University of L’Aquila
Department of Information Engineering, Computer Science and Mathematics
Center of Excellence DEWS

Diagnosability of Hybrid Dynamical Systems

Maria Domenica Di Benedetto

University of L’Aquila

Many thanks!

- Elena De Santis

- Giordano Pola
- Gabriella Fiore

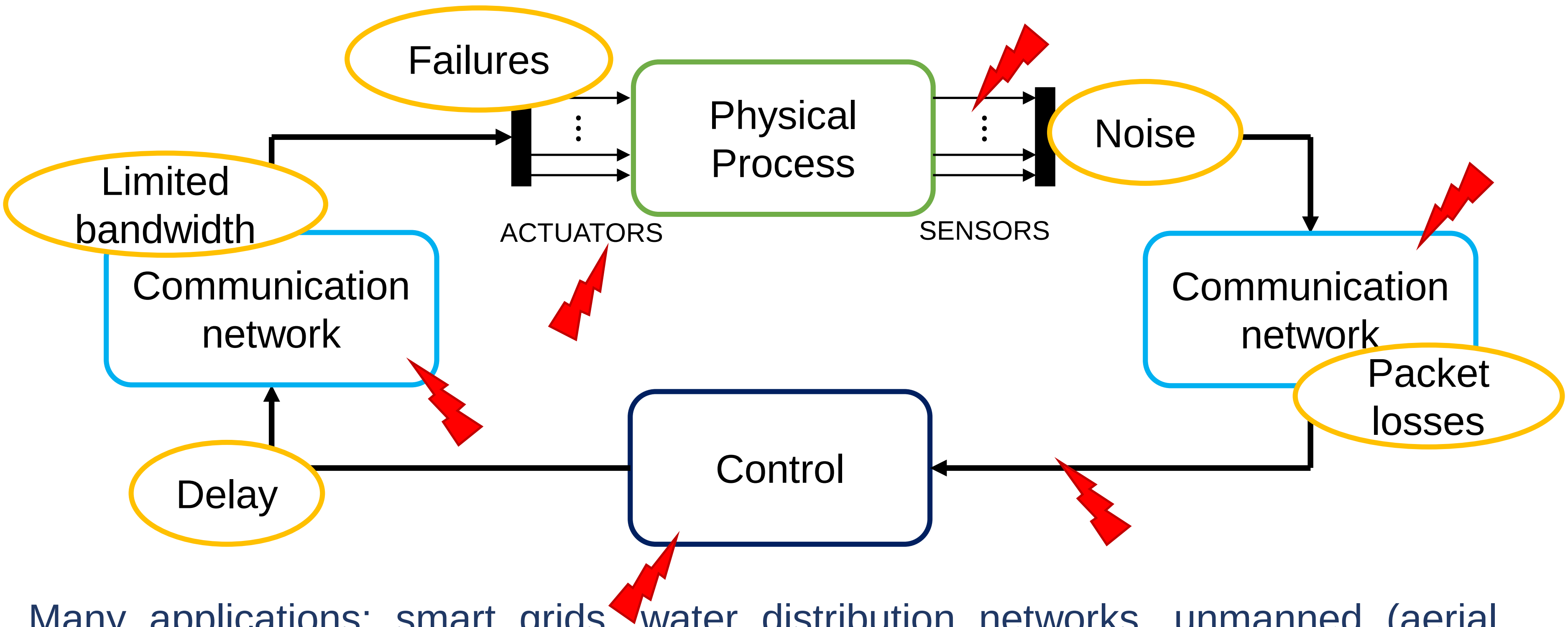
- Andrea Balluchi
- Luca Benvenuti
- Alberto Sangiovanni Vincentelli

Outline

- Motivation
  - Cyber-Physical Systems (CPS)
  - Security for CPS
- Modeling CPS as hybrid systems
- Secure state estimation for hybrid systems
  - Observability and diagnosability
  - Secure mode distinguishability
  - Secure diagnosability
  - Approximate diagnosability
- Conclusions and future work

Cyber-Physical Systems

Cyber-Physical Systems (CPSs) integrate physical processes, computational resources and communication capabilities.



Many applications: smart grids, water distribution networks, unmanned (aerial, ground, underwater) vehicles, biomedical and health care devices, air traffic management systems, and many others.

Security of CPS

SECURITY 6/01/2012 @ 3:59PM | 26,186 views

What Stuxnet's Exposure As An American Weapon Means For Cyberwar

How vulnerable are UAVs to cyber attacks?

Kevin G. Coleman, SilverRhino 11:50 a.m. EST February 23, 2015

Cyberattack Inflicts Massive Damage on German Steel Factory

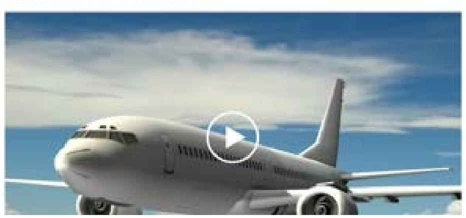
POSTED BY: PAUL DECEMBER 21, 2014 14:52 1 COMMENT

Keeping your car safe from hacking

Automakers and NHTSA scramble to protect your privacy and safety

Published: May 07, 2015 06:00 AM

FBI: Hacker claimed to have taken over flight's engine controls

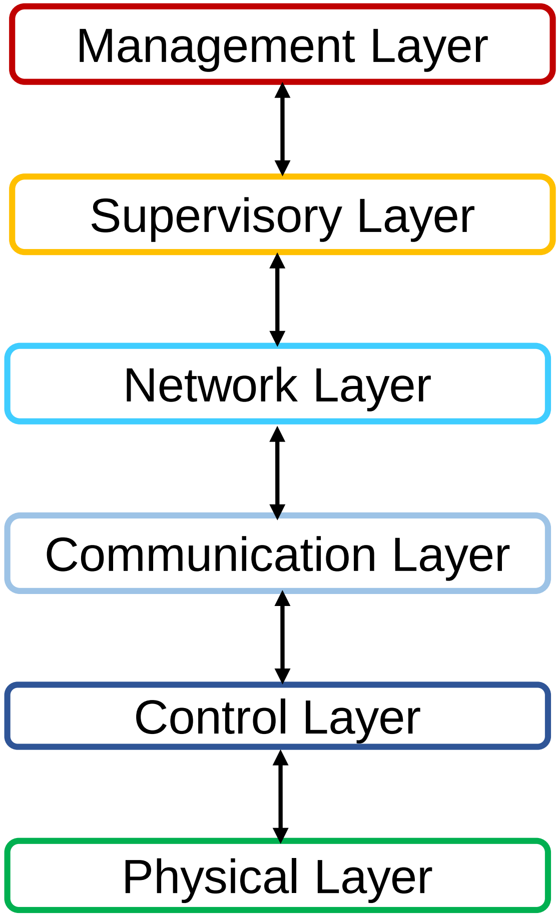


Emerging Technology From the arXiv
April 24, 2015

Security Experts Hack Teleoperated Surgical Robot

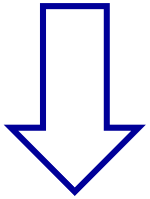
The first hijacking of a medical telerobot raises important questions over the security of remote surgery, say computer security experts.

Security of CPSs



[Q. Zhu and T. Basar, 2015]

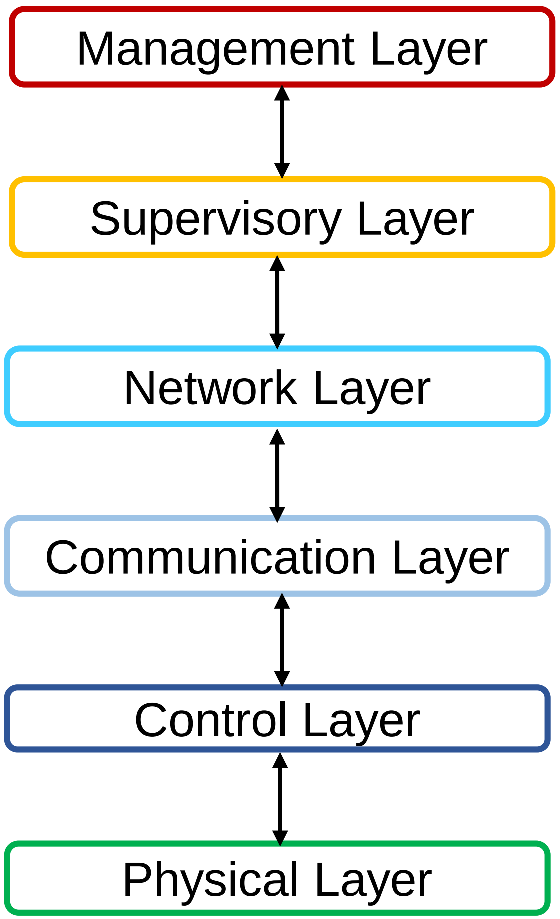
Security measures protecting only the computational and communication layers are **necessary but not sufficient** for guaranteeing the safe operation of the entire system

