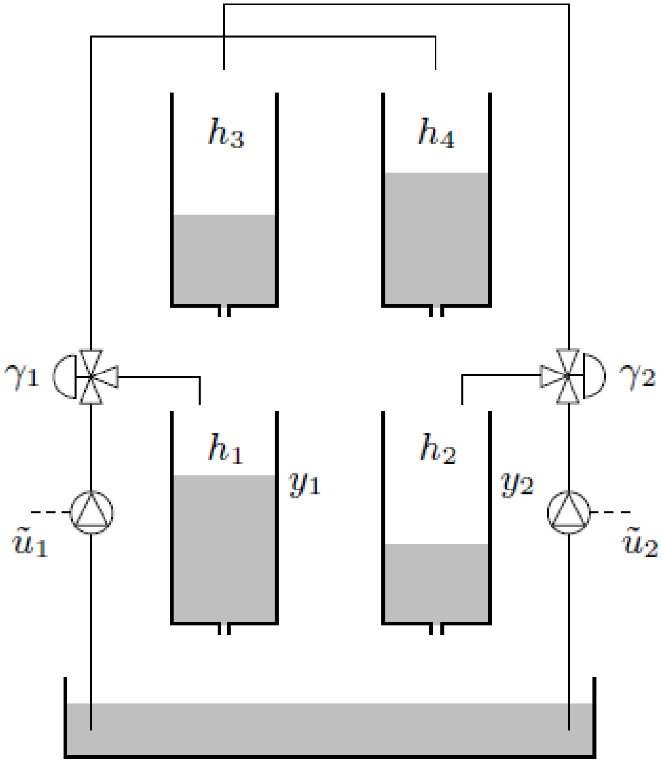
- Maximize impact (push $\|\mathbf{x}\|_\infty$ far away from equilibrium)
- No alarms (threshold δ_α)
- Attack no more than ϵ channels

$$h_p(\mathbf{a}) = [\|\mathbf{a}_{(1)}\|_{\ell_p}, \ldots, \|\mathbf{a}_{(i)}\|_{\ell_p}, \ldots, \|\mathbf{a}_{(q_a)}\|_{\ell_p}]$$

- Mixed Integer Linear Program (MILP)

[Teixeira *et al.*, 2013]

Numerical Example



$$\dot{h}_1 = -\frac{a_1}{A_1}\sqrt{2gh_1} + \frac{a_3}{A_1}\sqrt{2gh_3} + \frac{\gamma_1 k_1}{A_1}u_1,$$
$$\dot{h}_2 = -\frac{a_2}{A_2}\sqrt{2gh_2} + \frac{a_4}{A_2}\sqrt{2gh_4} + \frac{\gamma_2 k_2}{A_2}u_2,$$
$$\dot{h}_3 = -\frac{a_3}{A_3}\sqrt{2gh_3} + \frac{(1-\gamma_2)k_2}{A_3}u_2,$$
$$\dot{h}_4 = -\frac{a_4}{A_4}\sqrt{2gh_4} + \frac{(1-\gamma_1)k_1}{A_4}u_1,$$

