



ROYAL INSTITUTE  
OF TECHNOLOGY

# Secure Control and Applications in Power Systems

**Henrik Sandberg**

Automatic Control Lab, ACCESS Linnaeus Centre,  
KTH Royal Institute of Technology, Sweden

## Contributions From:

### **KTH**

André Teixeira

Kin Cheong Sou

György Dán

Karl Henrik Johansson

### **ETH Zürich**

Peyman Mohajerin

Esfahani

Maria Vrakopoulou

John Lygeros

Göran Andersson

### **UCSB**

Fabio Pasqualetti

Florian Dörfler

Francesco Bullo

Tutorial lecture, Cyber-Secure Control, ECC 2013  
Zürich, July 18, 2013

# Goals of Lecture

- To introduce centralized control and monitoring systems in transmission power grids (SCADA/EMS)
- To cover three scenarios:
  - Attacking/defending steady-state monitoring systems
  - Attacking/defending dynamic monitoring systems
  - Attacking/defending automatic generation control


*(Caveat: Other relevant scenarios can be found in the literature. This is an illustrative sample of available results)*
- To illustrate how control engineering can contribute to cyber-physical security in power systems

# Why We Should Care

- Northeast U.S. Blackout of August 14<sup>th</sup>, 2003: 55 million people affected
- Software bug in energy management system stalled alarms in energy management systems for over an hour
- Cyber attacks against the power network control center systems pose a real threat to society



# Background

- Modern society increasingly dependent on critical infrastructures
  - One of the most important is **electrical power system**
  - Operation and management of electrical power system based on computerized control and telecommunications (SCADA)
  - Keeping control system secure and resilient to external attacks vital for uninterrupted service
  - Critical component: Interaction between
    - Physical power flows
    - IT infrastructure
    - Human operators
-  **Cyber-physical system**
- (SCADA = Supervisory Control and Data Acquisition)

# Motivation



Attack



SCADA system



Power network

## Security issues

Power system: susceptible to operational errors and external attacks

Smart grid technology makes the system even more vulnerable



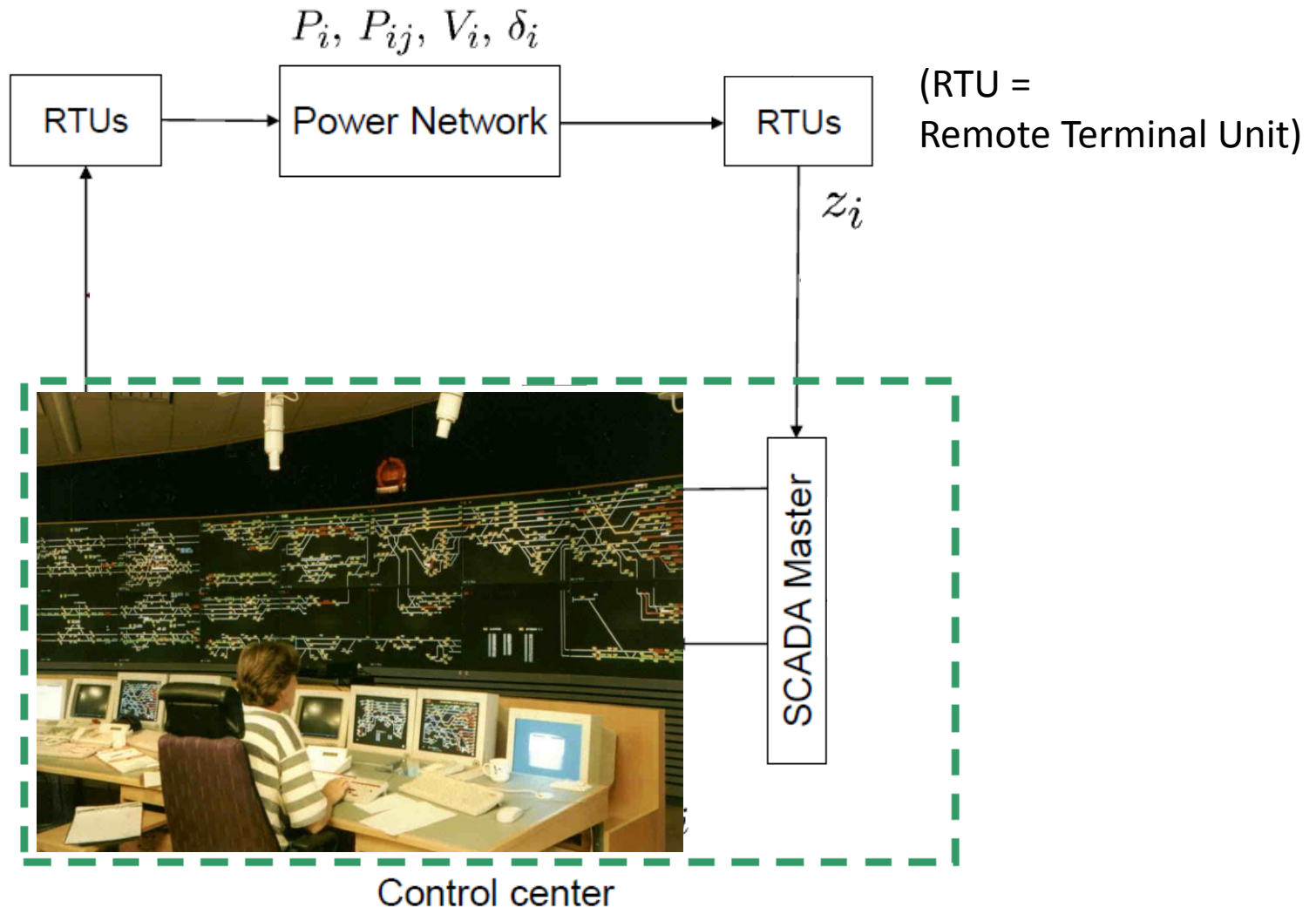
Societal cost

# Why Control Engineering?

1. Cyber security
  - Can attacker gain access to critical data streams?
  - Can they feed false data?
2. Impact analysis
  - What will be the effect?
  - How will human operators react?
3. Countermeasures
  - Can we tell something is wrong?
  - What can we do about it?

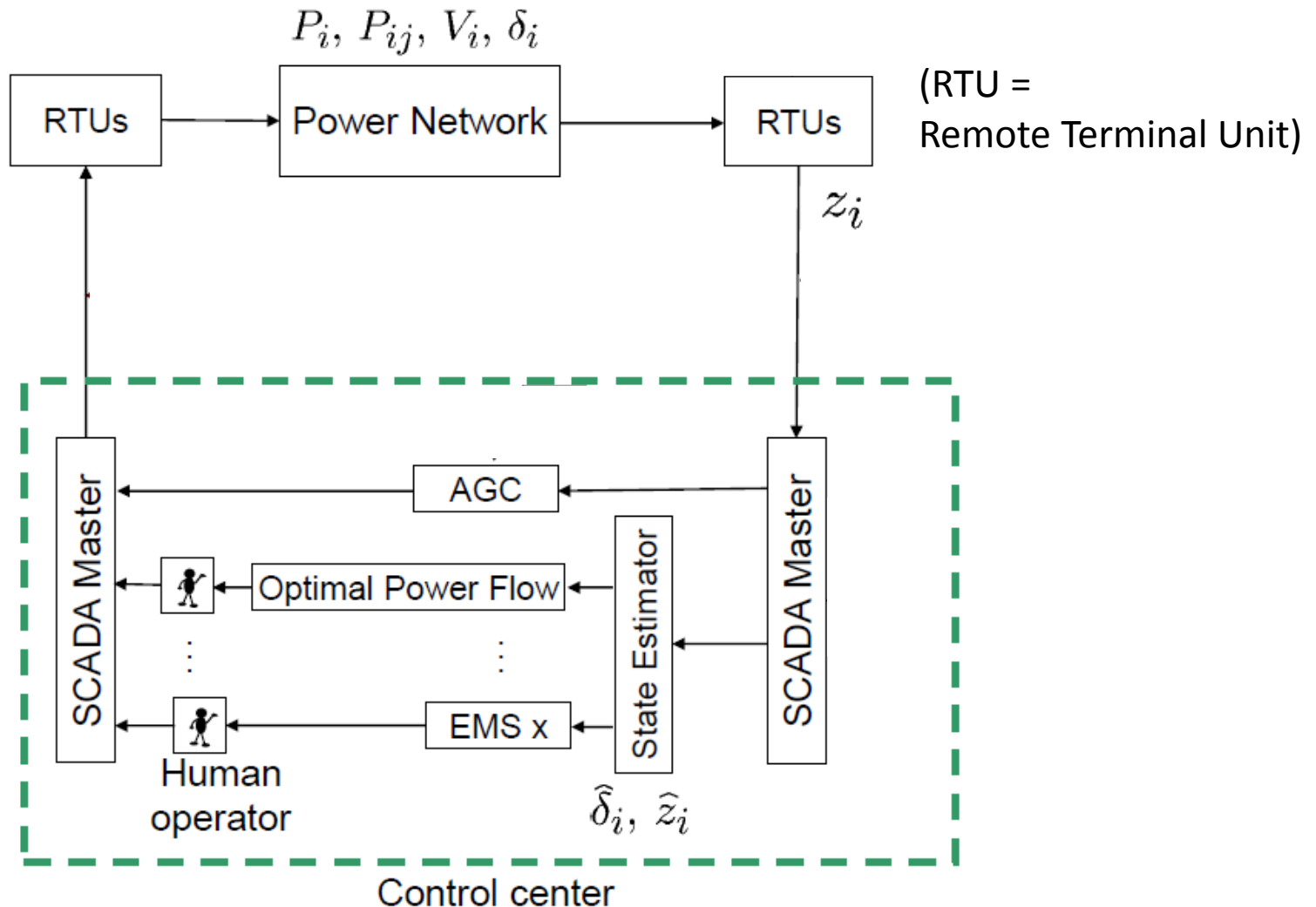
Control engineering central to items 2. and 3.

# The Cyber-Physical-Human Loop



(SCADA/EMS = Supervisory Control and Data Acquisition/Energy Management Systems)

# The Cyber-Physical-Human Loop



(SCADA/EMS = Supervisory Control and Data Acquisition/Energy Management Systems)



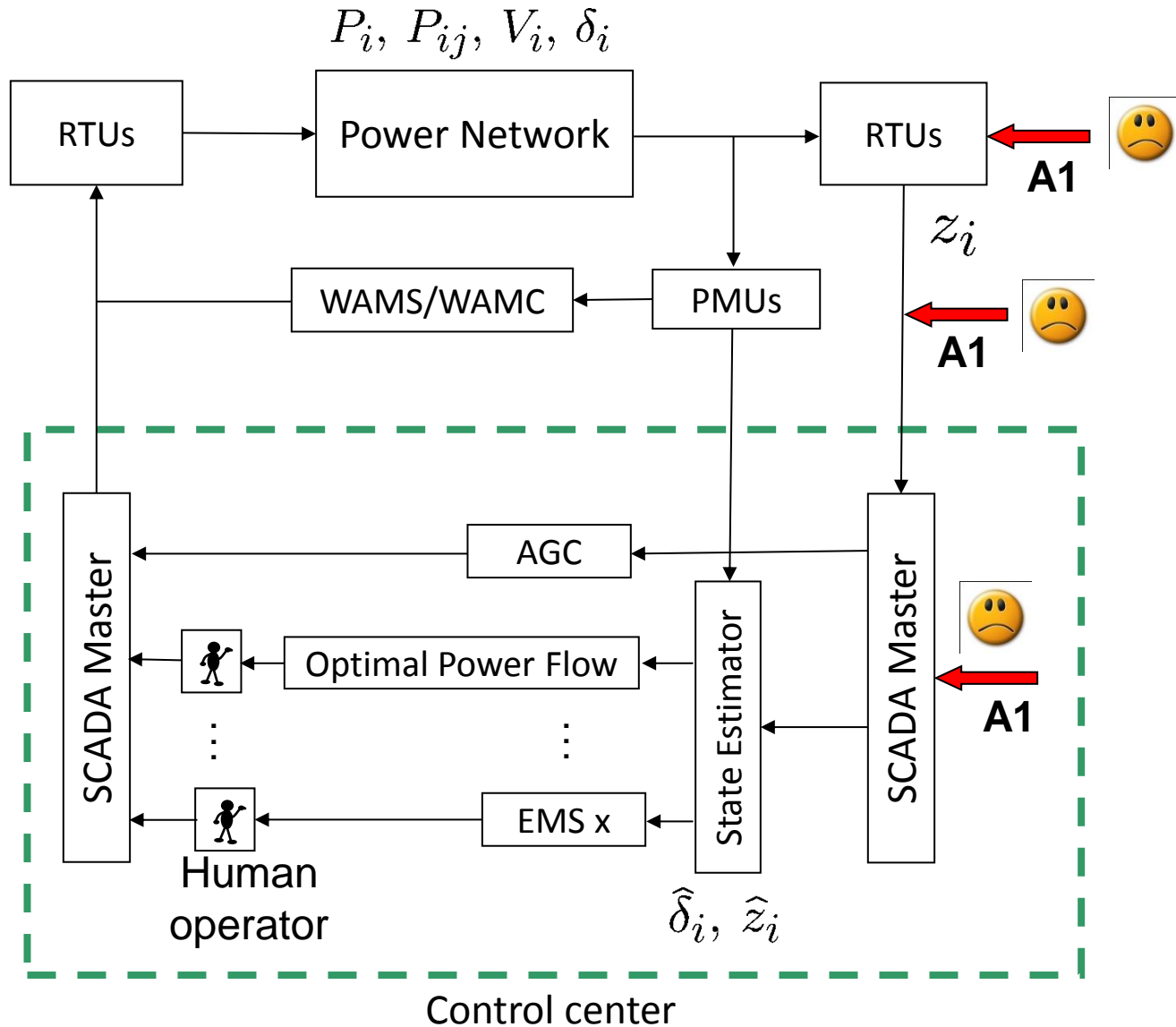
# Three Attack/Protection Scenarios

- Scenario 1: Steady-state monitoring systems
  - Time scale: minutes-hours
- Scenario 2: Dynamic monitoring systems
  - Time scale: seconds
- Scenario 3: Automatic Generation Control (AGC)
  - Time scale: seconds