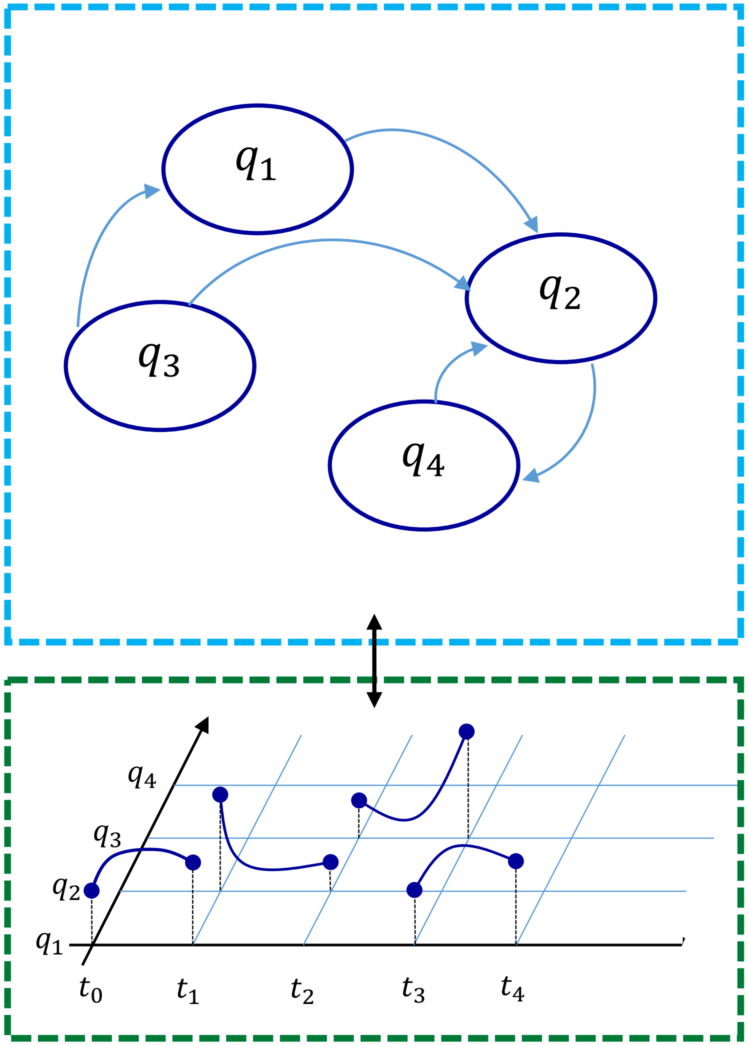


Exploit also system dynamics to

- assess correctness and compatibility of measurements,
- ensure robustness and resilience with respect to malicious attacks.

CPSs modeled as hybrid systems

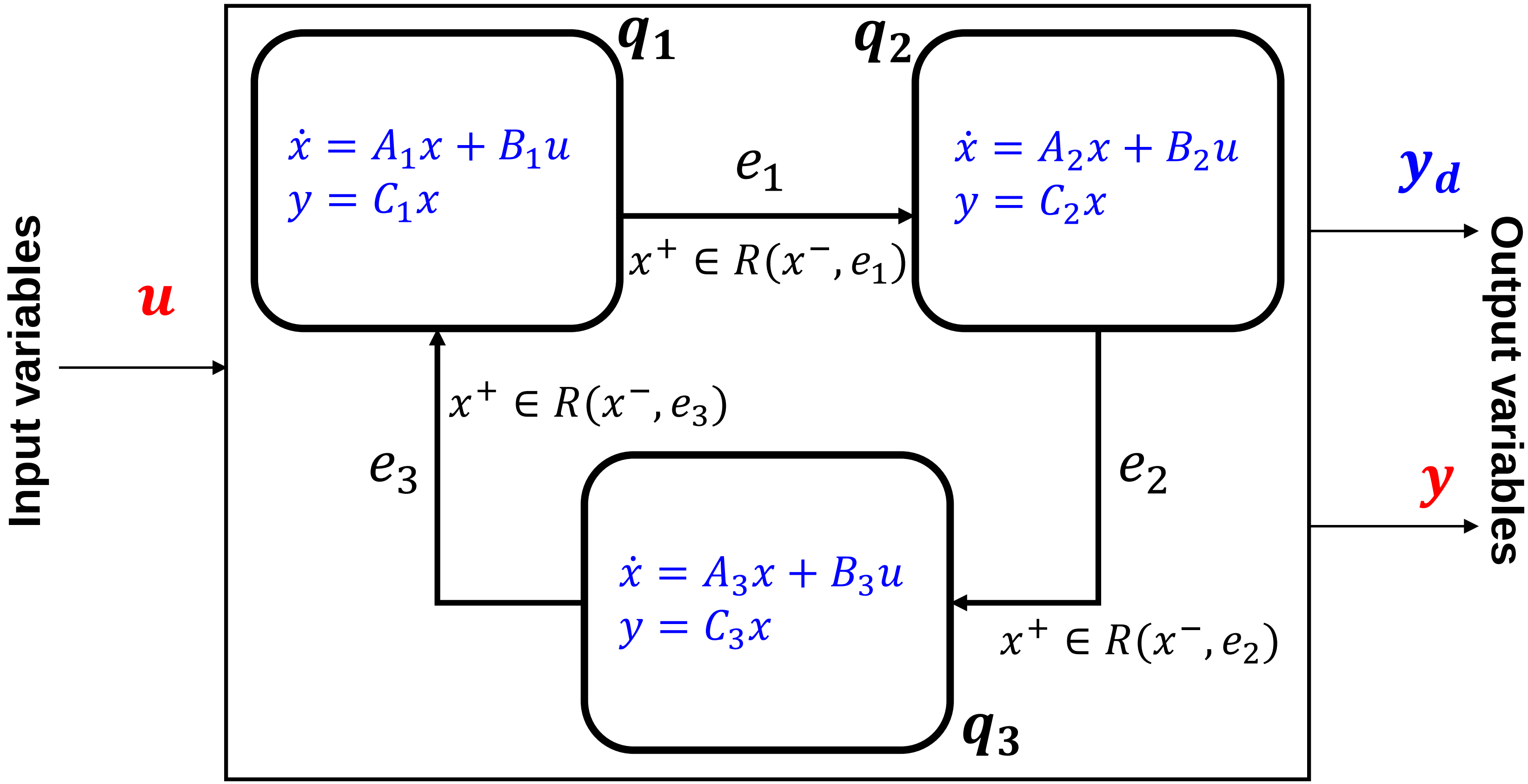




Outline

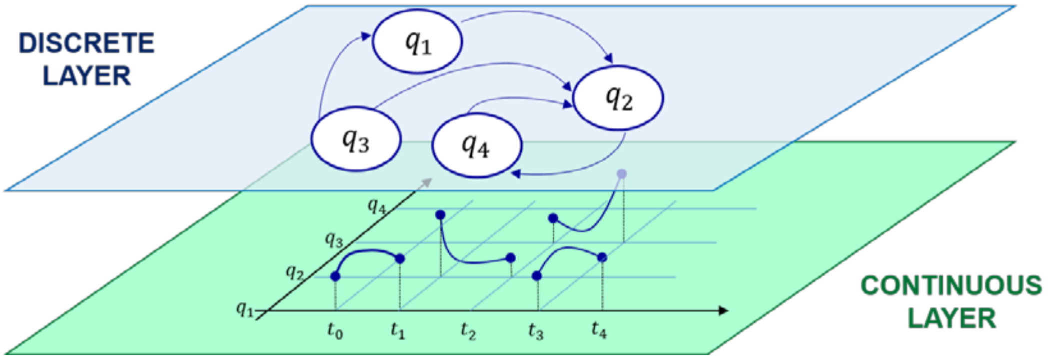
- Motivation
  - Cyber-Physical Systems (CPS)
  - Security for CPS
- Modeling CPS as hybrid systems
- Secure state estimation for hybrid systems
  - Observability and diagnosability
  - Secure mode distinguishability
  - Secure diagnosability
  - Approximate diagnosability
- Conclusions and future work

Linear Hybrid systems



Hybrid system modeling framework

Definition. An H-system is a tuple: $\mathcal{H} = (\Xi, \Xi_0, \Upsilon, h, S, E, G, R, \delta, \Delta)$



- $\Xi = Q \times X$ hybrid state space
- $\Xi_0 \subseteq \Xi$ set of initial hybrid states
- $\Upsilon = Y_d \times \mathbb{R}^p$ hybrid output space
- $h: Q \to Y_d$ discrete output function
- S associates to each discrete state a dynamical system $S(i)$ described by:

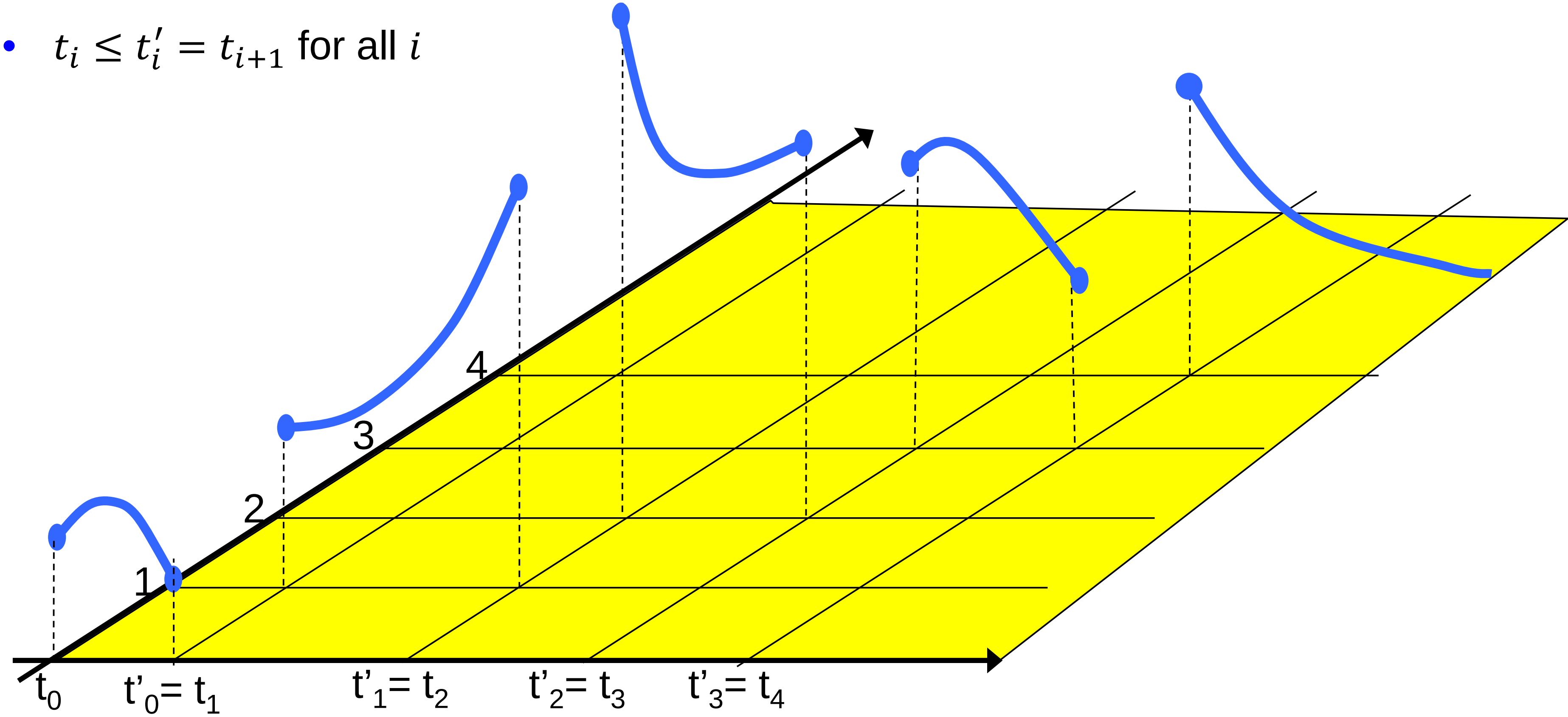
$$\begin{cases} \dot{x}_i = A_i x(t) + B_i u(t) \\ y(t) = C_i x(t) \end{cases}$$

- $E \subseteq Q \times Q$ admissible discrete transitions
- $G: E \to 2^X$ guard
- $R: E \times X \to 2^X$ reset
- $\delta: Q \to \mathbb{R}^+$ minimum dwell time associated to $i \in Q$
- $\Delta: Q \to \mathbb{R}^+ \cup \{\infty\}$ maximum dwell time associated to $i \in Q$

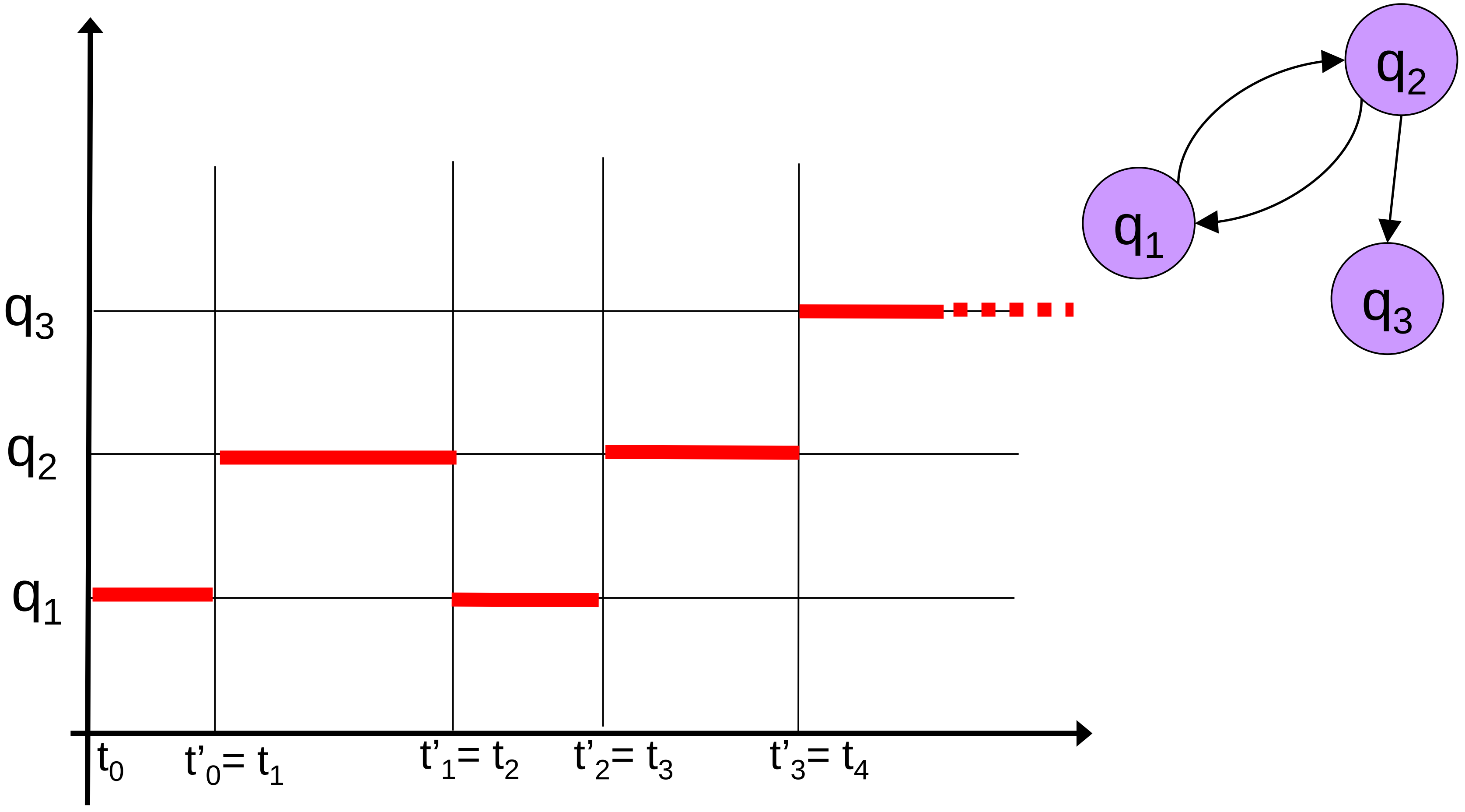
Continuous State Evolution

Definition: A **hybrid time basis** is a sequence of intervals $\tau = \{I_0, I_1, \ldots, I_N\} = \{I_i\}_{i=0}^{N}$, with $N < \infty$ or $N = \infty$, $I_i = [t_i, t_i']$ for all $i < N$ such that

- if $N < \infty$ then either $I_N = [t_N, t_N']$ or $I_N = [t_N, t_N')$
- $t_i \leq t_i' = t_{i+1}$ for all i