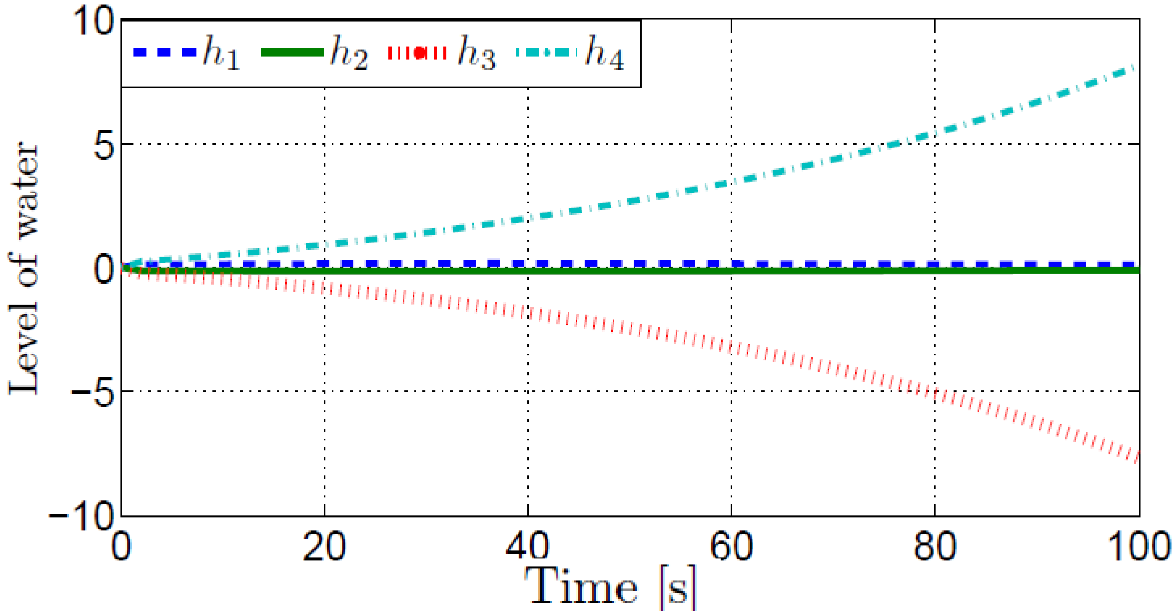
- Wireless LQG controller
- 4 channels: 2 actuators and 2 measurements
- Minimum phase or non-minimum phase depending on $\gamma_1,\ \gamma_2$

Numerical Example (Non-Min Phase)

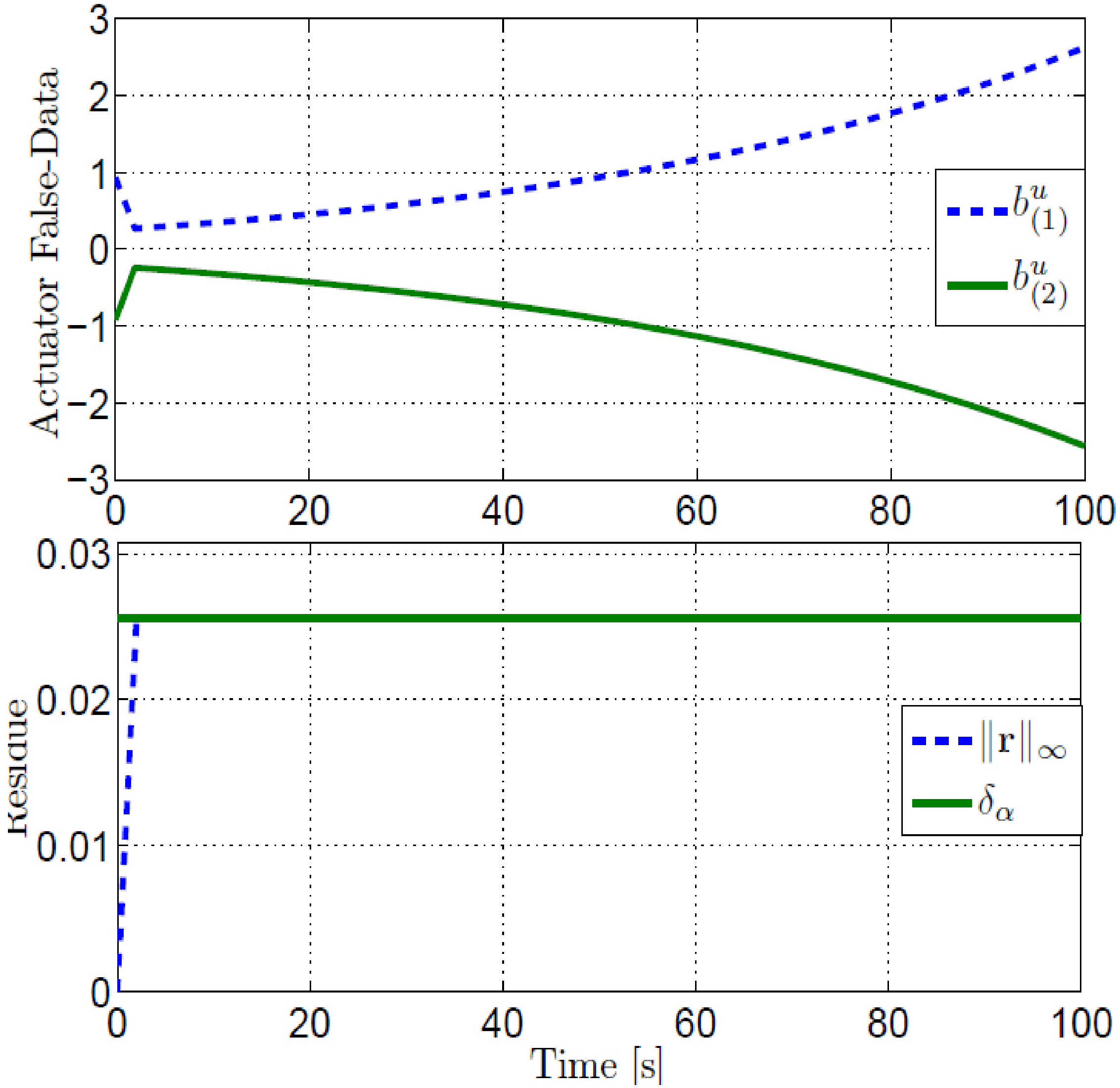
Values of $\|\mathbf{x}\|_2$ for max impact/bounded resource attack