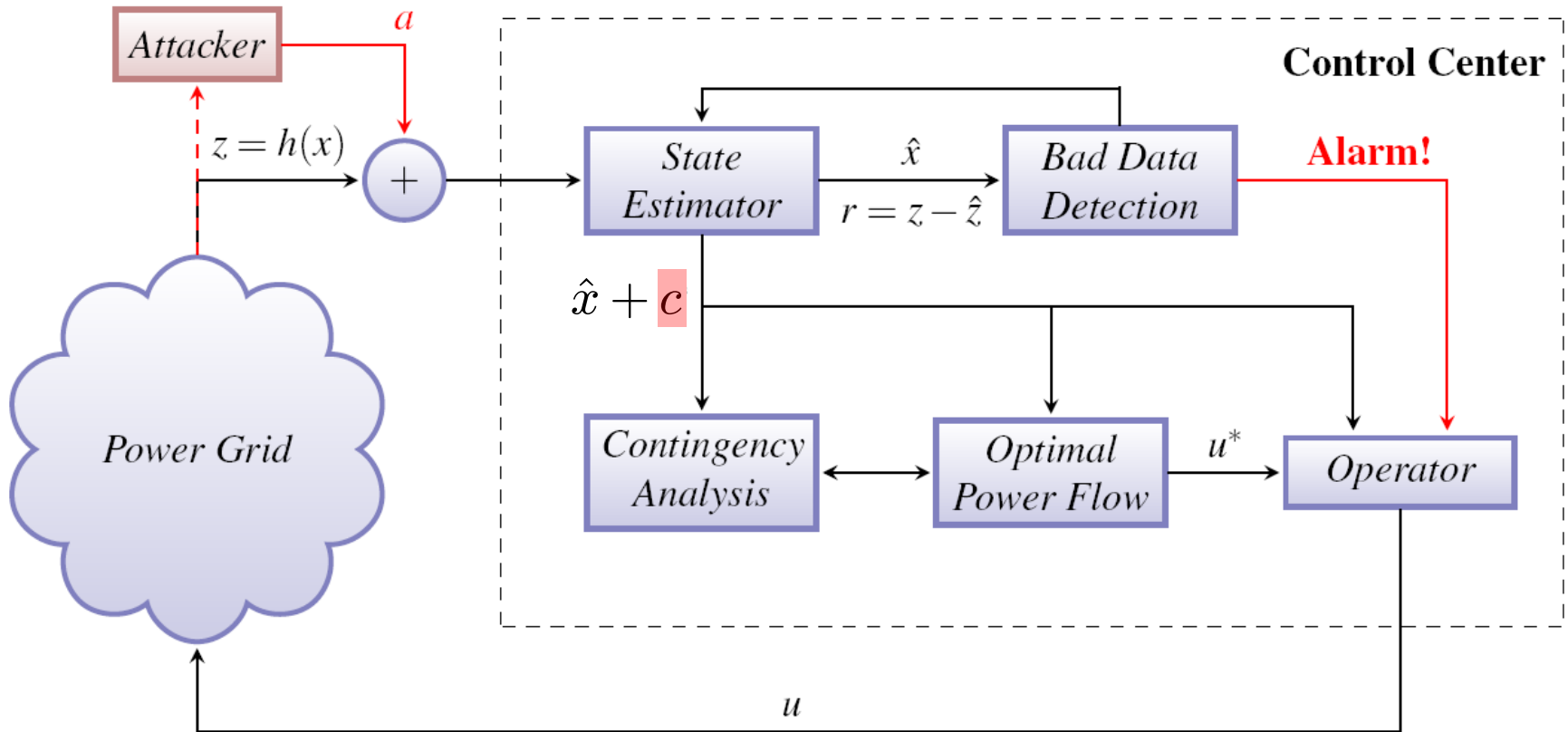
# Issues Considered in the Scenarios

- Model of the power grid
  - What is the relevant time scale?
  - What is the relevant spatial resolution?
- Attacker's capabilities
  - What model does the attacker have access to?
  - What signals do the attacker have access to?
  - **Often easier to assume powerful attacker  $\Rightarrow$  bounds on attack impact**
- Possible physical impact
  - Simulation and experiments in test beds
- Possible defense and mitigation strategies
  - Model-based residual generators
  - Protect the most critical signals

# Scenario 1

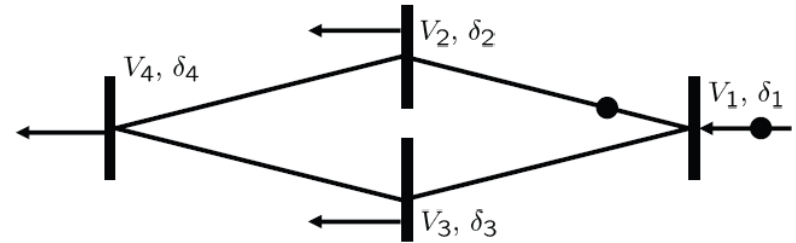


# Attacker Model and Bad Data Detection in Control Center



- Scenario: Attacker injects **malicious data**  $a$  to induce bias  $c$  in state estimate
- First characterize the set of **undetectable** malicious data  $a$

# Power Network and Estimator Models



- Steady-state models:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{V_1 V_2}{X_{12}} \sin(\delta_1 - \delta_2) + \frac{V_1 V_3}{X_{13}} \sin(\delta_1 - \delta_3) \\ \frac{V_1 V_2}{X_{12}} \sin(\delta_1 - \delta_2) \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = h(x) + e \in \mathbb{R}^m$$

- WLS-Estimates of bus phase angles  $\delta_i$  (in vector  $\hat{x}$ ):

$$\hat{x}^{k+1} = \hat{x}^k + (H_k^T R^{-1} H_k)^{-1} H_k^T R^{-1} (z - h(\hat{x}^k))$$

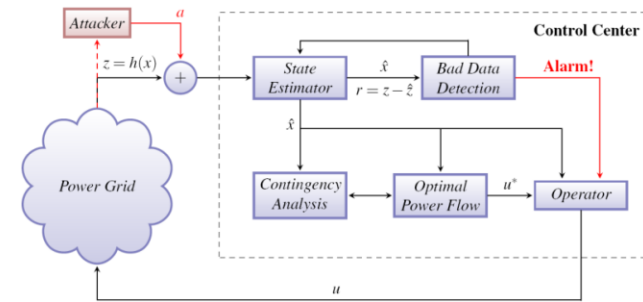
$$H_k := \frac{\partial h}{\partial x}(\hat{x}_k) \quad R := \mathbf{E} e e^T$$

- Linear approximation (DC-power flow model):

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} z$$

$$H := \left. \frac{\partial h(x)}{\partial x} \right|_{x=0}$$

# Bad Data Detection and Undetectable Attacks



- Bad Data Detection triggers when residual  $r$  is large

$$r := z - \hat{z} = z - H\hat{x} = z - H(H^T R^{-1} H)^{-1} H^T R^{-1} z$$

**Theorem:** A malicious data attack  $a$  is undetectable if and only if

$$z_a := z + a$$

$$a = Hc \in \mathcal{R}(H)$$

$$r = z - \hat{z} = z_a - \hat{z}_a$$

[Liu et al., 2009]

- The attacker has a lot of freedom in the choice of  $a$ !
- $a_k \neq 0$  means measurement device  $k$  is corrupted. Attacker likely to seek sparse solutions  $a$ !

# Security Metric $\alpha_k$

- Assume attacker wants to make undetectable attack against measurement channel  $k$

$$\alpha_k := \min_c \|a\|_0 \quad (\text{sparsest possible attack})$$

$$a = Hc \quad (\text{undetectable attack})$$

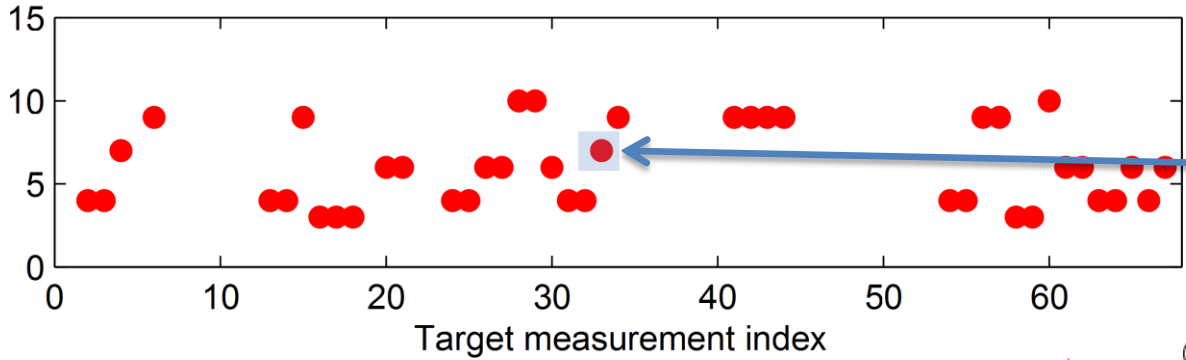
$$a_k = 1 \quad (\text{targets measurement } k)$$

$$(\|a\|_0 := \#\{a_i; a_i \neq 0\}) \quad [\text{Sandberg } et al., 2010]$$

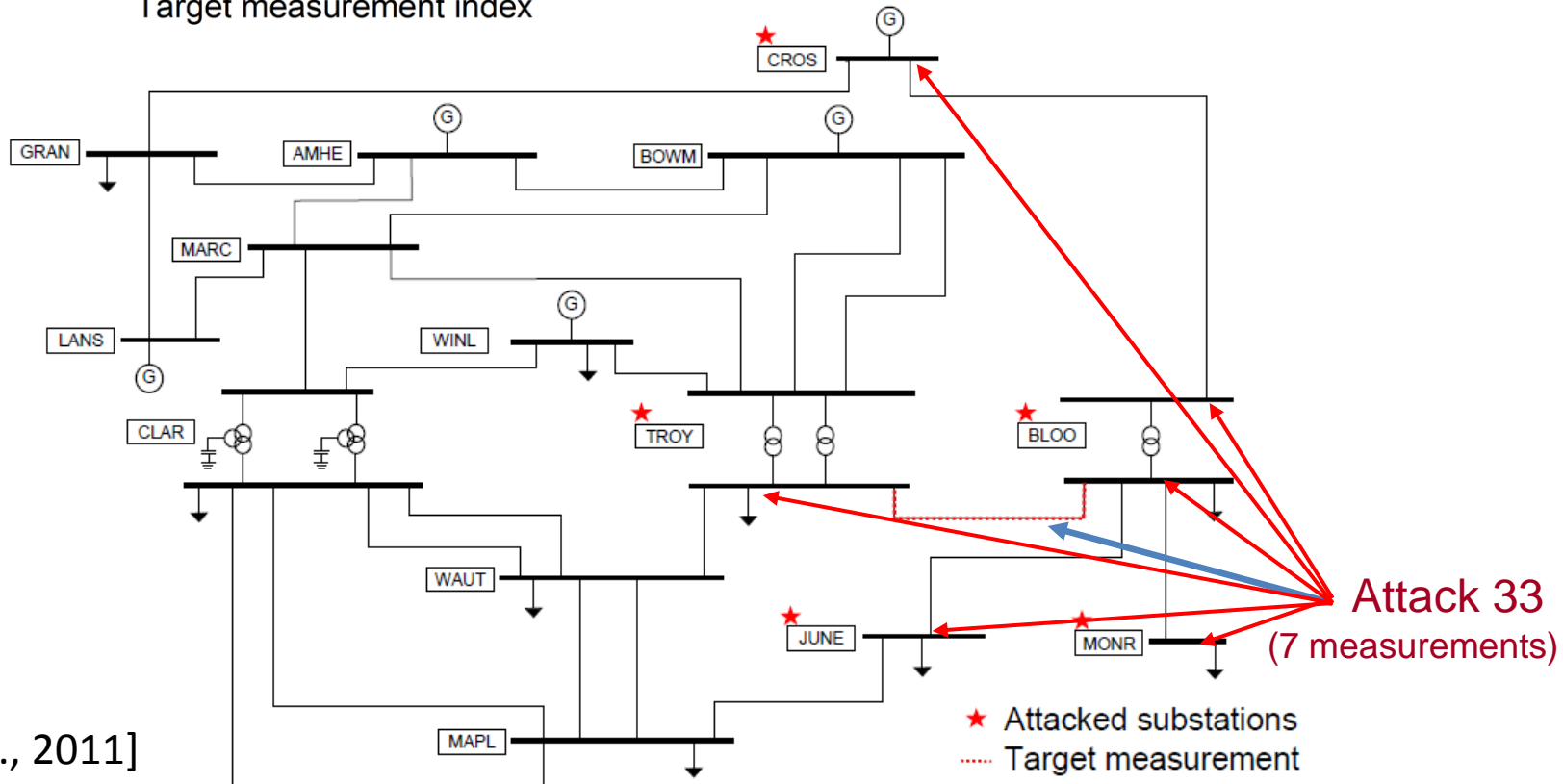
- Quantifies complexity of “least-effort undetectable attack” on measurement  $k$
- Example:**  $\alpha_1=2 \Rightarrow$  undetectable attack against measurement 1 involves *at least two* measurements
- Efficient min-cut/max-flow algorithm for computation exists  
[Hendrickx *et al.*, 2013]