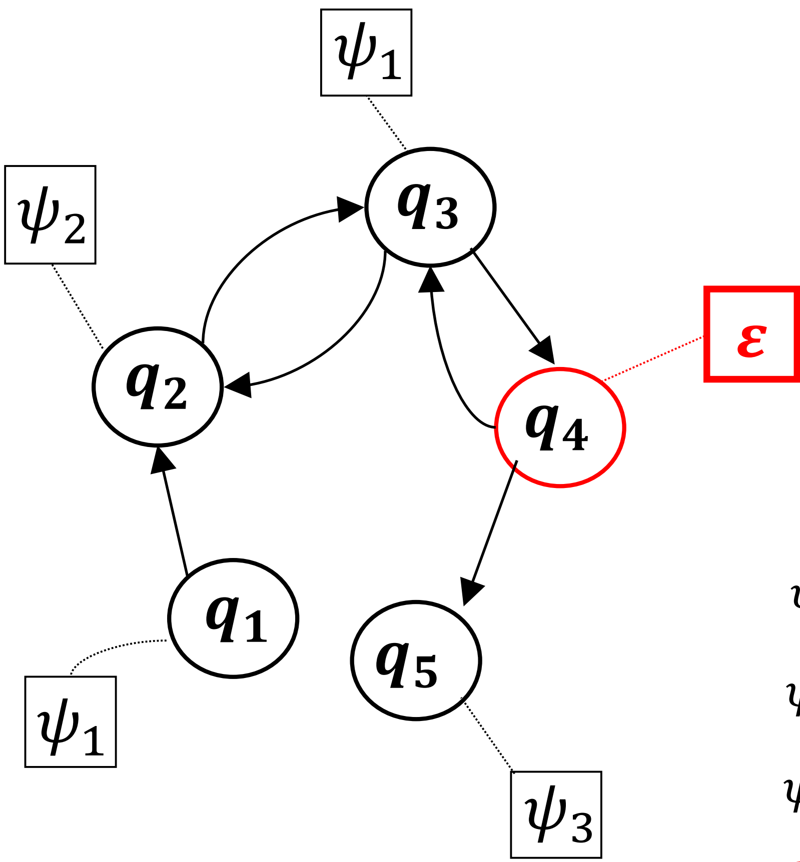


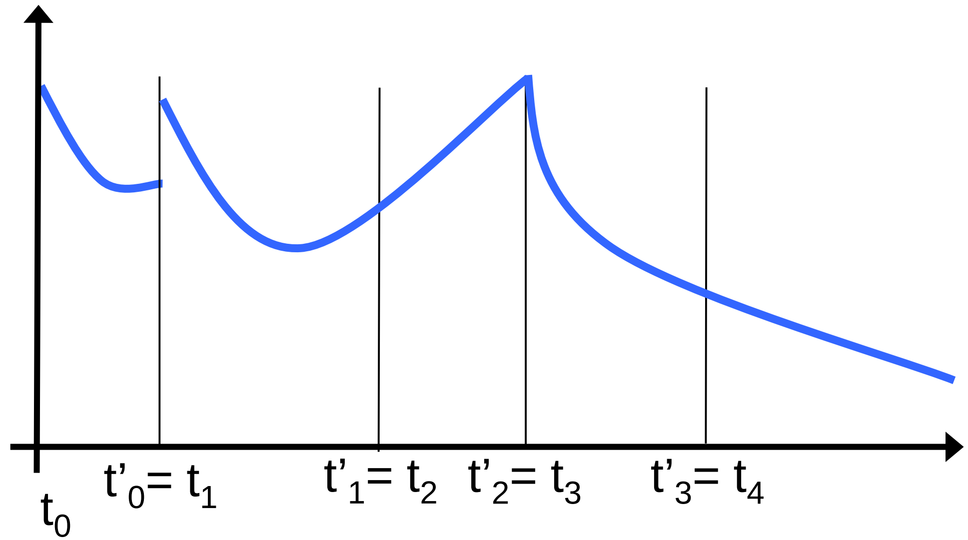
Discrete State Evolution

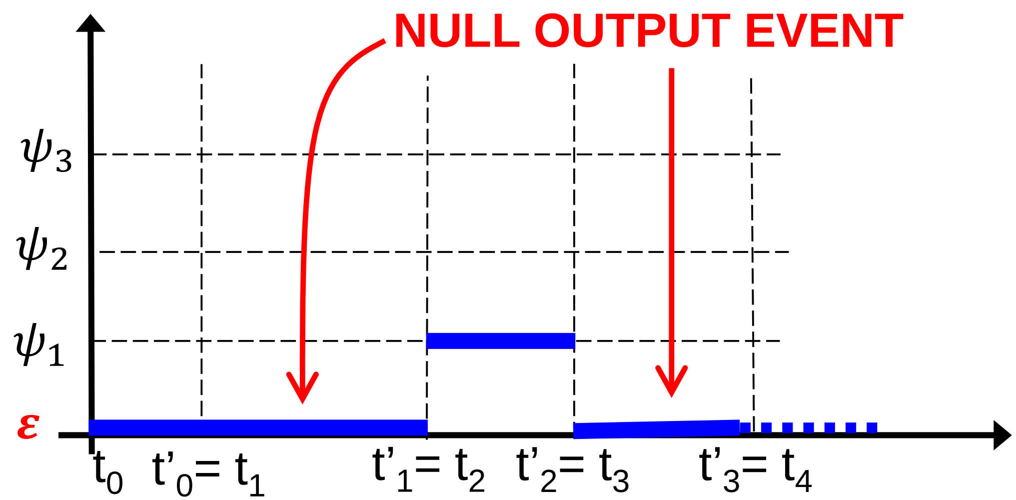


Observed output

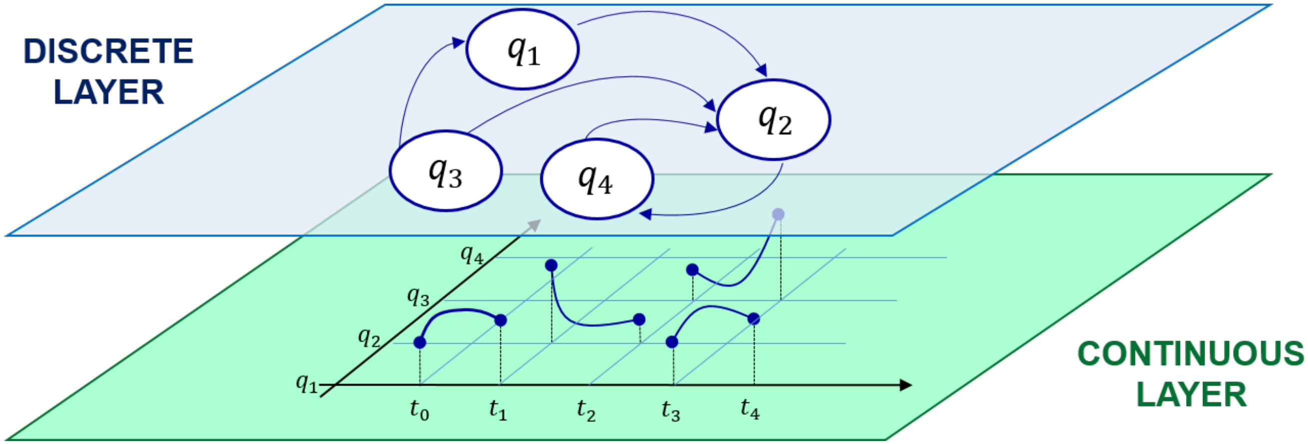
$h: Q \to Y$ is the **discrete output function**, where Y is the discrete output space







Observablity and diagnosability of H-systems



$\Xi = Q \times X$

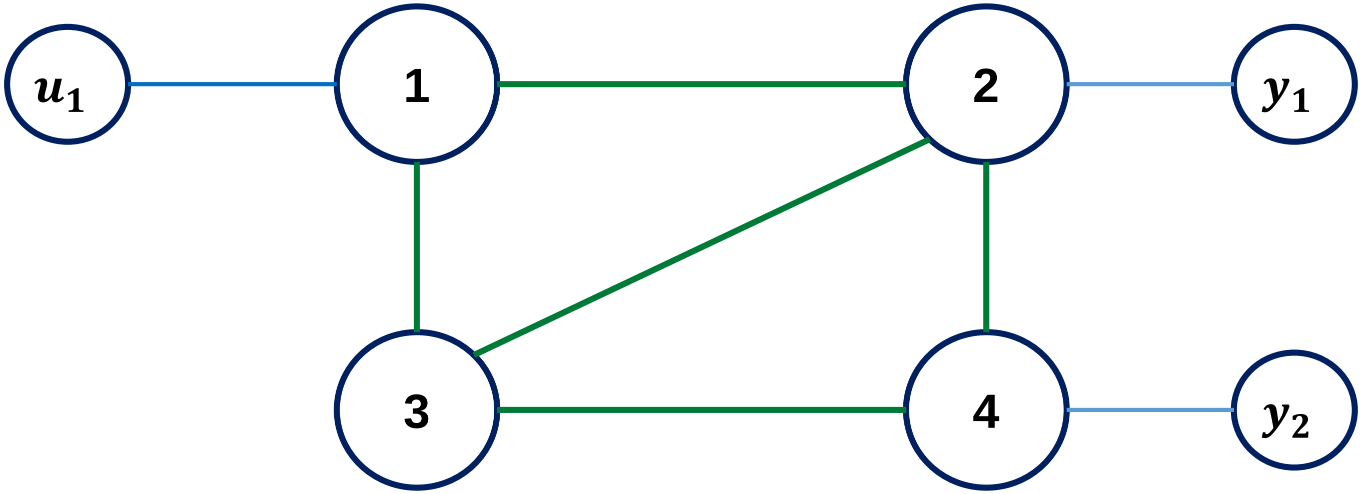
- **Observability:** possibility of determining the current discrete state and the continuous state, on the basis of the observed output information.

- **Diagnosability:** possibility of detecting the occurrence of particular subsets of hybrid states, for example faulty states, on the basis of the observations, within a finite time interval.

Outline

- Motivation
  - Cyber-Physical Systems (CPS)
  - Security for CPS
- Modeling CPS as hybrid systems
- Secure state estimation for hybrid systems
  - Observability and diagnosability
  - Secure mode distinguishability
  - Secure diagnosability
  - Approximate diagnosability
- Conclusions and future work

Observability and resilience: example 1

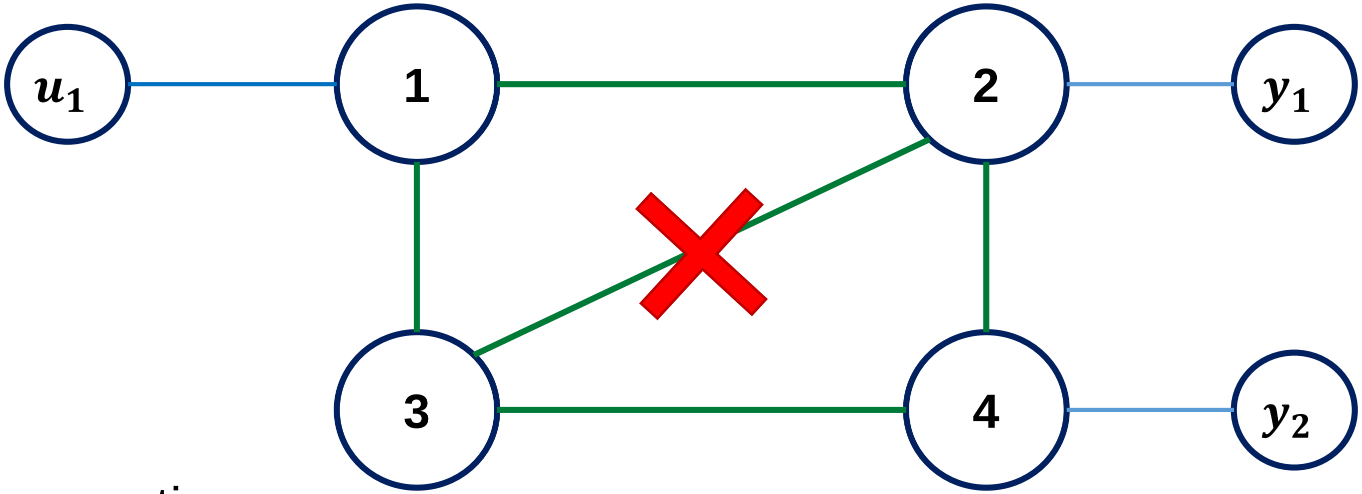


$$x(t+1) = -Lx(t) + Bu(t)$$
$$y = Cx(t)$$

$$l_{ij} = \begin{cases} 1 & j \in \mathcal{N}_i \\ -|\mathcal{N}_i| & j = i \\ 0 & \text{otherwise} \end{cases}$$

$$L = \begin{bmatrix} -2 & 1 & 1 & 0 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 0 & 1 & 1 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Observability and resilience: example 1