$\delta_\alpha = 0.15$

	$\|h_p(\mathbf{a})\|_0$			
	1	**2**	**3**	**4**
Minimum phase	1.15	140.39	∞	∞
Non-minimum phase	2.80	689.43	∞	∞



Numerical Example (Non-Min Phase)



ROYAL INSTITUTE OF TECHNOLOGY

Numerical Example

- Maximum Impact/Bounded Resource attack illustrated
- 2 channels allowed: MILP selects the actuators
- 3-4 channels allowed: Unbounded impact (any attack on actuators can be hidden by corrupting 2 measurements)
- Infinity norm criteria yields more aggressive attack than 2-norm criteria (bounds get saturated)
- Not surprisingly, non-min phase plant more sensitive

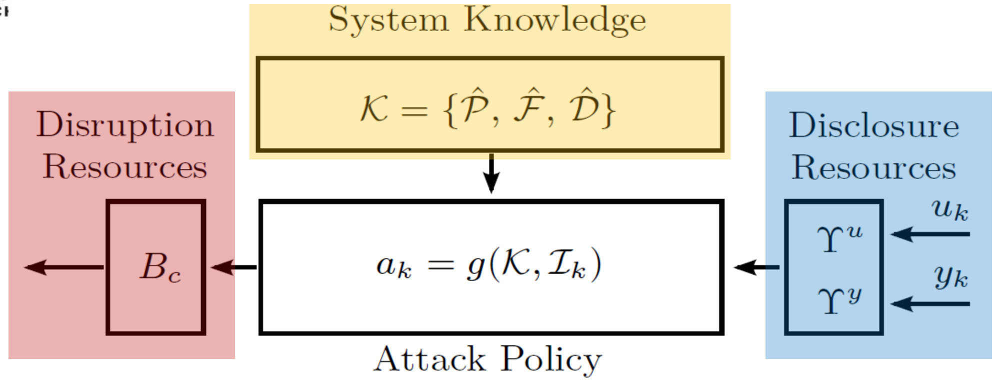
Steady-State Attacks

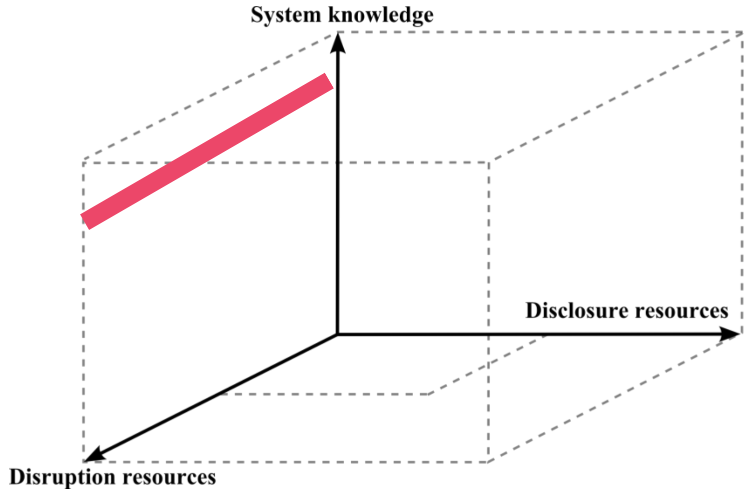
- Consider attacks over $[0, N]$ where
  - $N \to \infty$
  - $a_k = ge^{i\omega k}, \quad \omega \in \mathbb{R},\ g \in \mathbb{C}^{q_a}$ (sinusoidal attacks)