# Security Metric $\alpha_k$ for VIKING 40-bus Network

Security metric



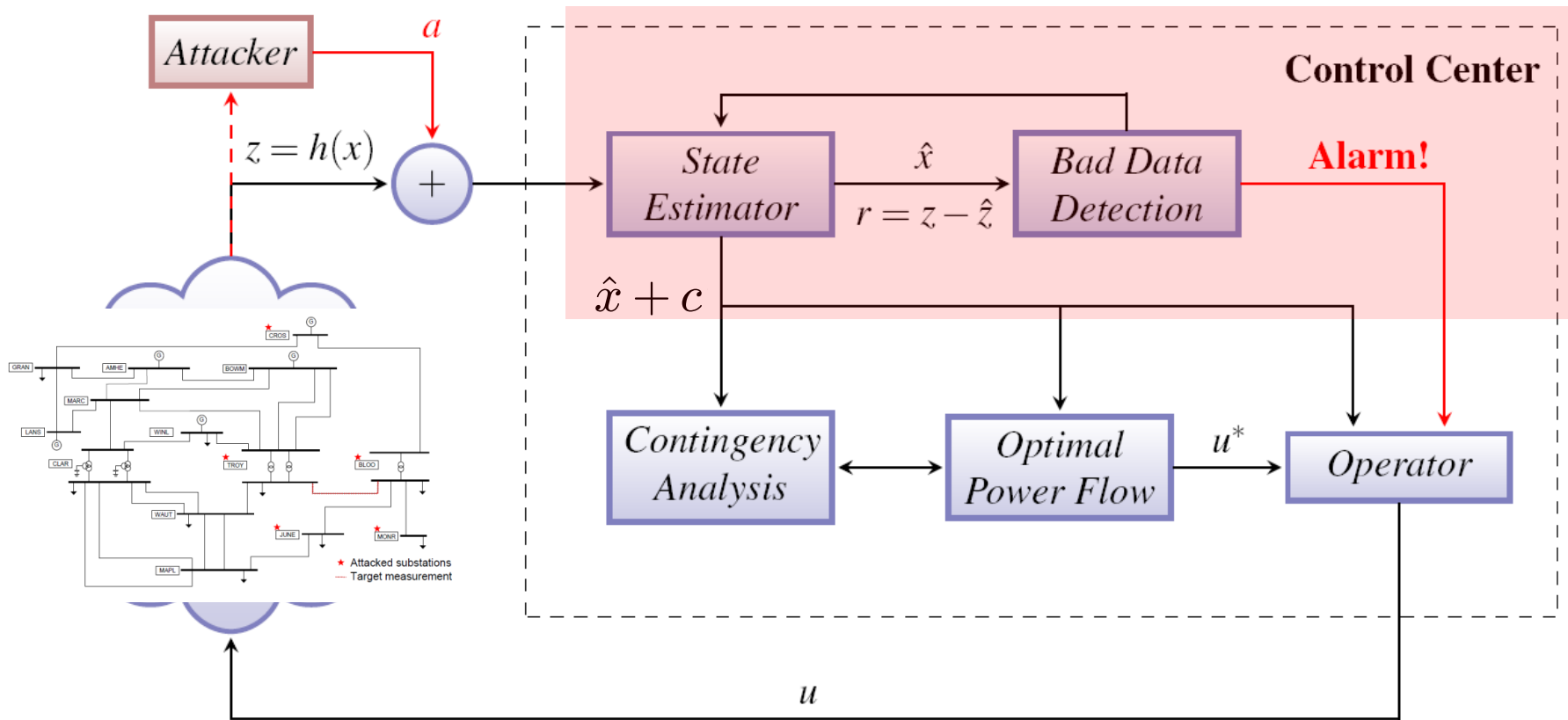
At least 7 measurement readings involved in an *undetectable* attack against measurement 33



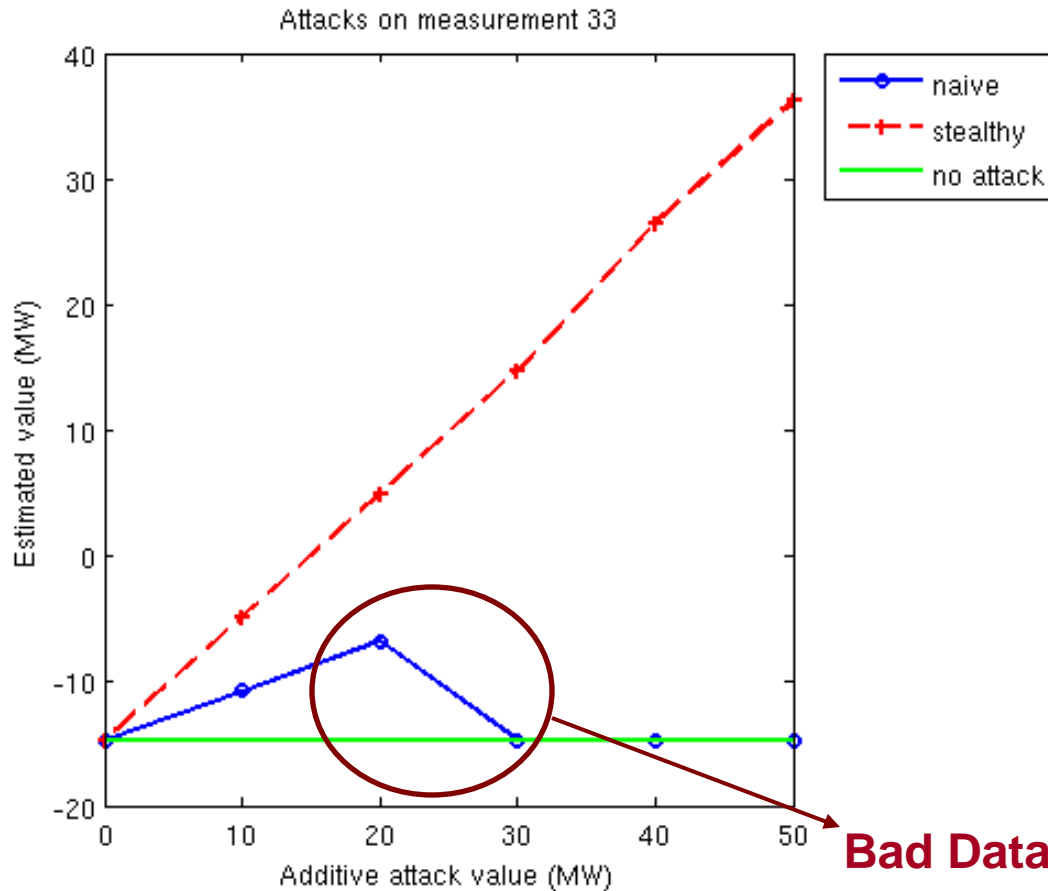
[Teixeira et al., 2011]



# Experiments on SCADA/EMS Testbed



# Experiments on SCADA/EMS Testbed



| False value (MW) | Estimated value (MW) | # BDD Alarms |
|------------------|----------------------|--------------|
| -14.8            | -14.8                | 0            |
| 35.2             | 36.2                 | 0            |
| 85.2             | 86.7                 | 0            |
| 135.2            | 137.5                | 0            |
| 185.2            | Non convergent       | -            |

**Bad Data Detected & Removed**

- Attacks of 150 MW ( $\approx 55\%$  of nominal value) pass undetected in a real system!



# Contingency Alarms

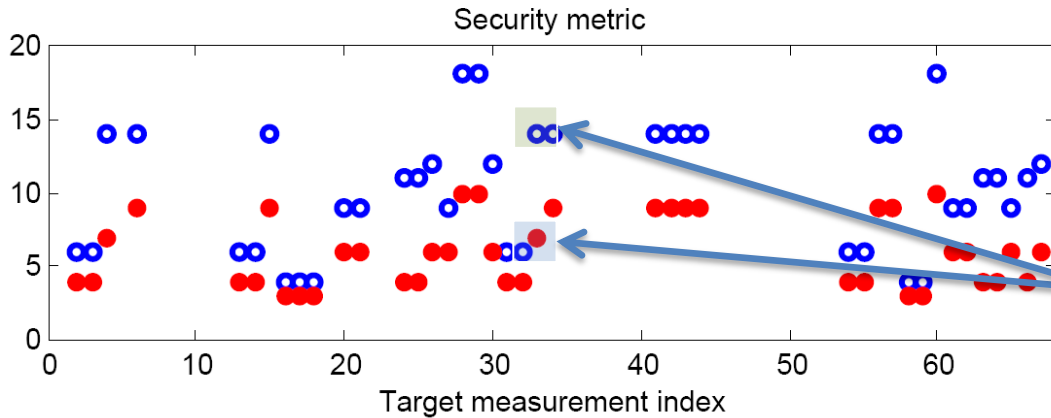
| Target bias,<br>$a_{33}$ | False value<br>(MW), $z_{33}^a$ | Estimate<br>(MW), $\hat{z}_{33}^a$ | #BDD<br>Alarms | #CA<br>Alarms |
|--------------------------|---------------------------------|------------------------------------|----------------|---------------|
| 0                        | -14.8                           | -14.8                              | 0              | 2             |
| 50                       | 35.2                            | 36.2                               | 0              | 2             |
| 100                      | 85.2                            | 86.7                               | 0              | 10            |
| 150                      | 135.2                           | 137.5                              | 0              | 27            |
| 200                      | 185.2                           | -                                  | -              | -             |

- 25 new CA alarms and no BDD alarms!
- What is the reaction in the control room?
  - The human operator may think the system is in a seriously bad state, but in reality the system is in the same state as before the attack

# Protection Strategies

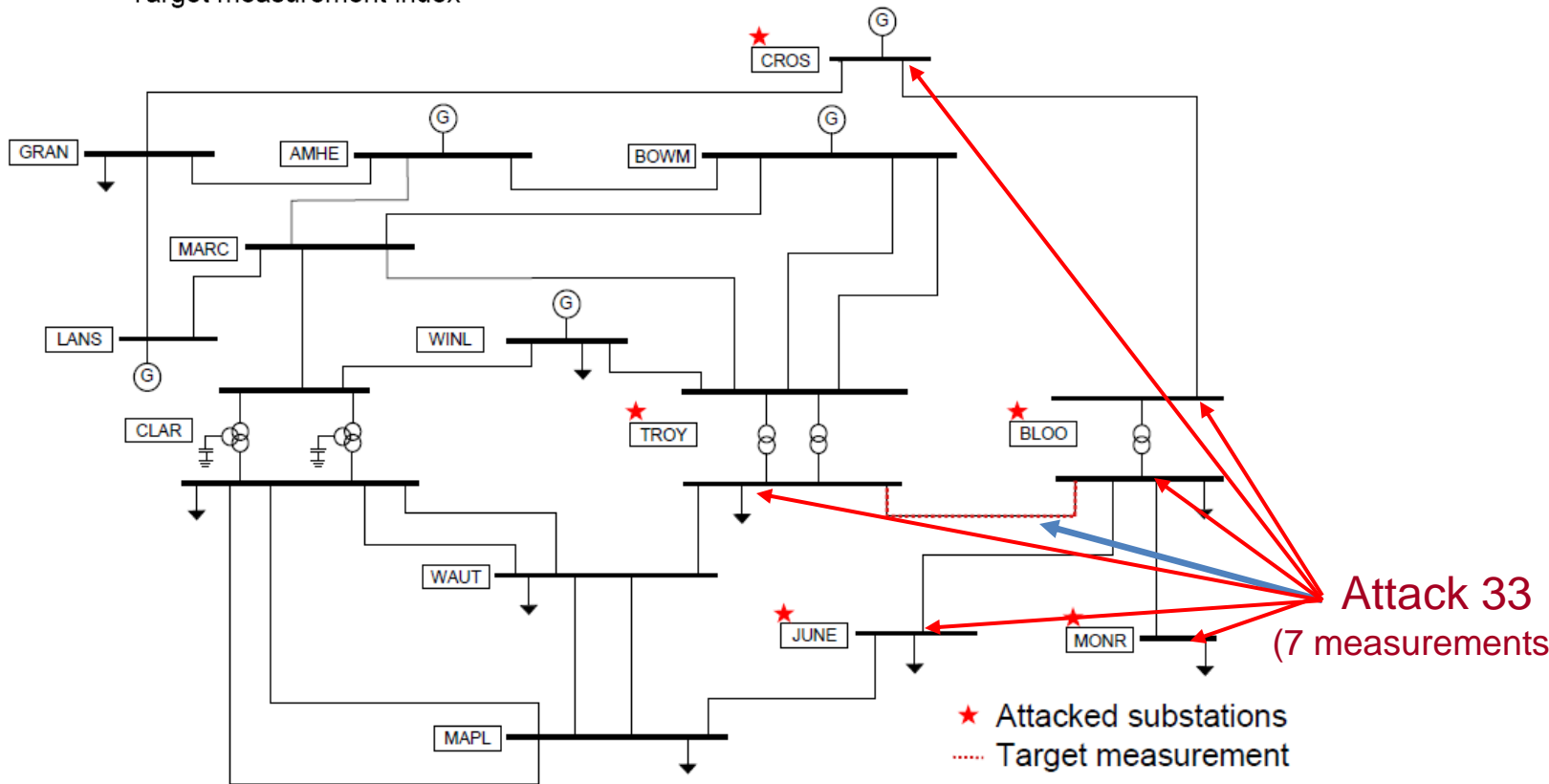
1. Introduce encryption and re-route communication
  - Use security metric to identify critical measurements
  - Possibly weight metric based on expected impact, using optimal power flow [Vukovic *et al.*, 2012], [Teixeira *et al.*, 2012]
2. Augment bad-data detection with cross-checking with historical data of operation [Kosut *et al.*, 2010]
3. Improve available bad-data detection algorithms
  - Discover model mismatch between active/reactive power flow measurements [Sou *et al.*, 2012]

# 1. Use Metric to Assign Protection



- = Current measurement config.
- = Upgraded measurement config.