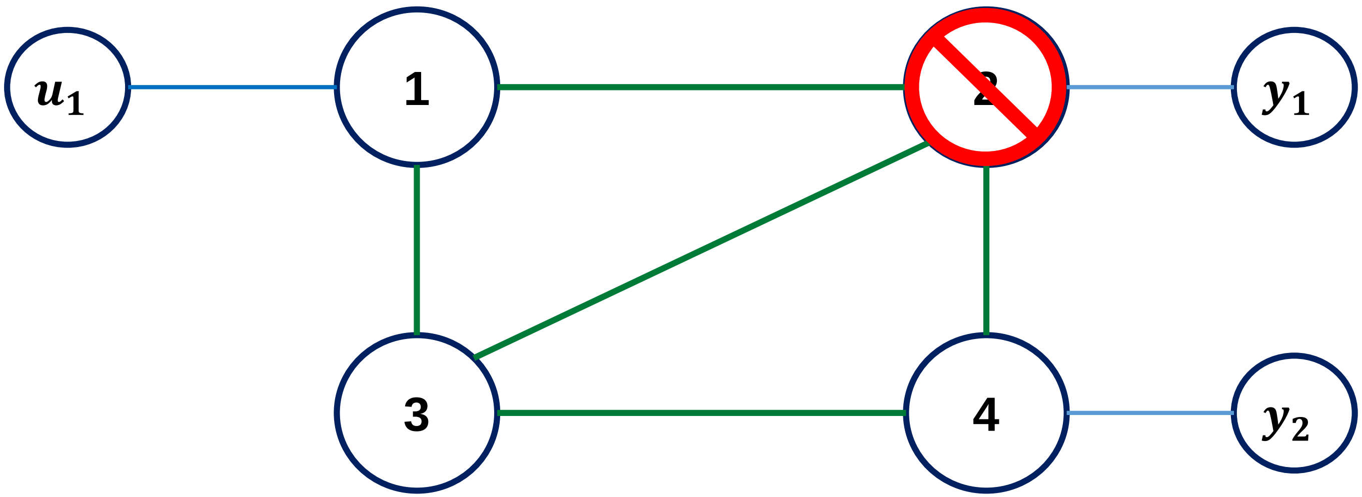
Link disconnection:

$$x(t+1) = -\bar{L}x(t) + Bu(t)$$
$$y = Cx(t)$$

$$l_{ij} = \begin{cases} 1 & j \in \mathcal{N}_i \\ -|\mathcal{N}_i| & j = i \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{L} = \begin{bmatrix} -2 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \\ 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Observablity and resilience: example 1



Node disconnection:

$$x(t+1) = -\bar{L}x(t) + \bar{B}u(t)$$
$$y = \bar{C}x(t)$$

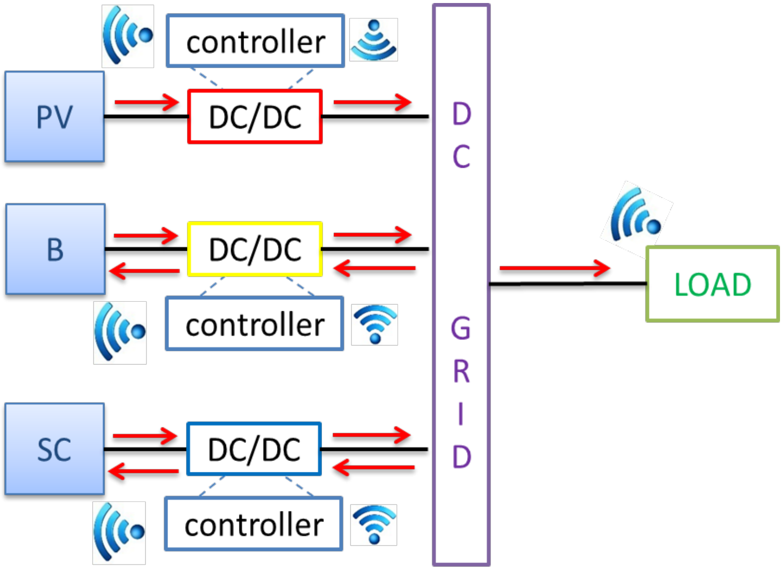
$$l_{ij} = \begin{cases} 1 & j \in \mathcal{N}_i \\ -|\mathcal{N}_i| & j = i \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{L} = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \qquad \bar{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \bar{C} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

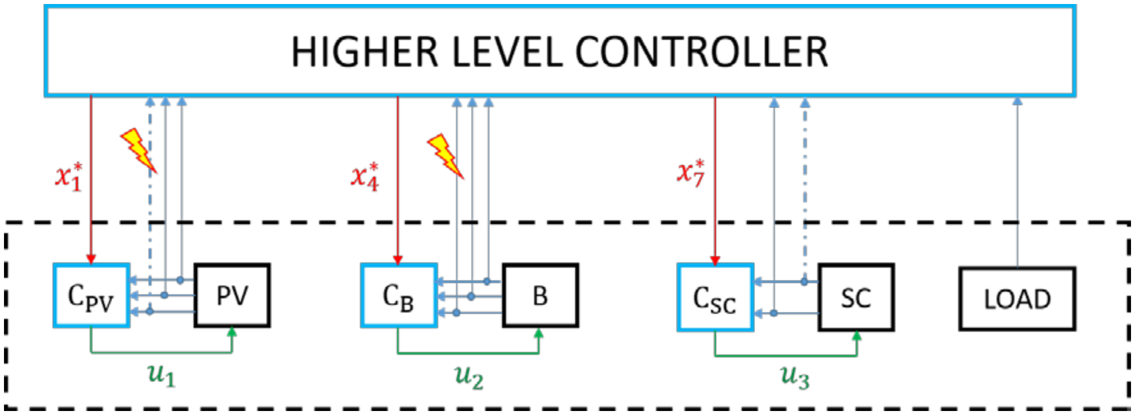
Observability and resilience: example 2

Objectives:

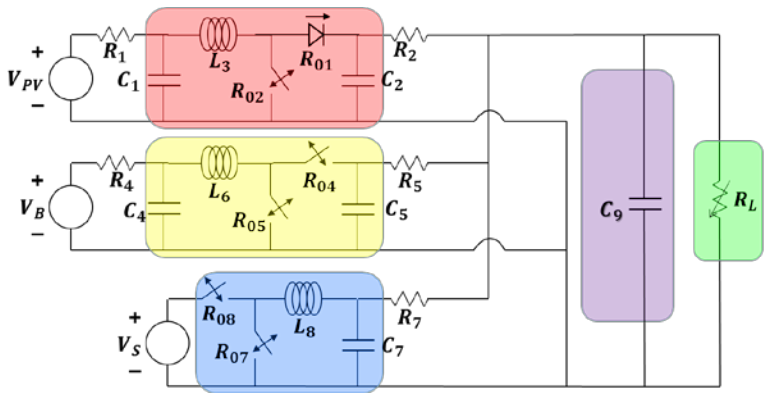
- Extract the maximum available power from renewable sources
- Provide/absorb the power when needed by means of the battery
- Stabilize grid and load voltage (also in case of disturbances)



[Iovine et al. 2017]



Observability and resilience: example 2



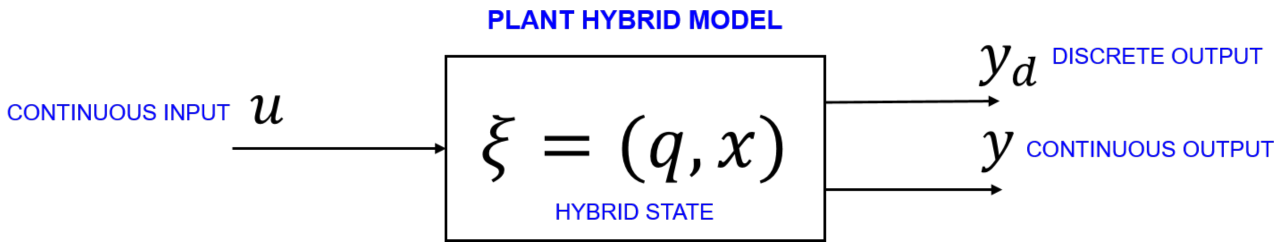
Linearized digital model

$$S = \begin{cases} x(k+1) = Ax(k) + [B_b \quad D]\begin{bmatrix} b(k) \\ d_x(k) \end{bmatrix} = Ax(k) + Bu(k) \\ y(k) = Cx(k) + w(k) \end{cases}$$

→ Sparse attack $w(k) \in \mathbb{S}_\sigma^p$

$$x(k) \in \mathbb{R}^n, u(k) \in \mathbb{R}^m, y(k) \in \mathbb{R}^p$$

Observability of H-systems



Definition. The system H is **observable** if there exists a function $\hat{\xi}: \Upsilon \times \mathrm{U} \to \Xi$ which, by setting

$$\hat{\xi}\left(\eta|_{[0,t]}, \hat{u}|_{[0,t]}\right) = \left(\hat{q}(t), \hat{x}(t)\right)$$

satisfies the following conditon:

- there exists $\hat{t} > 0$ such that:
  - $\widehat{q}(t) = q(t) \qquad \forall \;\; t > \hat{t}$
  - $\|\hat{x}(t) - x(t)\| = \; 0 \qquad \forall \;\; t > \hat{t}$

DETERMINATION OF THE HYBRID STATE

for any generic input $\hat{u} \in U$, for any execution χ with $u = \hat{u}$.