- Similar analysis carries through but make substitutions
  - $\mathcal{T}_r \to G_r(e^{i\omega})$
  - $\mathcal{T}_x \to G_x(e^{i\omega})$

- Yields worst-case attack frequency ω etc.

Summary





- Tools for quantitative trade-off analysis between attacker's impact and resources, also important for cyber defense prioritization

- For dynamical systems there are *temporal* as well as *spatial (channel) constraints* for attacker to fulfill
  - Enforced through lifting and frequency-response models
  - Solved using MILP. No well-working relaxation known by us

ROYAL INSTITUTE OF TECHNOLOGY

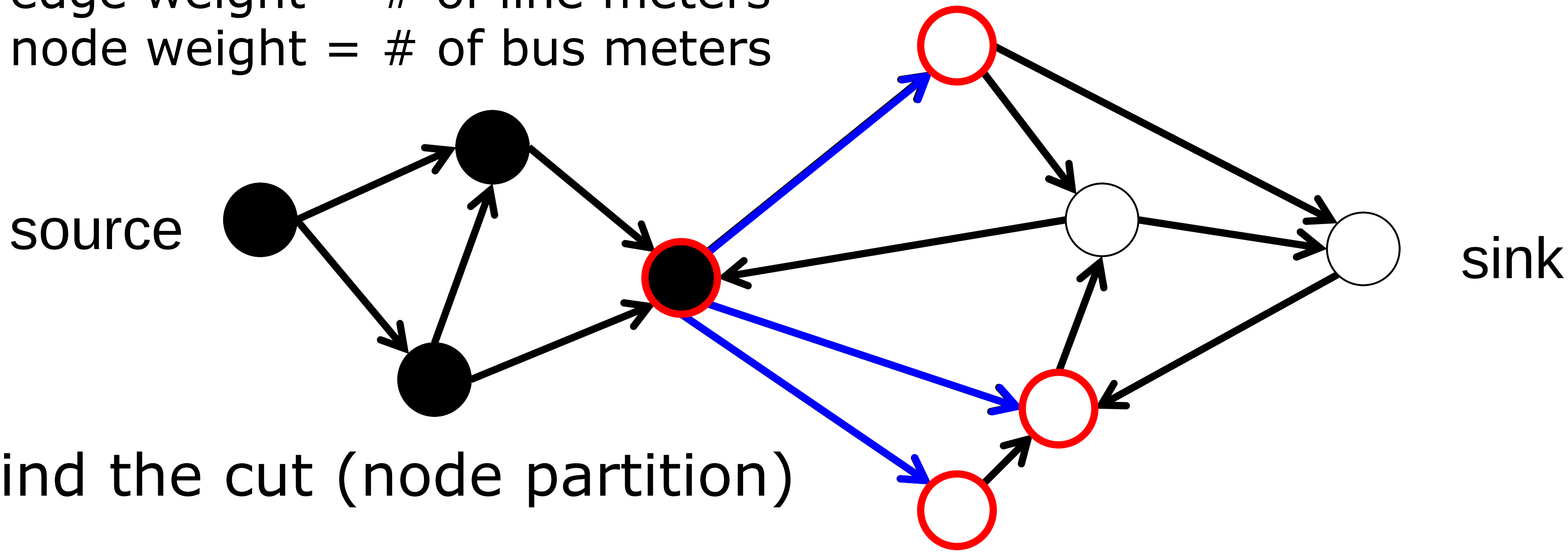
References

- Adversary models and quadruple tank process
  - Teixeira *et al*., "Attack models and scenarios for networked control systems", Proc. of HiCoNS, ACM, 2012
  - Teixeira *et al*., "*Quantifying Cyber-Security for Networked Control Systems*", Workshop on Control of Cyber-Physical Systems, Springer Verlag, 2013 (to appear)
- The security index problem
  - Sandberg *et al*., "On Security Indices for State Estimators in Power Networks", Preprints of 1st workshop on Secure Control Systems, CPSWEEK, 2010
  - Vukovic *et al.,* "Network-aware Mitigation of Data Integrity Attacks on Power System State Estimation", IEEE JSAC, 2012
- Efficient computation, and Min Cut relaxation
  - Sou *et al.*, "On the Exact Solution to a Smart Grid Cyber-Security Analysis Problem", IEEE Trans. on Smart Grid, 2013
  - Hendrickx *et al.,* "Efficient Computations of a Security Index for False Data Attacks in Power Networks", arXiv:1204.6174

Generalized Min Cut with Costly Nodes

Focus on directed graph (undirected = bi-directed)