With a few measurements protected, 14 instead of 7 measurements has to be involved in an undetectable attack!



# Summary Scenario 1

- Multiple interacting attacked measurements may be undetectable
- Security metric  $\alpha_k$  identifies measurements that are relatively “easy” to attack (locates weak spots)
- Experimental validation shows significant possible impact in control center (CA alarms)
- Protection strategies include encryption and re-routing of critical measurements

# References Scenario 1

## Undetectable attacks and security index

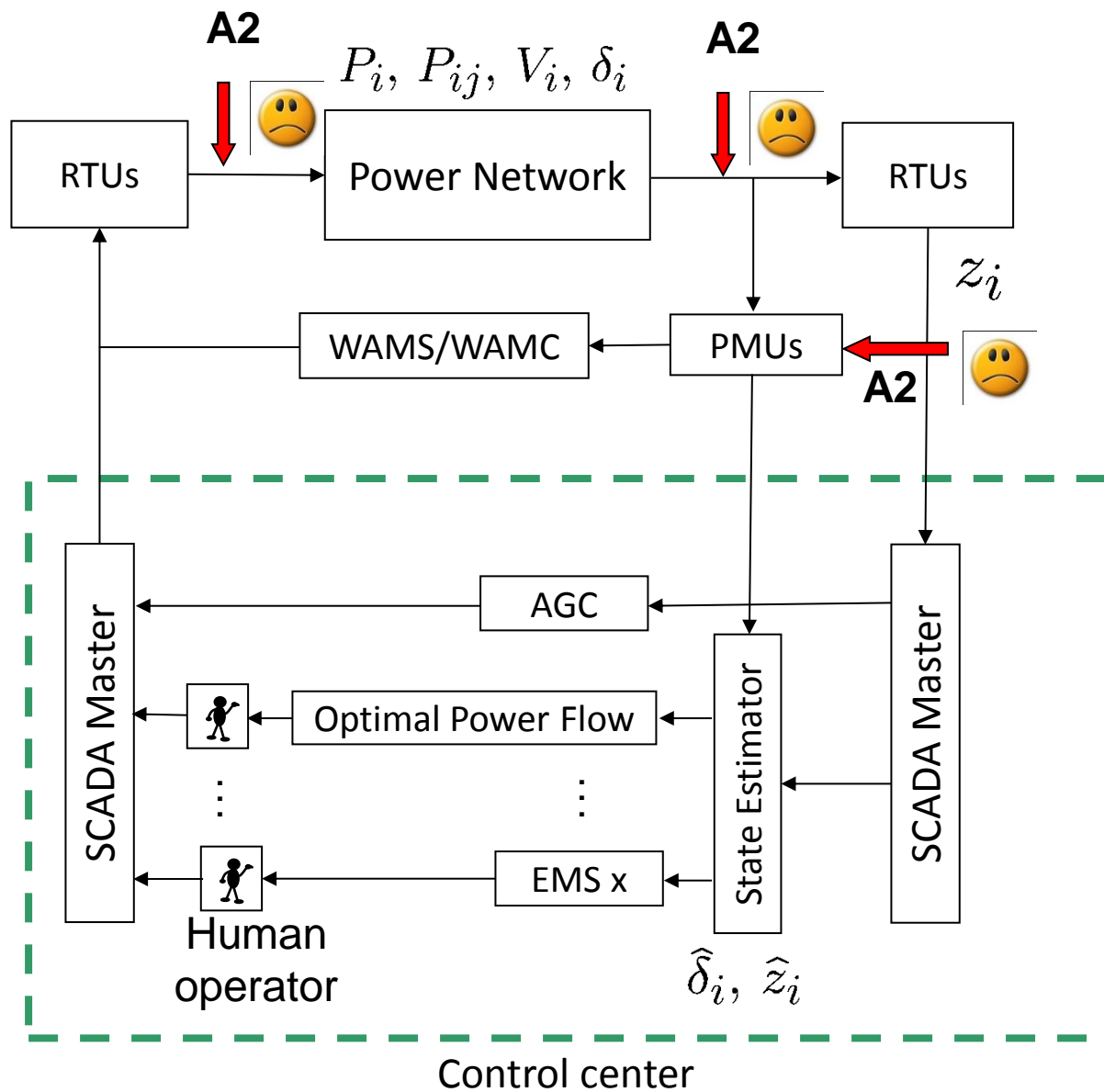
- Y. Liu, P. Ning, M. Reiter, “False data injection attacks against state estimation in electric power grids,” Proceedings of the 16<sup>th</sup> ACM conference on Computer and communications security, 2009
- A. Teixeira, G. Dán, H. Sandberg, K. H. Johansson, "A Cyber Security Study of a SCADA Energy Management System: Stealthy Deception Attacks on the State Estimator," Preprints of the 18th IFAC World Congress, 2011
- J. M. Hendrickx, K. H. Johansson, R. M. Jungers, H. Sandberg, K. C. Sou, “Efficient Computations of a Security Index for False Data Attacks in Power Networks,” 2013. Preprint arXiv:1204.6174

## Protection strategies

- O. Vukovic, K. C. Sou, G. Dán, H. Sandberg: "Network-aware Mitigation of Data Integrity Attacks on Power System State Estimation". IEEE Journal on Selected Areas in Communications (JSAC), 2012
- K. C. Sou, H. Sandberg, and K. H. Johansson, “Detection and identification of data attacks in power system,” American Control Conference (ACC), 2012



# Scenario 2



# Difference to Scenario 1

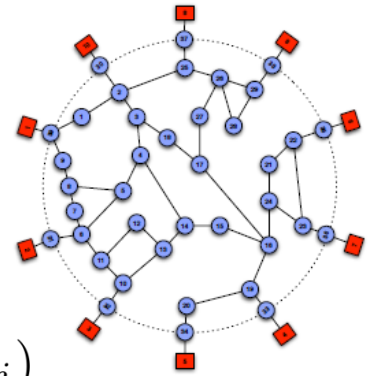
- SCADA/EMS system now has access to high sampling rate Phasor Measurement Units (PMUs, ~50 Hz sampling)
- Power system transients can be observed
  - ⇒ SE/BDD can use dynamical model

- To remain undetected an attacker needs to take dynamic model constraints into account

- Centralized monitoring solution puts severe constraints on IT infrastructure
- Distributed solutions have been considered

[Pasqualetti *et al.*, 2011], [Shames *et al.*, 2011]

# Active Power Flow Model



- Generators (■) 
$$M_i \ddot{\delta}_i = P_{\text{mech.in},i} - \sum_j \overbrace{\frac{V_i V_j}{X_{ij}}}^{Y_{ij}} \sin(\delta_i - \delta_j)$$
$$\approx P_{\text{mech.in},i} - \sum_j Y_{ij} \cdot (\delta_i - \delta_j)$$

inertia  $M_i$ , phase  $\delta_i$ , voltage  $V_i$ , reactance  $X_{ij}$

- Loads (●) 
$$0 = P_{\text{load},i} - \sum_j \frac{V_i V_j}{X_{ij}} \sin(\delta_i - \delta_j) \approx P_{\text{load},i} - \sum_j Y_{ij} \cdot (\delta_i - \delta_j)$$
- Time scale is sub-second. Resistive losses neglected
- Linear differential-algebraic equation:  $E\dot{x} = Ax$

# Cyber-Physical Attack

$$E\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

- $y(t)$  models continuous-time measurement (PMUs)
- $u(t)$  models continuous-time perturbation
  - Regular faults
  - Physical attack
  - Cyber attack
- Is it possible to detect these attack/faults through measurements  $y(t)$ ?

# Undetectable Deterministic Faults and Attacks

## Assumptions:

- $(E, A)$  is regular, i.e.,  $\det(Es - A) \neq 0$  for some  $s$
- Initial state  $x(0)$  is consistent
- Unknown perturbation  $u(t)$  is smooth

**Definition:** A perturbation  $u(t)$  is *undetectable* if there exists two consistent initial states such that for all times

$$y(x_1(0), u(t), t) = y(x_2(0), 0, t)$$

# Undetectable Deterministic Faults and Attacks

- Linearity implies that  $y(x_1(0) - x_2(0), u(t), t) = 0$   
 $\Rightarrow$  zero-dynamics

**Theorem:** There exists an undetectable perturbation if and only if

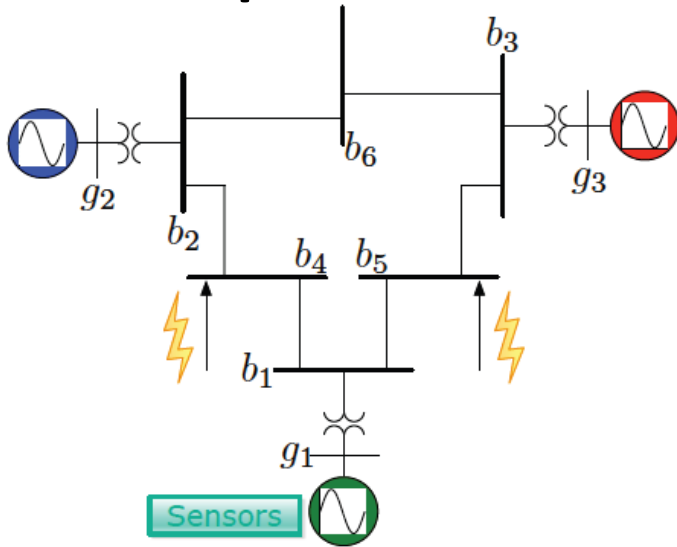
$$\begin{bmatrix} sE - A & -B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ g \end{bmatrix} = 0$$

for some  $s$ ,  $x \neq 0$ , and  $g$

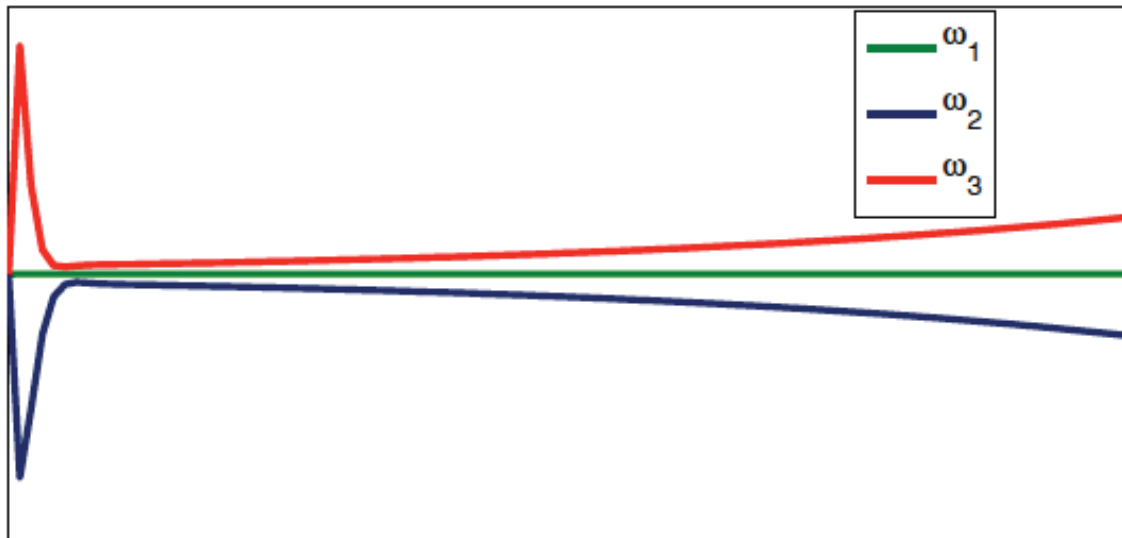
[Pasqualetti *et al.*, 2011]

- **Generalization of steady-state condition in Scenario 1**

# Example: WECC 3-Machine 6-Bus System



- **Measurement**  $y(t)$  :  
angle and frequency of gen.  $g_1$
- **Attack**  $u(t)$  : modify real power injections at buses  $b_4$  and  $b_5$



← Undetectable attack!