Role of the input

For an input $u \in \mathcal{U}$, with $\mathcal{U}$ set of piecewise continuous functions, define the norm of u as:

$$||u|| = sup_{t \in \mathbb{R}} ||u(t)||$$

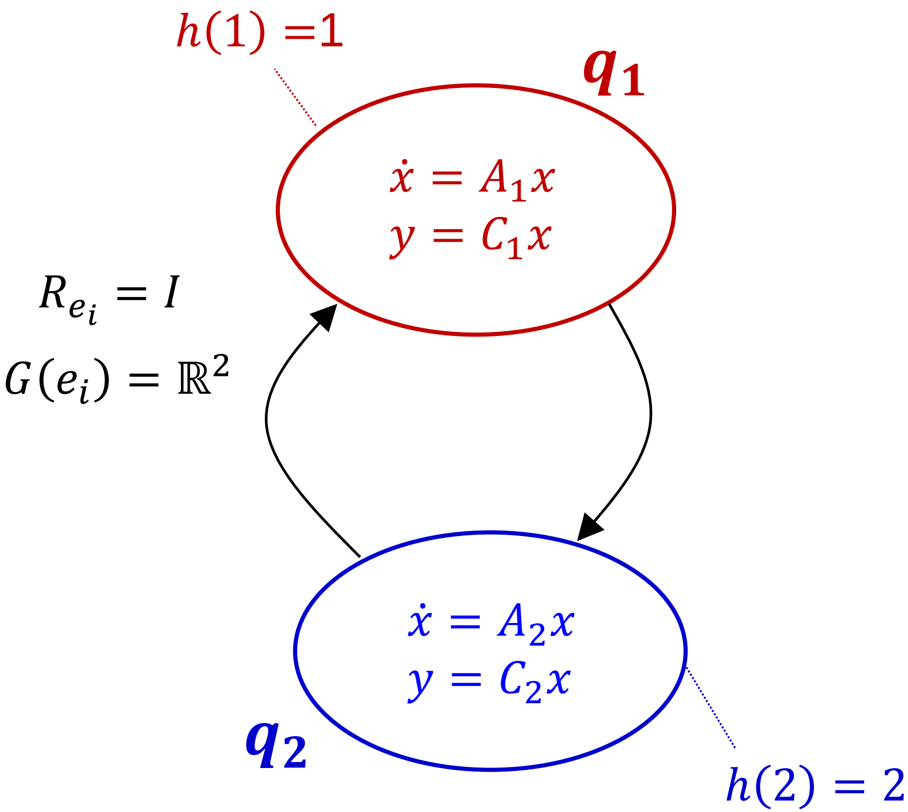
where $||u(t)||$ standard Euclidean norm of the vector $u(t)$ in the space $\mathbb{R}^m$.

A **generic input** $\hat{u} \in \mathcal{U}$ is any input function that belongs to a dense subset of the set $\mathcal{U}$ equipped with the above defined norm.

Role of dwell time

Is observability of each pair (A_i, C_i) necessary and sufficient for the observability of H?

Example:



$$x \in \mathbb{R}^2, \Delta(\boldsymbol{i}) = \Delta \neq \infty$$

$$h(i) = i, \qquad \forall i \in Q$$

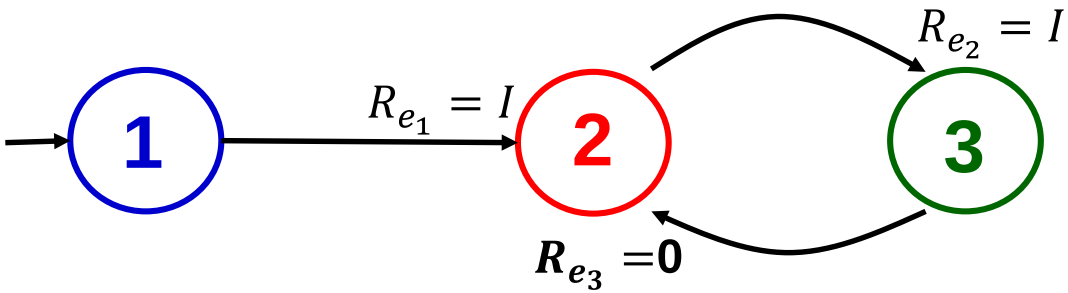
$$S(1) = \begin{cases} \dot{x}_1 = x_1 \\ \dot{x}_2 = x_2 \\ \boldsymbol{y = x_1} \end{cases} \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S(2) = \begin{cases} \dot{x}_1 = x_1 \\ \dot{x}_2 = x_2 \\ \boldsymbol{y = x_2} \end{cases} \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad C_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

The pairs (A_i, C_i) are not observable, however H is observable!

Role of reset, graph topology

Example:



$x \in \mathbb{R}^2, \Delta(\boldsymbol{i}) = \Delta \neq \infty$

$h(i) = i, \qquad \forall i \in Q$

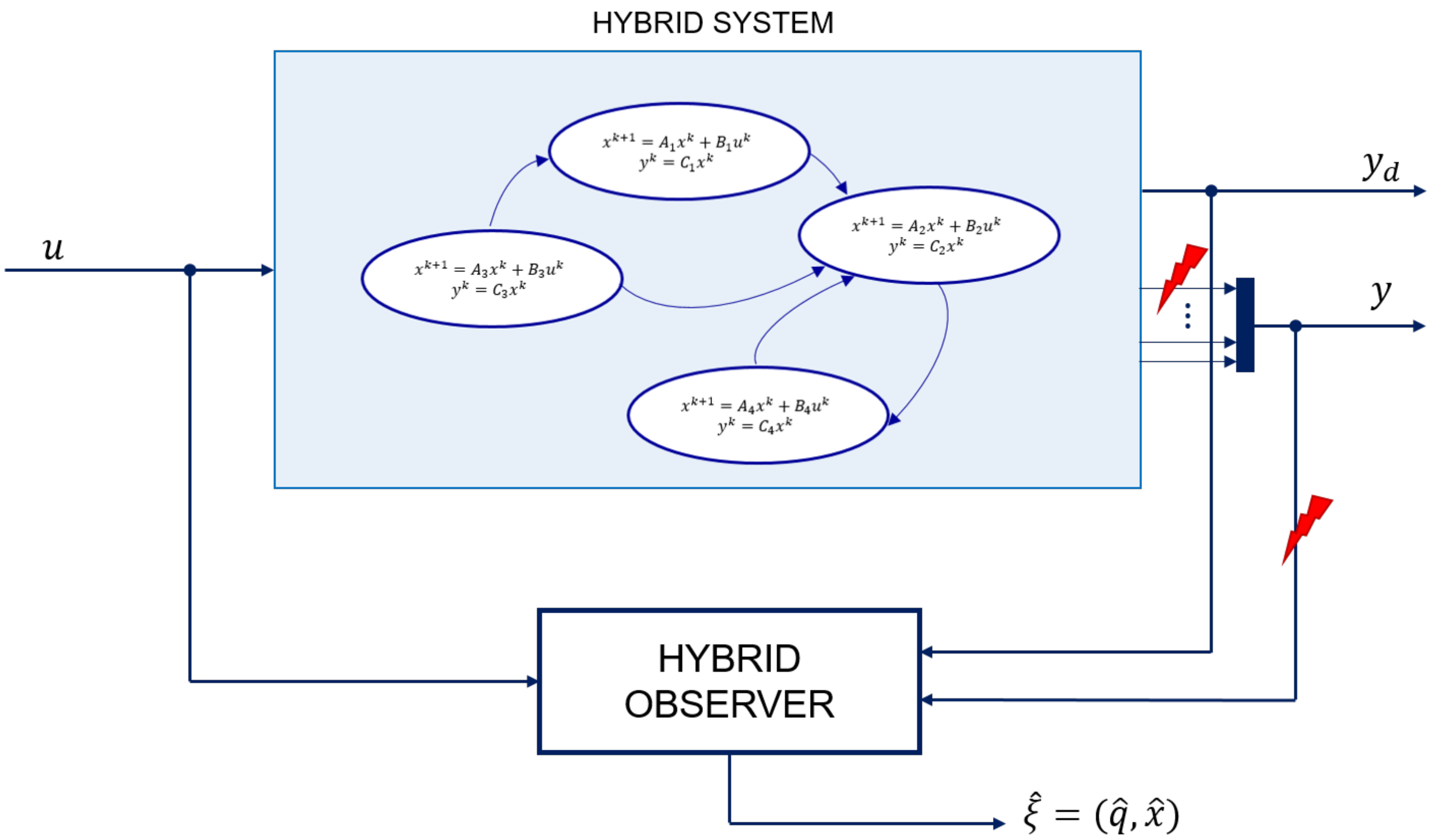
$A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ $\quad A_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ $\quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