edge weight = # of line meters
node weight = # of bus meters

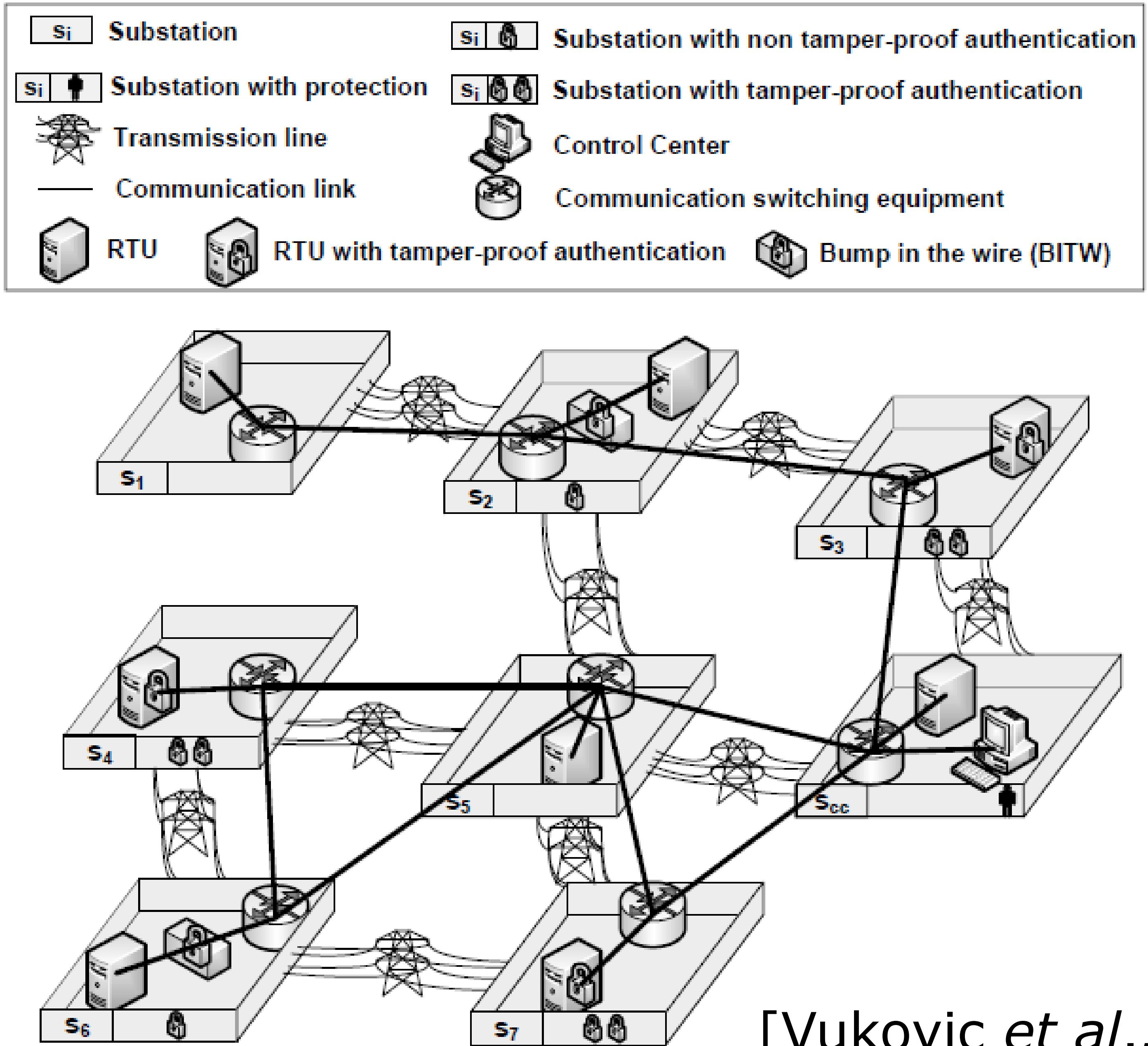


Find the cut (node partition)

to minimize weights of cut edge + incident node weights…

Generalization of standard Min Cut!