[Pasqualetti *et al.*, 2011]

# Possible Mitigation

- Similar to Scenario 1:
  - Protect selected communication channels, etc.
- *Detectable* perturbations can be detected and sometimes identified:

System model

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

Dynamic Bad Data Detector

$$\begin{aligned} E\dot{w}(t) &= (A + GC)w(t) - Gy(t) \\ r(t) &= Cw(t) - y(t) \end{aligned}$$

**Theorem:** Assume  $x(0) = w(0)$ ,  $(E, A + GC)$  is Hurwitz, and perturbation detectable. Then  $r(t) = 0$  if and only if  $u(t) = 0$

[Pasqualetti *et al.*, 2011]



# Summary Scenario 2

- High sampling rate measurements leads to dynamic models in monitoring system
- Harder for attacker to remain undetectable: More constraints to satisfy
- Perturbation is **undetectable** iff measurements can be explained with **no perturbation** and a **proper initial state**
- Similar protection strategies as in Scenario 1 possible

# References Scenario 2

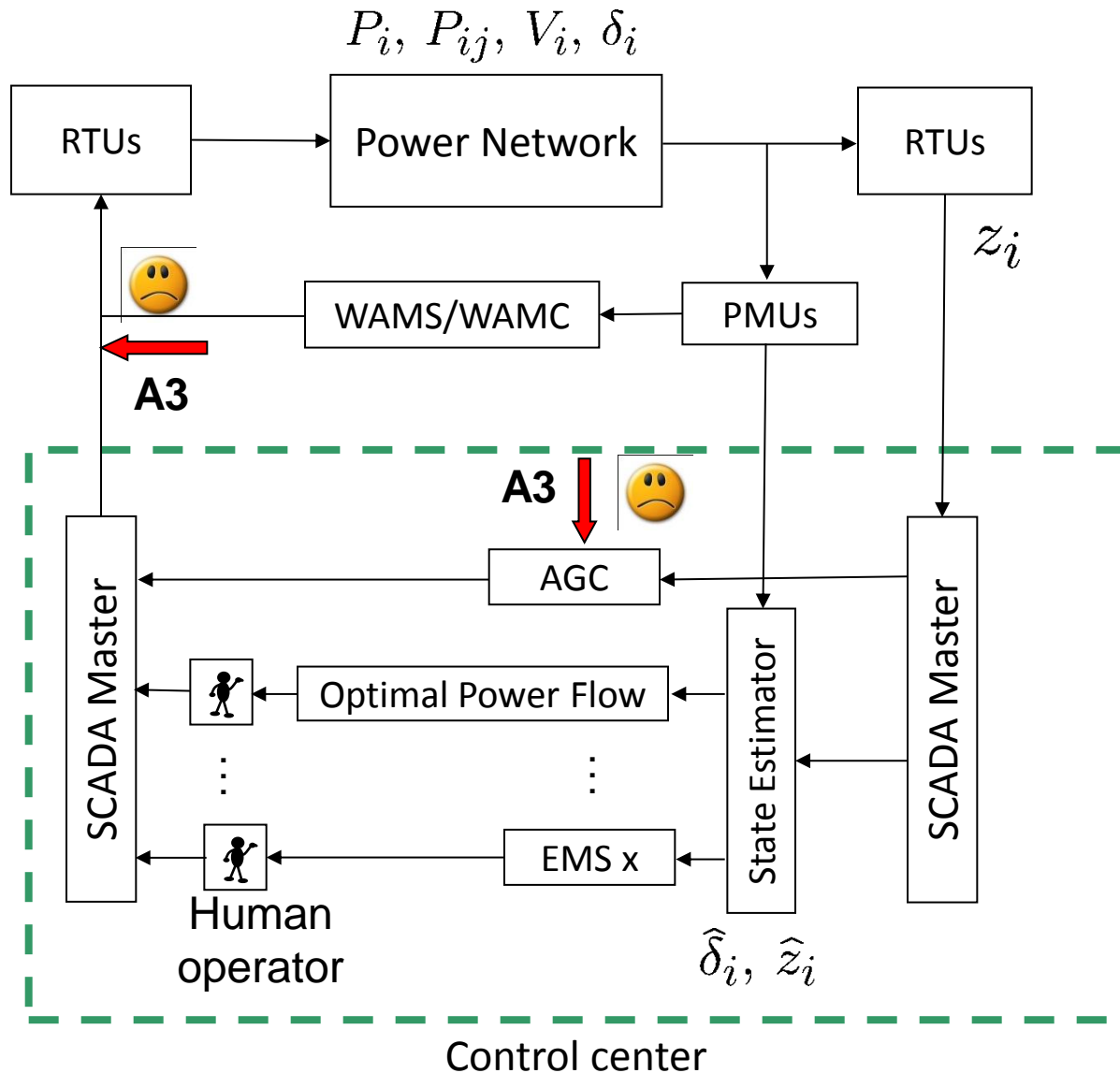
## **Cyber-physical attacks**

- F. Pasqualetti, F. Dörfler, and F. Bullo. “Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design.” In IEEE Conf. on Decision and Control, 2011
- F. Pasqualetti, F. Dörfler and F. Bullo. “Attack Detection and Identification in Cyber-Physical Systems,” IEEE Transactions on Automatic Control, 2012. To appear

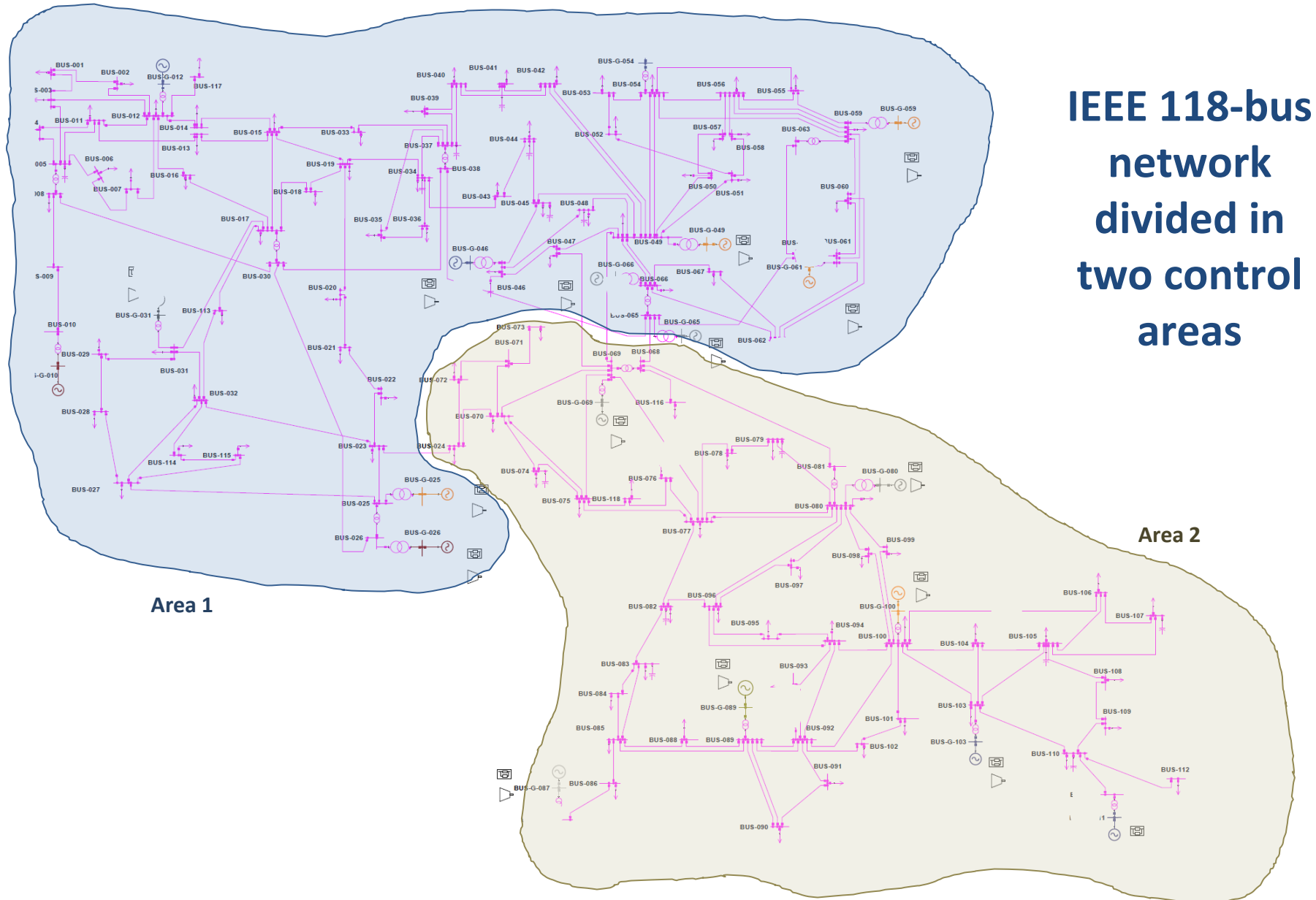
## **Distributed monitoring**

- F. Pasqualetti, R. Carli, and F. Bullo. “Distributed estimation and false data detection with application to power networks.” Automatica, 2011
- I. Shames, A. Teixeira, H. Sandberg, and K. H. Johansson: "Distributed Fault Detection for Interconnected Systems". Automatica, 2011