$C_1 = [0 \quad 0]$ $\quad C_2 = [0 \quad 0]$ $\quad C_3 = [0 \quad 0]$

The pairs (A_i, C_i) are not observable

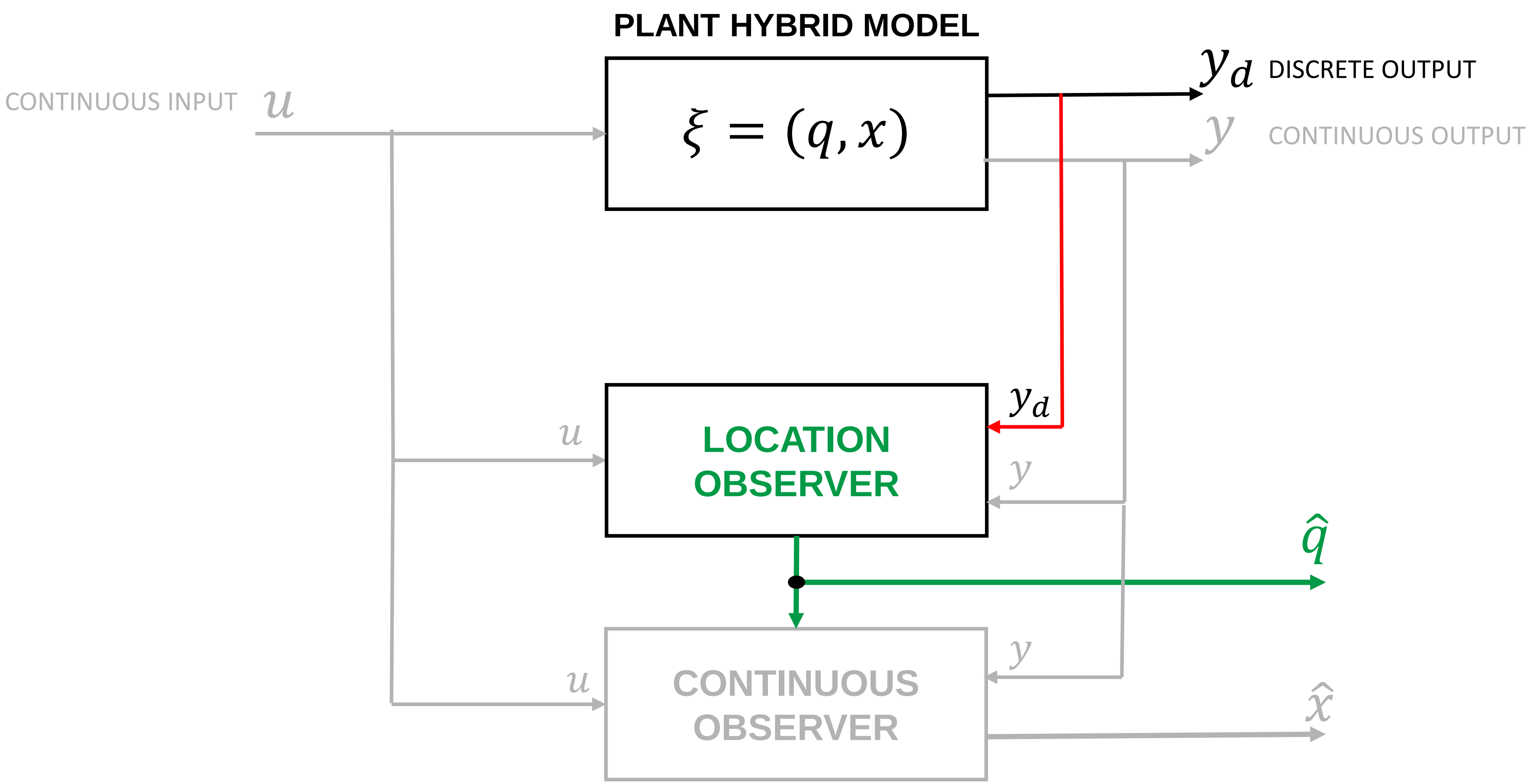
At most after $\mathbf{3}\Delta$ units of time the state is equal to **0** because of the reset function definition. Hence, H is observable!

State estimation of H-systems



Location observer design

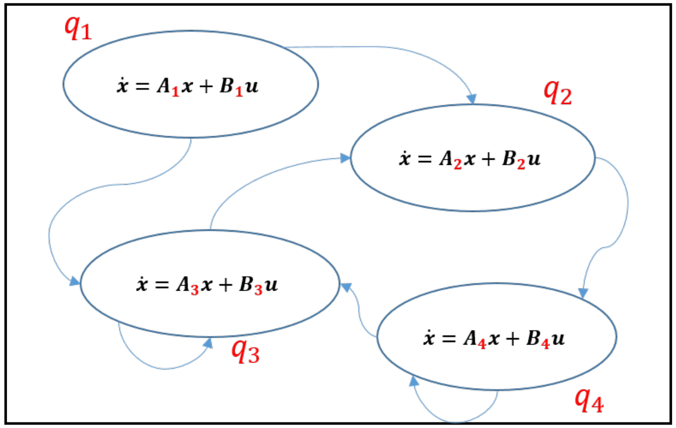
Goal: Determine current **discrete state** of H by using discrete output information either independently from continuous output evolution or by using also continuous evolution.



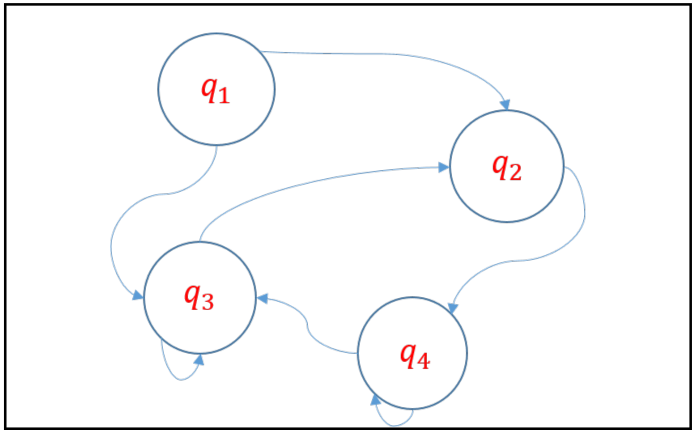
Discrete information only

Finite state machine associated to H

HYBRID SYSTEM $\mathrm{H} = (\Xi = (\boldsymbol{Q}, X), \Xi_0 = (\boldsymbol{Q_0}, X_0), \Upsilon = (\boldsymbol{Y}, \mathbb{R}^p), \boldsymbol{h}, S, \boldsymbol{E}, G, R, \delta, \Delta)$



FINITE STATE MACHINE



Nondeterministic **finite state machine** (FSM) that abstracts the dependence of the discrete dynamics of *H* from its continuous evolution:

$$M = (Q, Q_0, Y, h, E)$$

Finite state machine associated to H

$$M = (Q, Q_0, Y, h, E)$$

Given the evolution in time of the H-system $\chi = (q_0, \tau, q)$, where τ is a time basis with $\mathrm{card}(\tau) = L$, the **event-based evolution** of the FSM is a string σ

- State execution of M:

$$\sigma(1) \in Q$$
$$\sigma(k) = q(t_{k-1}), \qquad k = 1,2, \dots, L$$
$$\sigma(k+1) \in succ\big(\sigma(k)\big), \qquad k = 1, \dots, L-1$$

- $\mathcal{X}^*$ set of all state executions
- $\mathcal{X}$ set of infinite state executions with $\sigma(1) \in Q_0$
- Liveness: $succ(i) \neq \emptyset \quad \forall\ i \in Q$
- Discrete output of M:

$$h\big(\sigma(k)\big) = h\big(q(t_{k-1})\big) = y_d(t_{k-1})$$

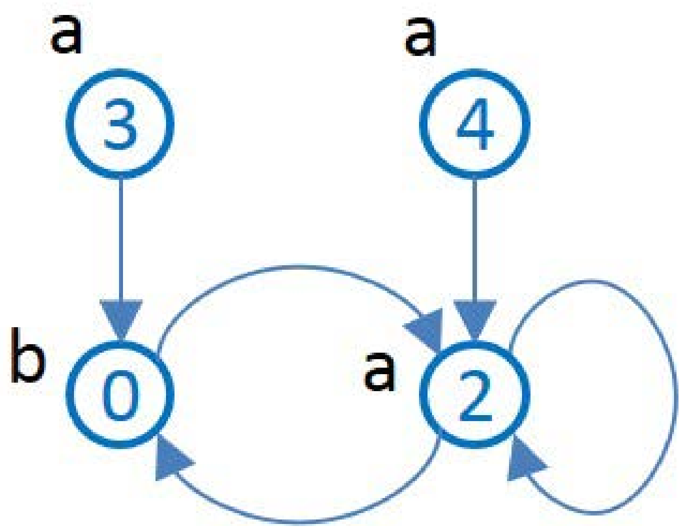
- Output string of M:

$$\mathbf{h}: \mathcal{X}^* \to (Y \setminus \{\varepsilon\})^*$$

where for $\sigma \in \mathcal{X}^*$, $\mathbf{h}(\sigma) = P(s)$, $s = (h\big(\sigma(1)\big) \dots h\big(\sigma(|\sigma|)\big))$
where for an output string $s \in Y^*$, $P(s)$ denotes the string obtained from s by erasing all ε symbols.