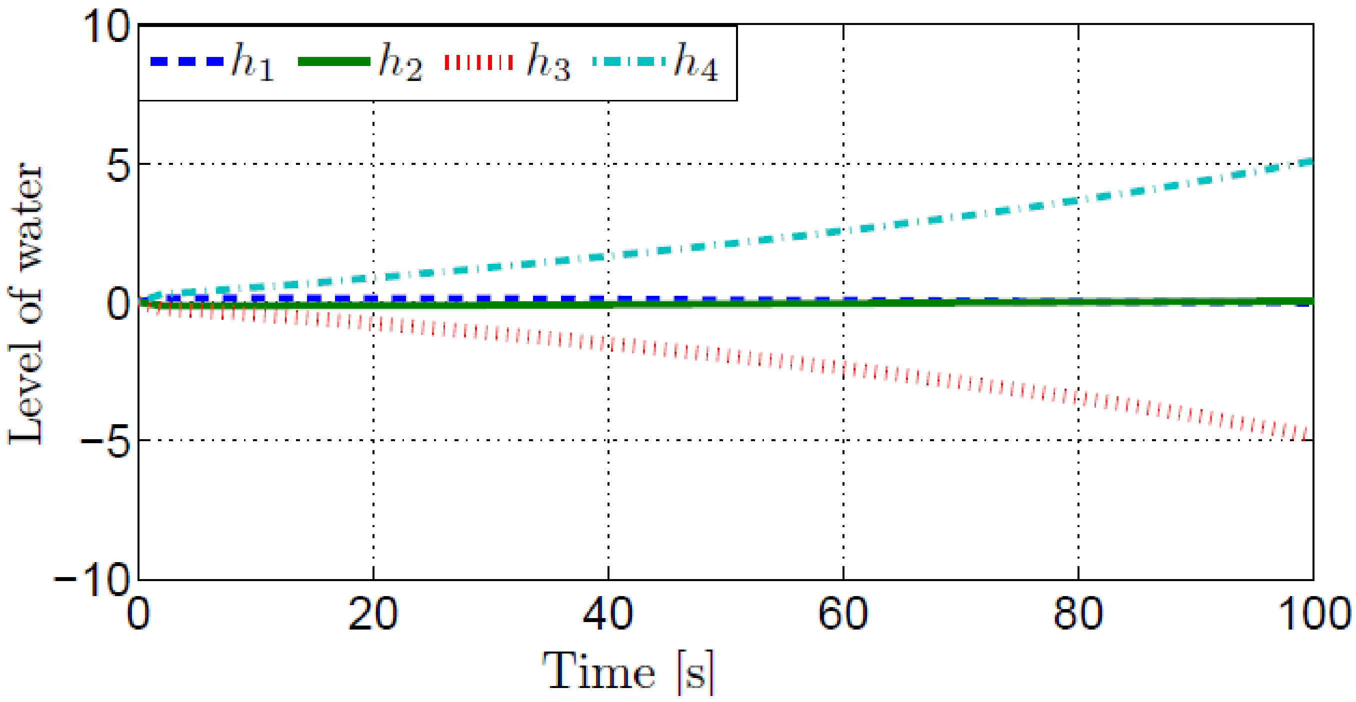
The Network: SCADA System