# Scenario 3



# Test system: Two-Area Power Network



**IEEE 118-bus  
network  
divided in  
two control  
areas**

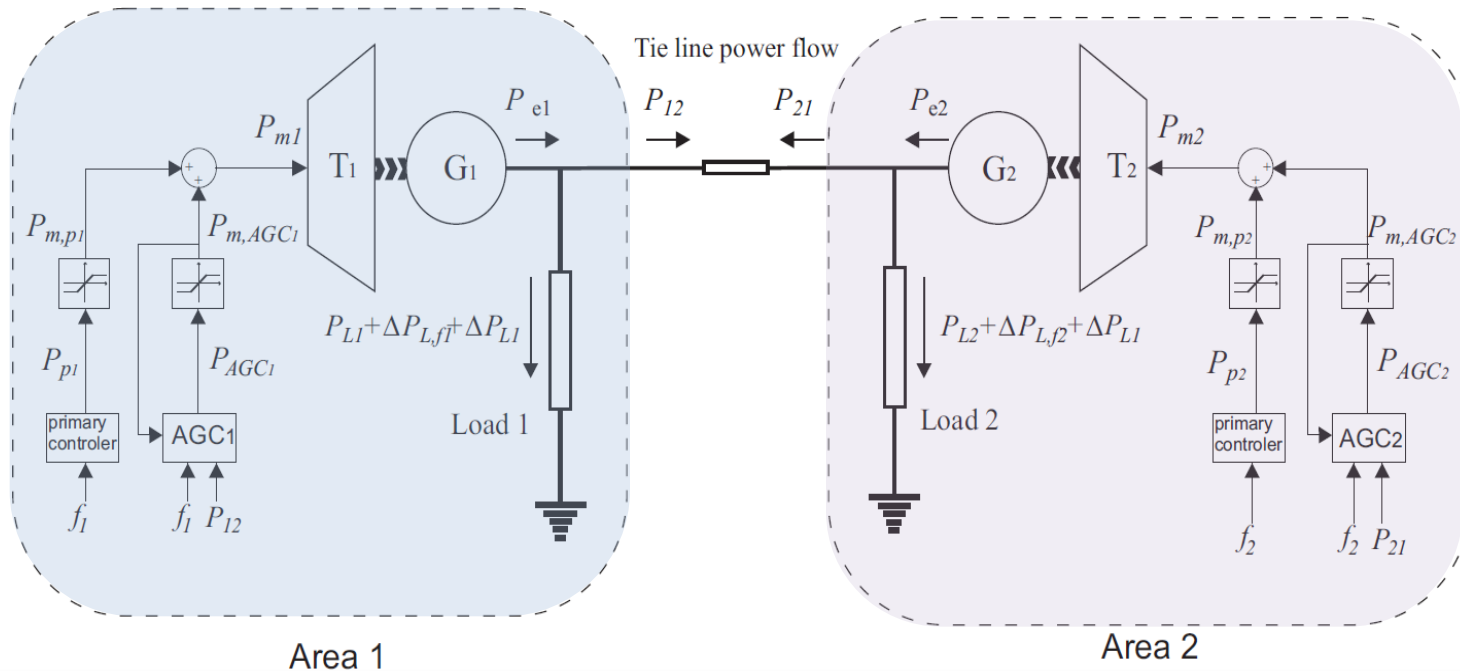
**Area 1**

**Area 2**

# Two-Area Power Network

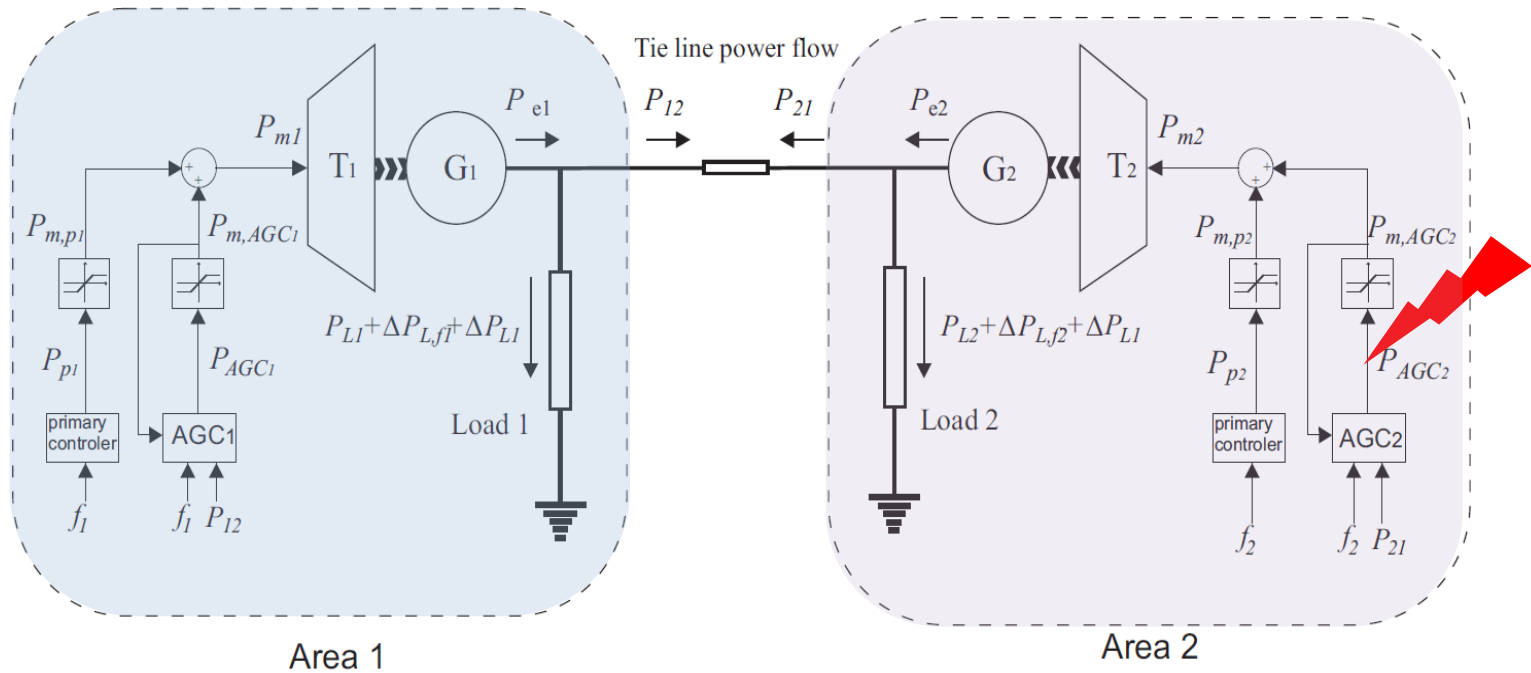
- Areas represent countries or regions
- Intra-area
  - Generation of power
  - Consumption of power
  - Primary control for maintaining frequency
- Inter-area: Automatic Generation Control (AGC)
  - Secondary control to regulate power exchange
  - Maintain frequency at desired value
  - Maintain power flow on tie lines to agreed value
- AGC loop closed automatically over SCADA system

# Automatic Generation Control (AGC)



- Only automatic control loop commonly closed over the SCADA system
- Used to control:
  - Frequency in areas ( $f_1, f_2$ )
  - Tie line power flow ( $P_{12}, P_{21}$ )

# Two machine frequency model



- What can attacker do with access to AGC signal in one area?
- Can they cause frequency or power exchange range violations?
  - ➔ Load or generator shedding

# Dynamic and Safety Considerations

Dynamic Model

$$\Delta \dot{f}_1 = \frac{f_0}{2H_1 S_{B_1}} \left( \Delta P_{m,p_1} + \Delta P_{m,AGC_1} - \frac{1}{D_{l_1}} \Delta f_1 - P_T \sin(\Delta\phi + \phi_0) + P_{0_{12}} \right),$$

$$\Delta \dot{f}_2 = \frac{f_0}{2H_2 S_{B_2}} \left( \Delta P_{m,p_2} + \mathbf{u} - \frac{1}{D_{l_2}} \Delta f_2 + P_T \sin(\Delta\phi + \phi_0) - P_{0_{12}} \right),$$

$$\Delta \dot{\phi} = 2\pi(\Delta f_1 - \Delta f_2),$$

$$\Delta \dot{P}_{AGC_1} = \left( \frac{1}{D_{l_1}} \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}} - \frac{1}{S_1} \frac{1}{T_{N_1}} \right) \Delta f_1$$

$$- \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}} \Delta P_{m,p_1} - \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}} \Delta P_{m,AGC_1}$$

$$- \left( \frac{1}{T_{N_1}} - \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}} \right) (P_T \sin(\Delta\phi + \phi_0) - P_{0_{12}})$$

$$- 2\pi C_{p_1} P_T (\Delta f_1 - \Delta f_2) \cos(\Delta\phi + \phi_0) - \frac{K_{a_1}}{T_{N_1}} p_1.$$

Attacker

$$\dot{x} = f(x, w) + g(x, w)u$$

Safety Constraints

$$x \in K \subseteq \mathbb{R}^4$$

$$\Delta f_1 \in [-1.5Hz, 1.5Hz]$$

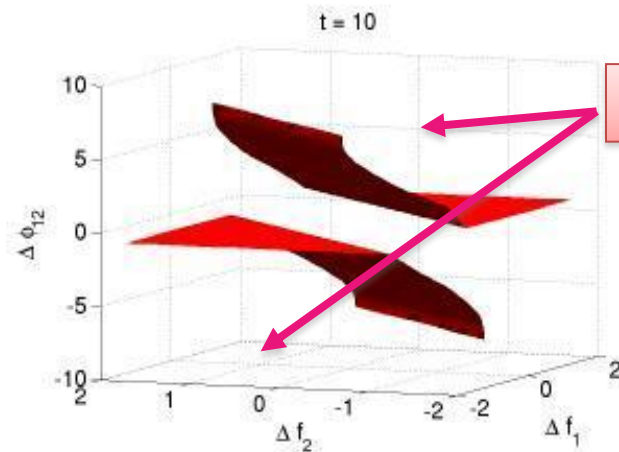
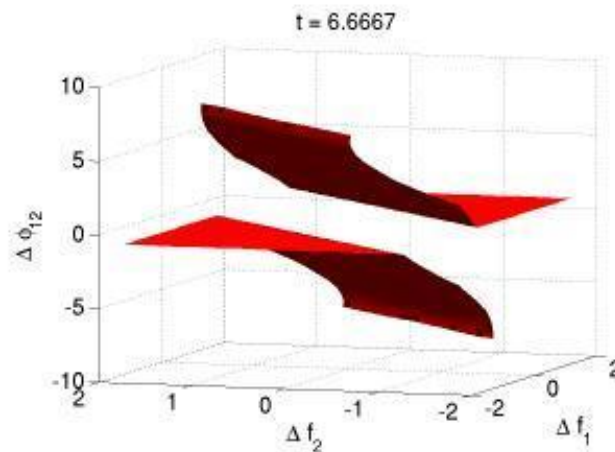
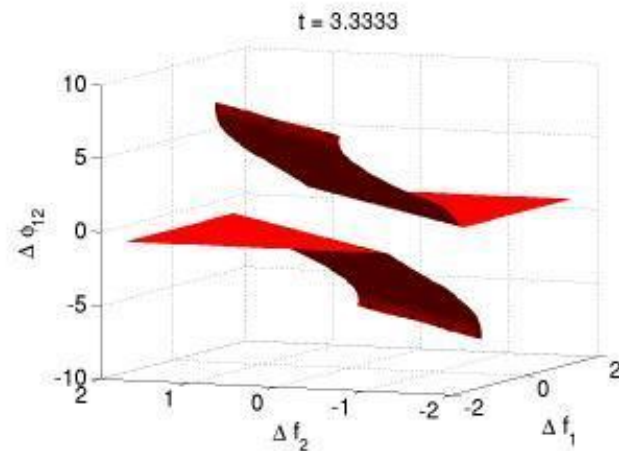
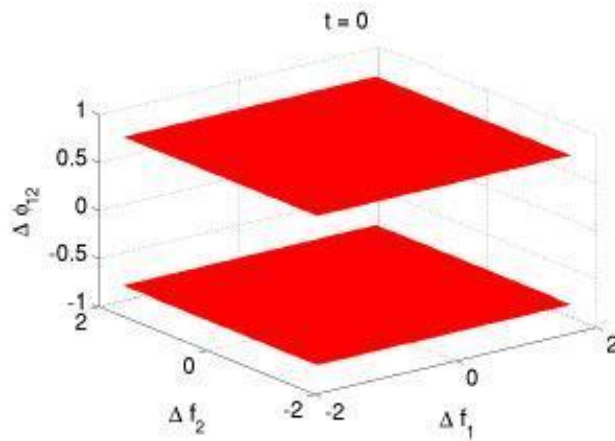
$$\Delta f_2 \in [-1.5Hz, 1.5Hz]$$

$$\Delta \phi \in [-44^\circ, +44^\circ]$$

Reachability analysis:  
Phase constraint can be violated  
and kept in unsafe region forever



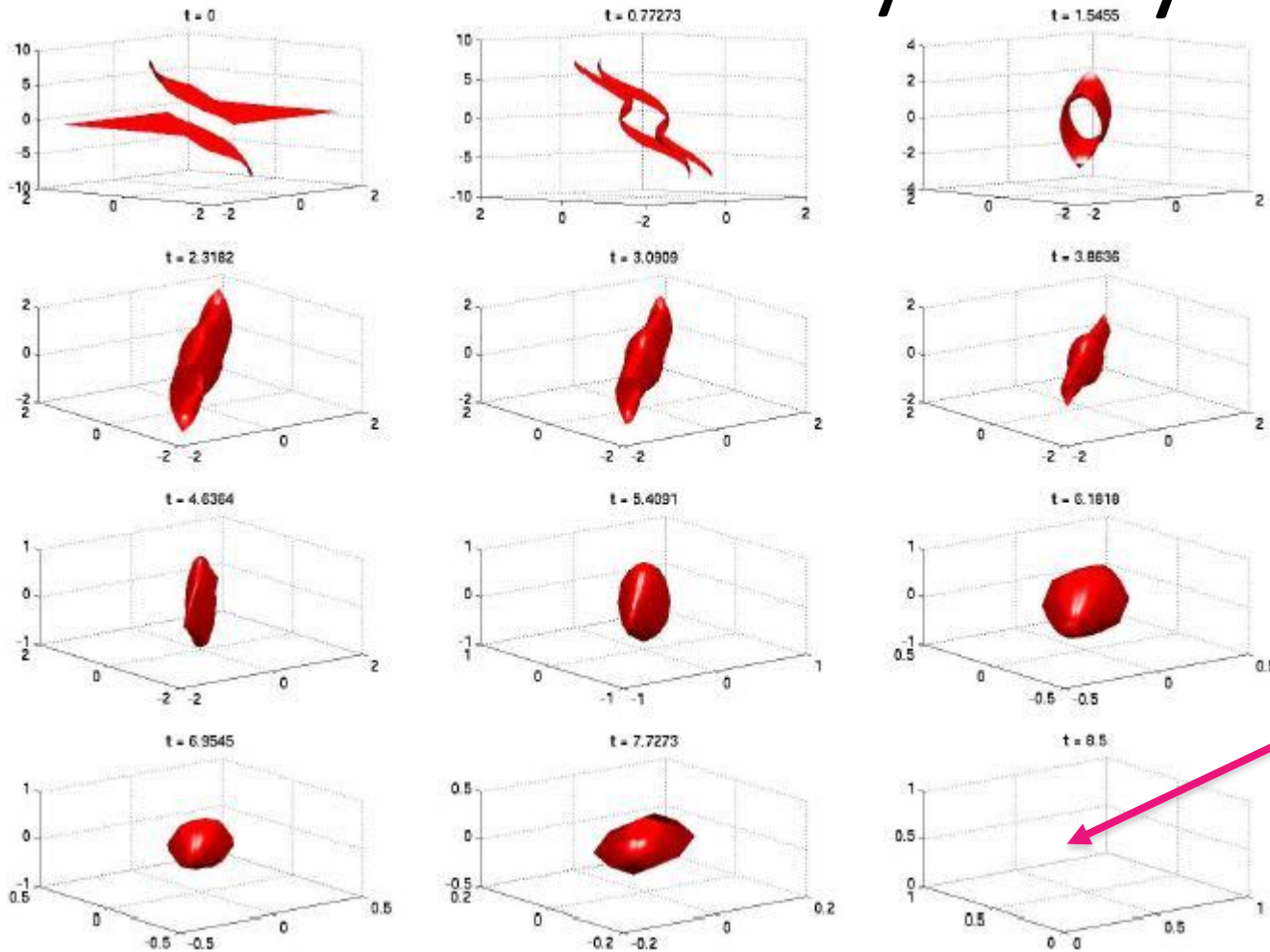
# Reachability Analysis



Unsafe initial sets

The set of initial conditions from which the attacker can keep the phase constraints violated forever

# Reachability Analysis



Can reach to unsafe region from everywhere

Reachability analysis of abstract model suggests system vulnerable

# Outline

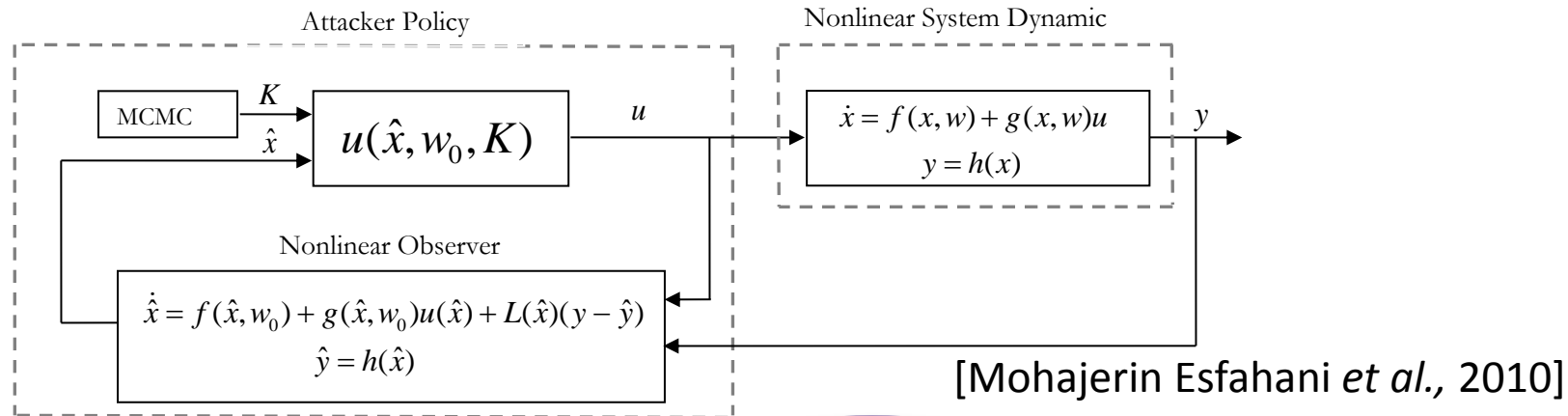
- Can attacker cause problems by manipulating AGC?

Yes he can!

- How?

# Synthesizing an Attack Signal

## Feedback policy (Feedback Linearization + MCMC)



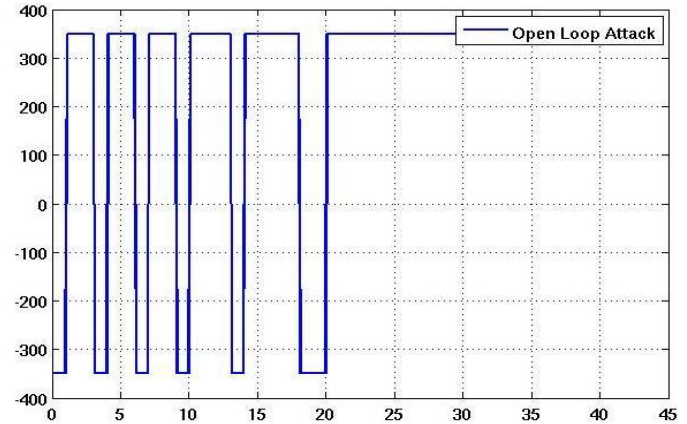
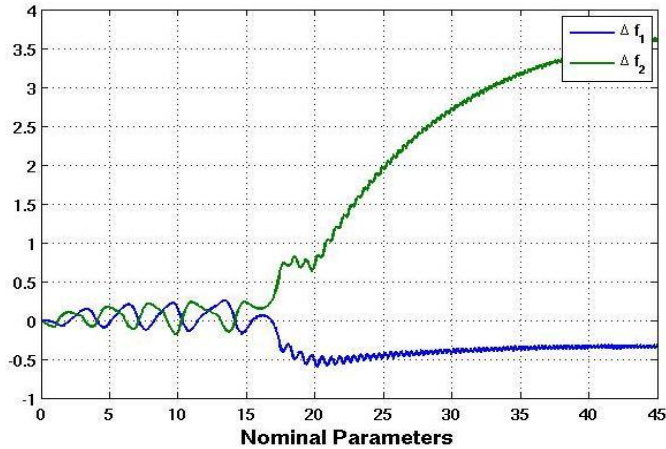
Measurement of two frequencies enough to build observer

- What information is needed to construct this policy?
- Does the feedback policy self-correct the parameter uncertainties?

# Synthesizing an Attack Signal

## Feedback policy (Feedback Linearization + MCMC)

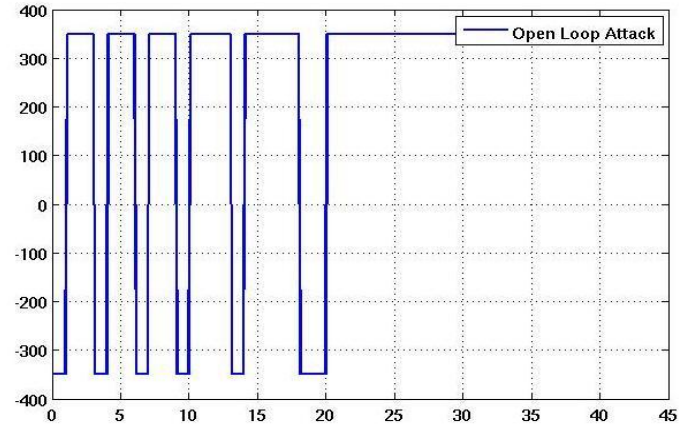
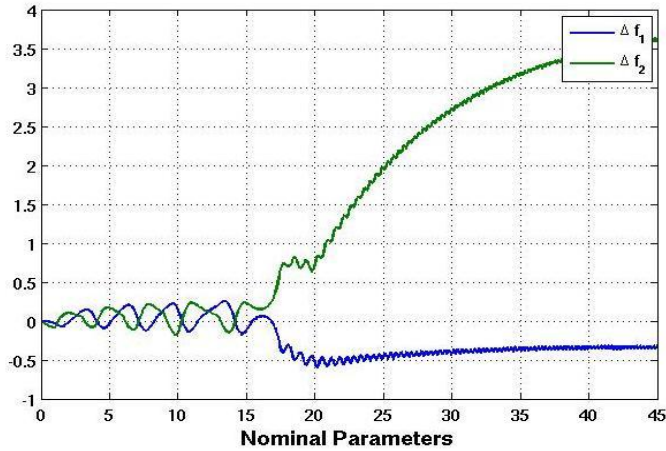
Perfect Model



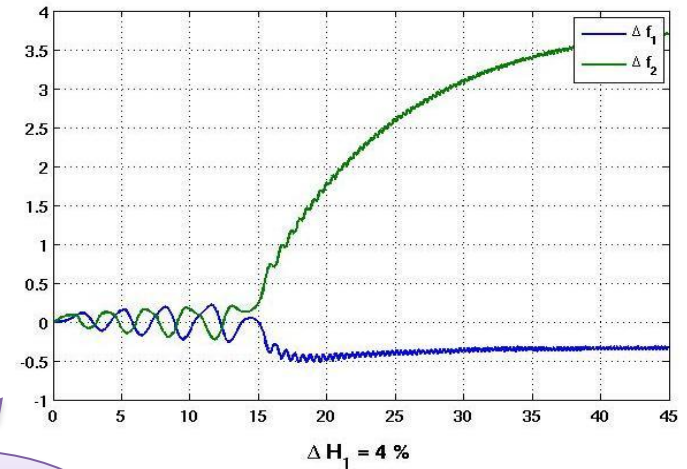
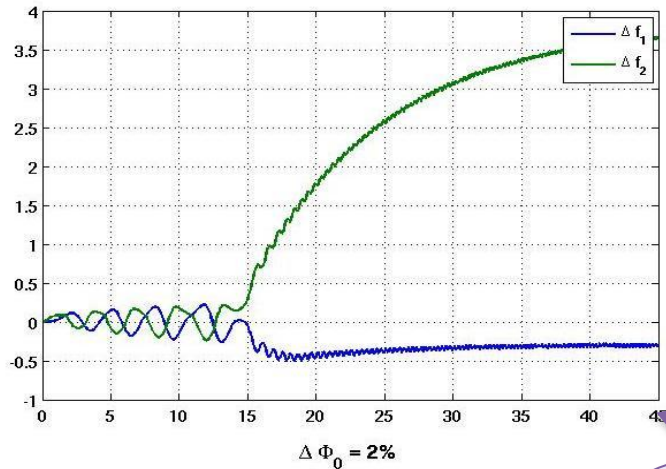
# Synthesizing an Attack Signal

## Feedback policy (Feedback Linearization + MCMC)

Perfect Model



Almost the same performance with imperfect information



Imperfect Model

# Outline

- Can attacker cause problems by manipulating AGC?

Yes he can!

- How?

With a fairly sophisticated feedback controller

- What can we do about it?

# Mitigation and Protection Ideas

## Main objective:

Develop mitigation strategy against AGC intrusion

Pseudo-mitigation ideas based on previous analysis. It helps to:

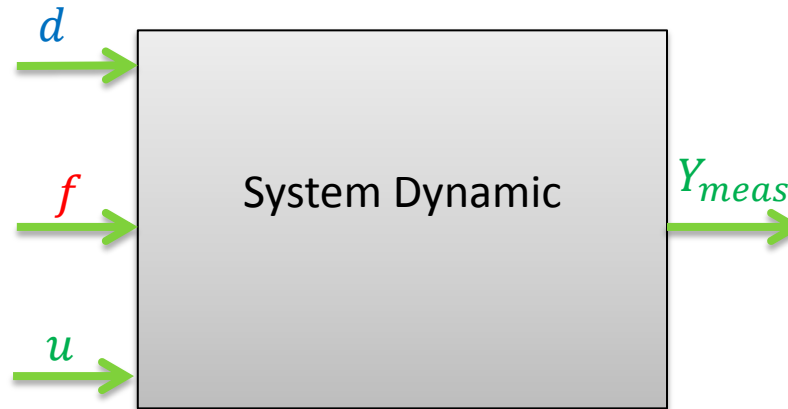
- Increase damping ratio of network areas
- Lower scheduled active power exchange between the areas
- In case of attack, disconnect AGC

## Question:

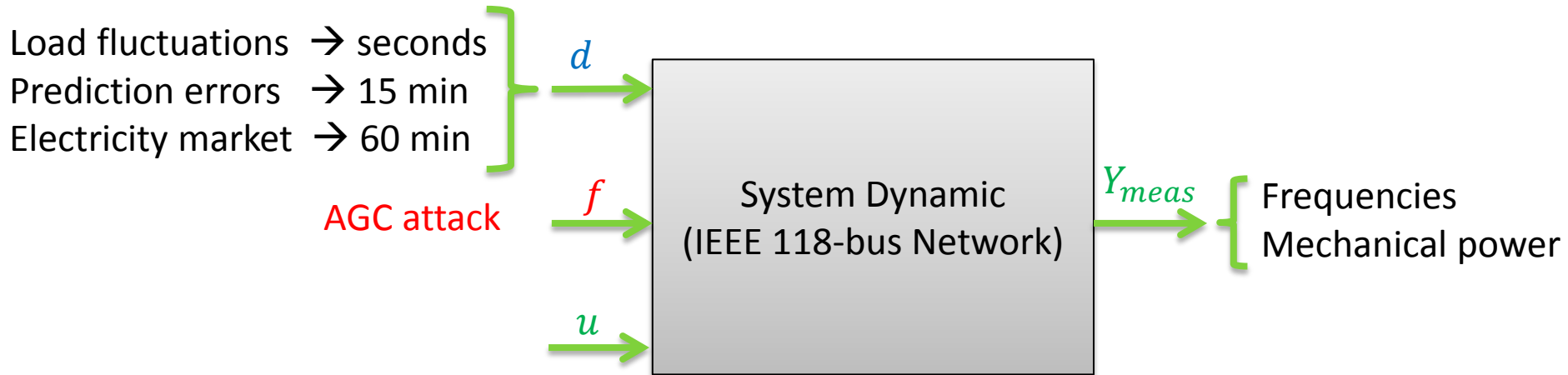
Would it be possible to diagnose AGC intrusion sufficiently fast?