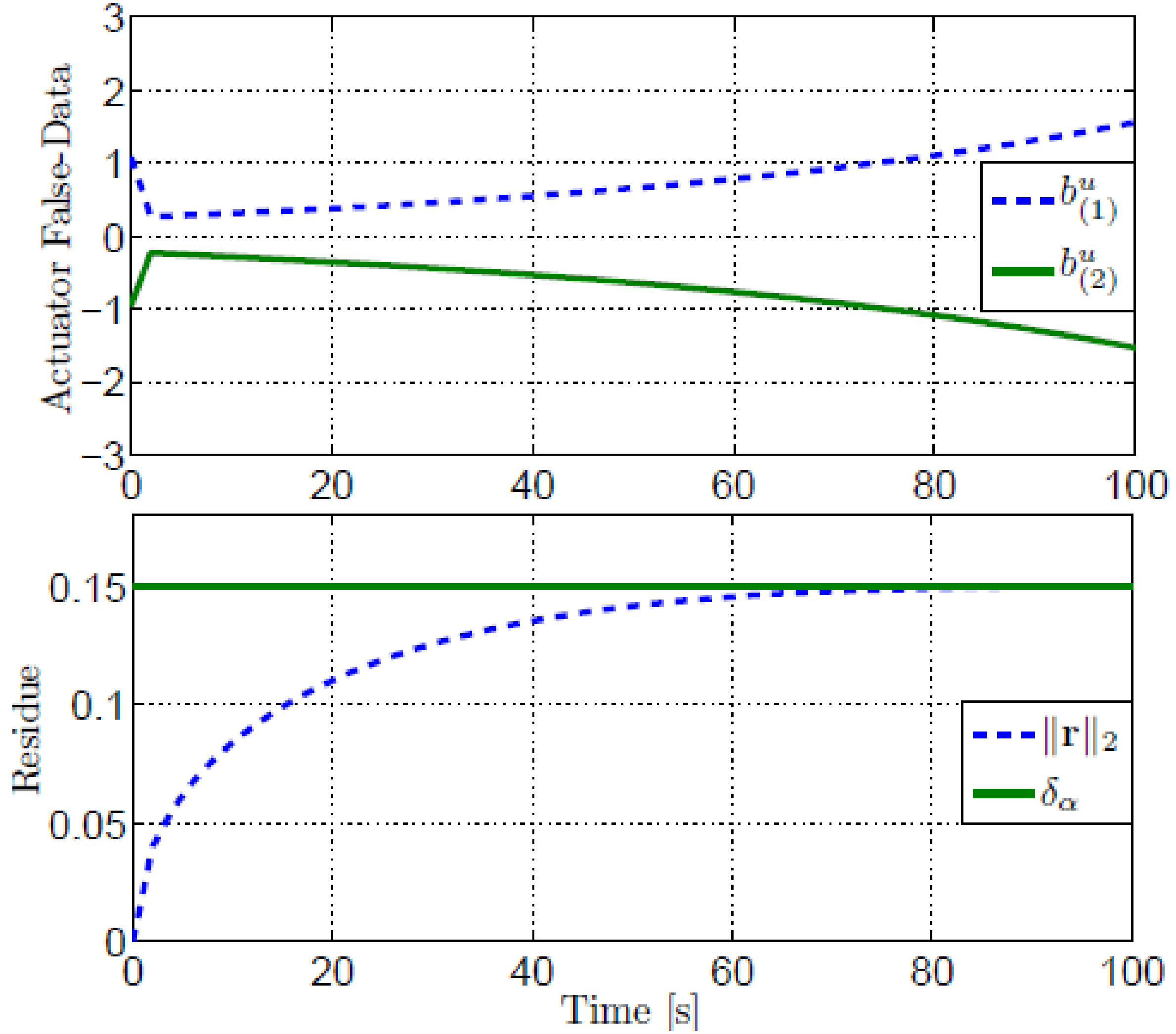
[Vukovic *et al*., 2012]

ROYAL INSTITUTE OF TECHNOLOGY

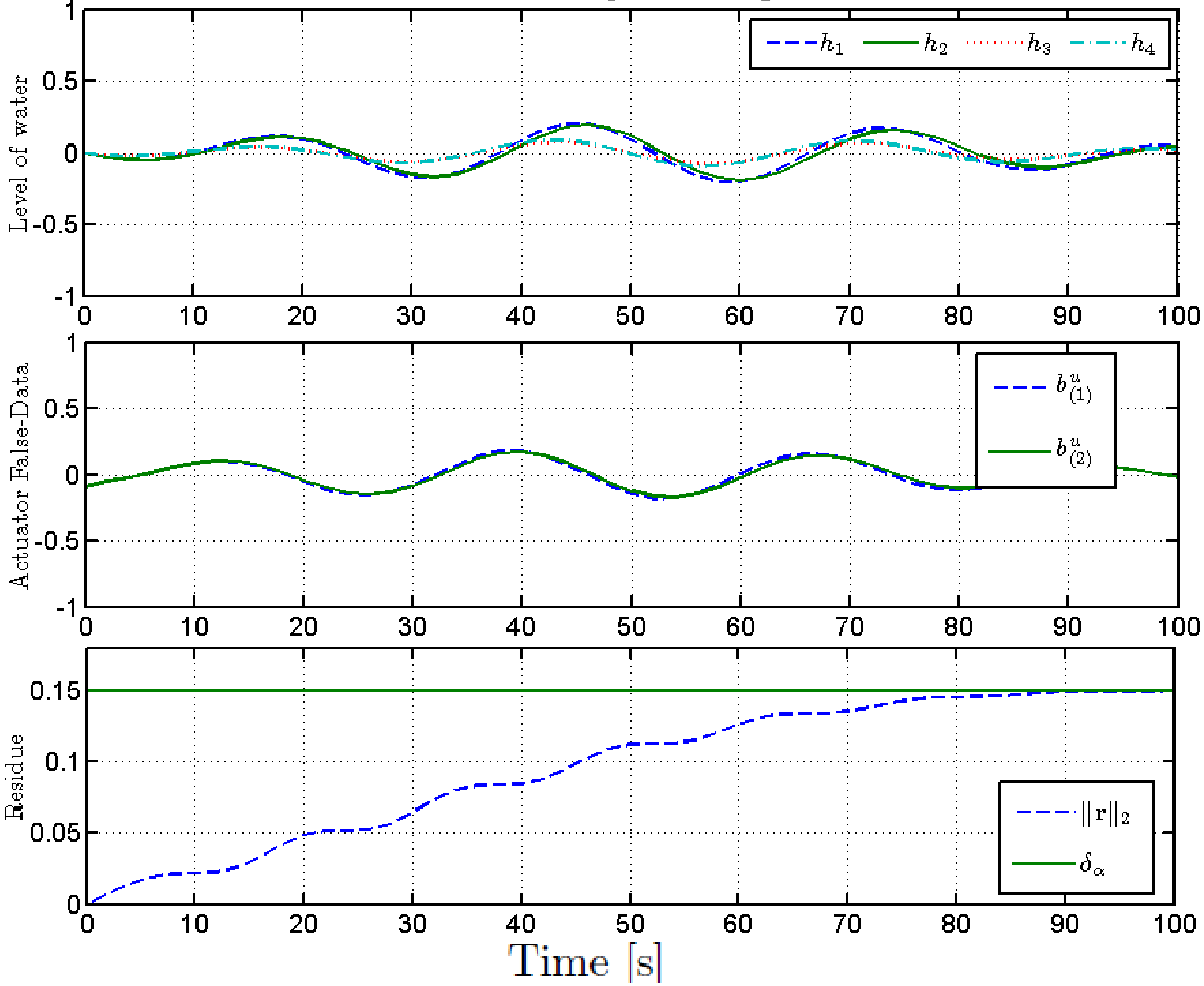
Numerical Example (2-Norm, Non-Min Phase)



Numerical Example (2-Norm, Non-Min Phase)



Numerical Example (Min Phase 2-norm)



Numerical Example (Min Phase inf-norm)