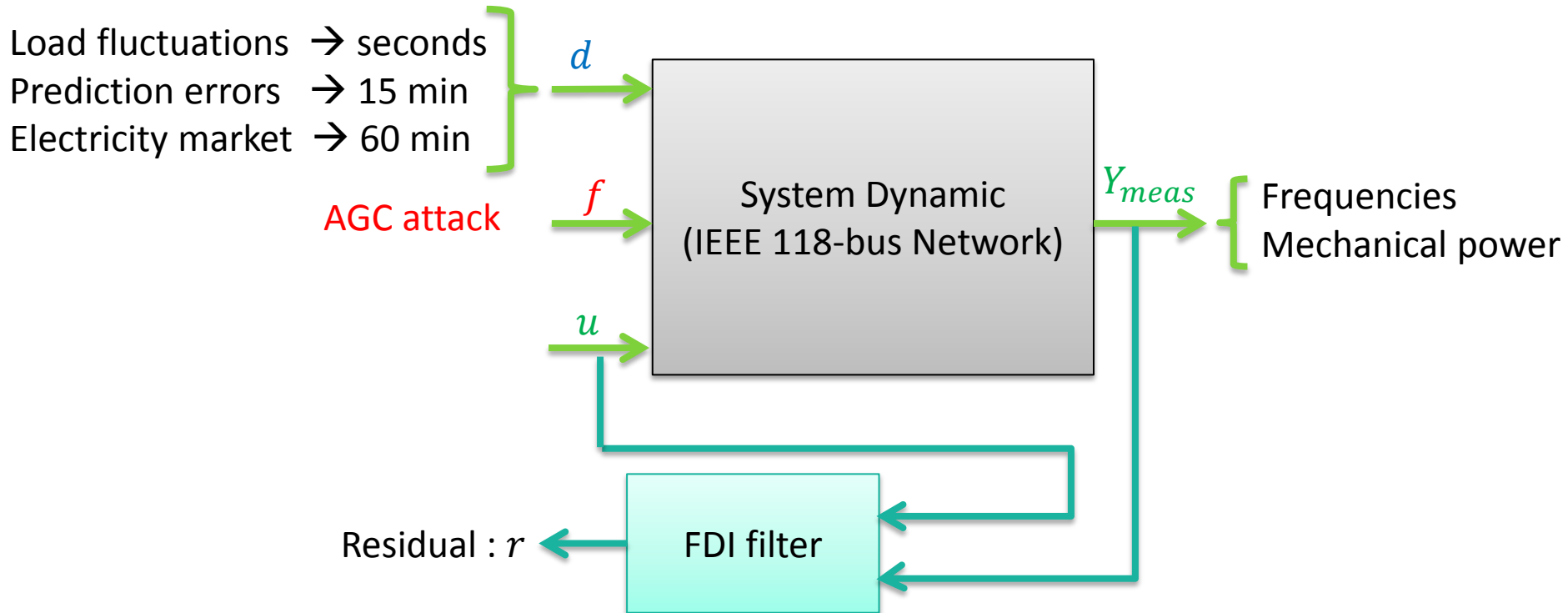
# Fault Detection and Isolation (FDI) Problem



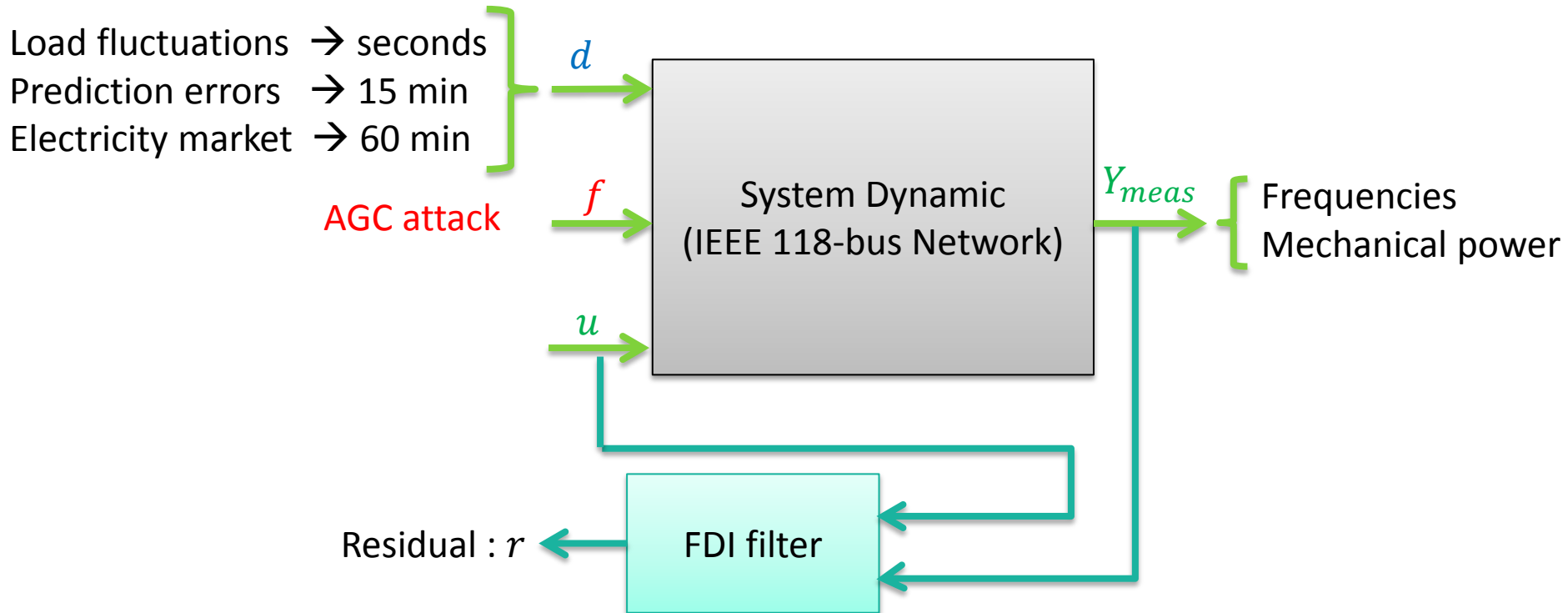
# Fault Detection and Isolation (FDI) Problem



# Fault Detection and Isolation (FDI) Problem

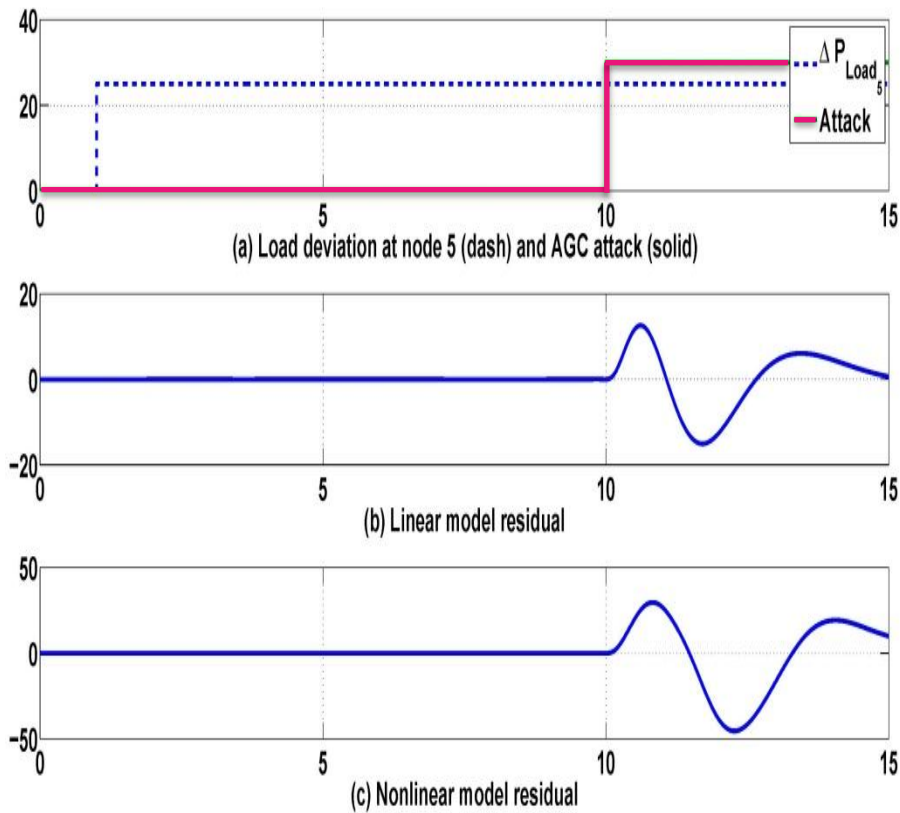


# Fault Detection and Isolation (FDI) Problem



FDI Problem:  $\begin{cases} d \mapsto r \equiv 0 & \text{Disturbance rejection} \\ f \mapsto r \neq 0 & \text{Fault sensitivity} \end{cases}$

# Simulation Results Based on Multi-Machine Frequency Model (19 generators, 59 states)



## Trained nonlinear filter

- (II-a) Disturbance and intrusion signal
- (II-b) Filter performance in linear setup
- (II-c) Filter performance in nonlinear setup

[Mohajerin Esfahani *et al.*, 2013]

# Summary Scenario 3

- Impact of cyber-attack on two-area power system AGC
  - Reachability  $\rightarrow$  system is vulnerable to the attack
  - Feedback attack policy under partial state information
  
- Mitigation ideas for AGC security
  - Robust FDI filter for high dimensional nonlinear systems
  - Designed for the multi-machine frequency model (59 states)
  - Tested on full model (567 states + 236 algebraic equations)

# References Scenario 3

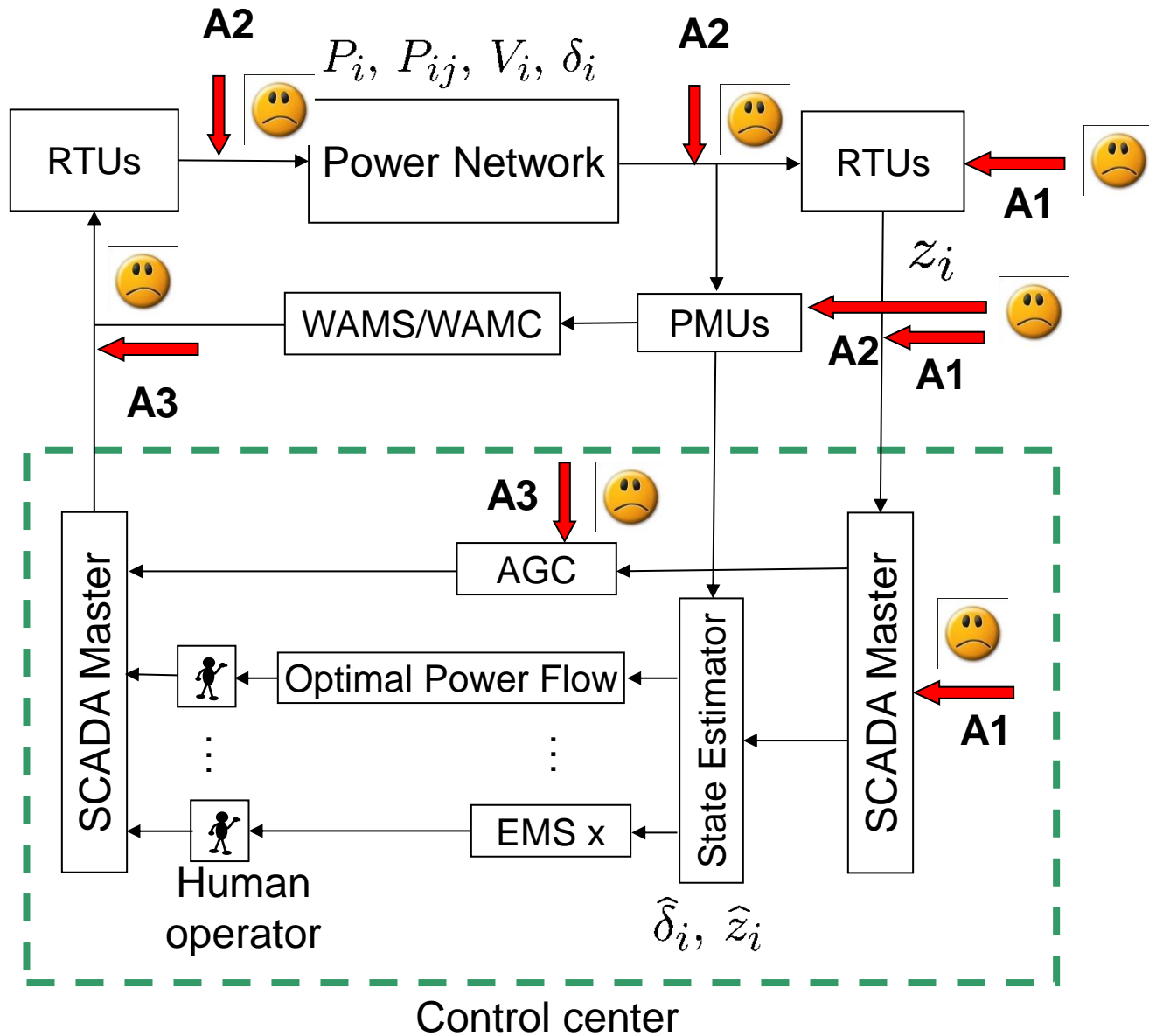
## Impact Analysis of AGC attack

- P. Mohajerin Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, G. Andersson, "A robust policy for Automatic Generation Control cyber attack in two area power network," IEEE Conference on Decision and Control (CDC), 2010
- P. Mohajerin Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, G. Andersson, "Cyber attack in a two-area power system: Impact identification using reachability," American Control Conference (ACC), 2010

## Countermeasures, Fault Detection and Isolation

- P. Mohajerin Esfahani, T. Sutter, J. Lygeros, "Performance Bounds for the Scenario Approach and an Extension to a Class of Non-convex Programs", Submitted for publication, 2013, arXiv:1307.0345.
- P. Mohajerin Esfahani, J. Lygeros, "A Tractable Fault Detection and Isolation Approach for Nonlinear Systems with Probabilistic Performance", Submitted for publication, Feb. 2013, [Tech. Rep.](#)

# Scenarios Considered





# Conclusions

- Power systems need **cyber-physical security**
- Control engineering can contribute to
  - Estimate of impact of attacks
  - Identify critical resources
  - Synthesis of novel attack detection schemes
- Three advanced cyber attacks against power system considered
  - What about the possible simple attacks?