Current location observability of M

Definition: The FSM M is **current location observable** if there exists $\bar{k} \in Z$, such that for any string $\sigma \in \mathcal{X}$ with unknown $\sigma(1) \in Q_0$, the knowledge of the output string $\mathbf{h}(\sigma|_{[1,k]})$ makes it possible to infer that $\sigma(k) = i$, for some $i \in Q$, for all $k \geq \bar{k}$.



Current location observable!

[Ramadge, CDC 1986]

Current location observability

Theorem. The FSM M is **current location observable** if and only if for every persistent state $i \in Q_p$ of M:

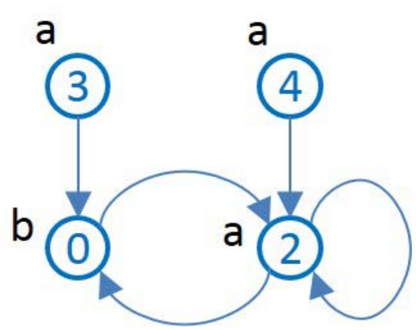
1) $h(i) \neq \varepsilon$;
2) there exists a singleton state $\{i\}$ in the observer O_M and it is the only persistent state of O_M containing i.

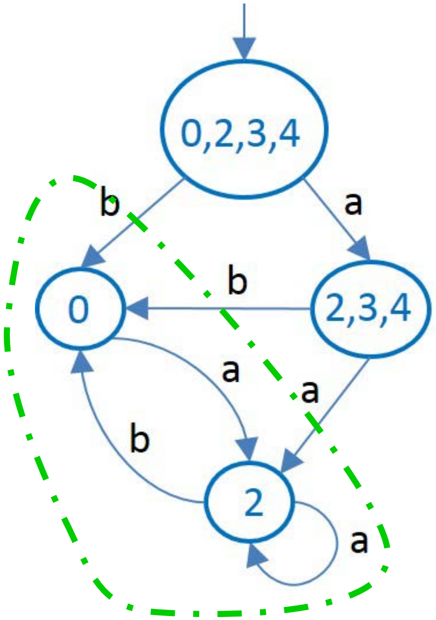
$\Downarrow$

M and O_M have the same set of persistent states!

M

$h(i) \neq \varepsilon$;





O_M

Current location observability of H (using discrete output only)

H-system

Current location observability

$\Delta < \infty$

=

FSM

Current location observability

Assuming **finite maximum dwell time**, current location observability of M is equivalent to current location observability of H.

Current location observability of H (using discrete output only)

H-system | **FSM**

$\Delta < \infty$

Current location observability = Current location observability

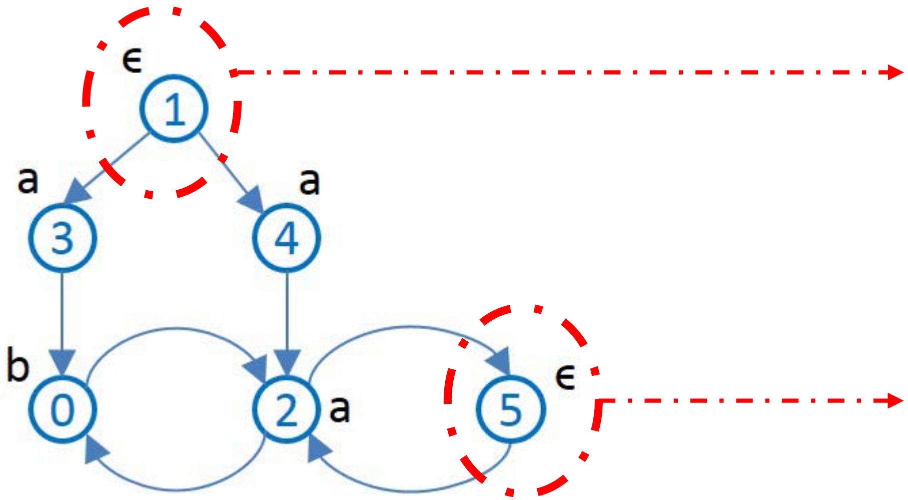
What if the maximum dwell time is $\Delta = \infty$**?**

Critical location observability is needed!

Critical observability of M

Definition: The FSM M is $\{\boldsymbol{i}\} -$ **critically location observable** if, for any $k \in Z$, whenever $\sigma(k) = i$, the knowledge of the output string $\mathbf{h}\left(\sigma|_{[1,k]}\right)$ makes it possible to infer that $\sigma(k) = i$. If M is $\{i\} -$ critically location observable for all $i \in Q$, then it is called **critically location observable**.

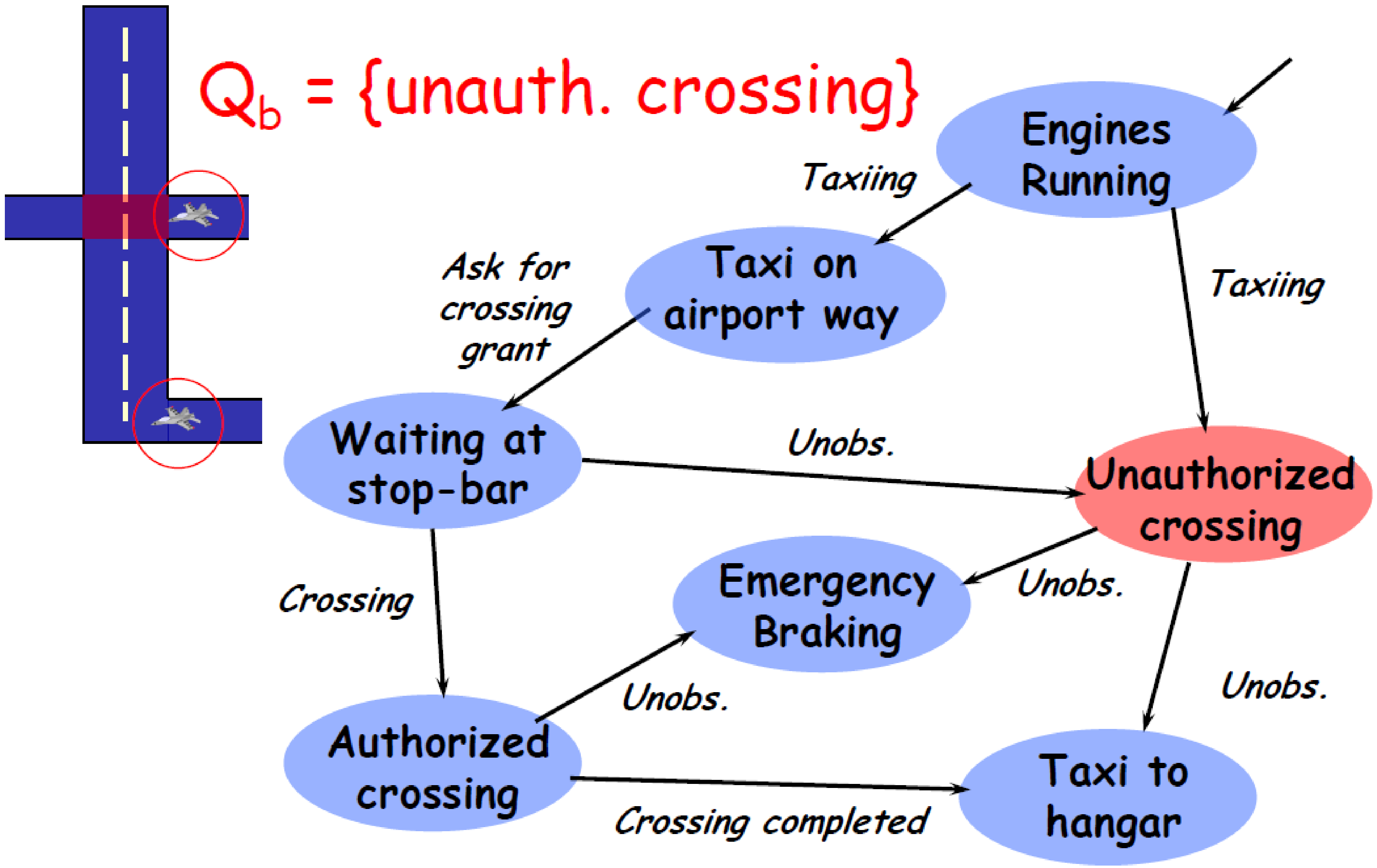
Theorem: The FSM M is $\{\boldsymbol{i}\} -$ **critically location observable** only if $h(i) \neq \varepsilon$.



Not $\{1\} -$ critically location observable

Not $\{5\} -$ critically location observable

Observability of critical states



Critical observability of H

Definition. The H-system is $\{i\}$ **−critically location observable** if there exists a function $\hat{\xi}: \Upsilon \times \mathrm{U} \rightarrow \Xi$ such that, by setting

$$\hat{\xi}\left(\eta|_{[0,t]}, \hat{u}|_{[0,t]}\right) = \left(\hat{q}(t), \hat{x}(t)\right)$$

whenever $q(t_k) = i$

$$\widehat{q}(t) = i \qquad\qquad \forall \;\; t \in (t_k, t_{k+1})$$

for any generic input $\hat{u} \in U$ and for any execution χ with $u = \hat{u}$.

The H-system is **critically location observable** if it is $\{\boldsymbol{i}\}$−critically location observable for all $i \in Q$.

Theorem. The H-system is **critically location observable** if and only if it is current location observable with $\hat{t} = 0$.

Current location observability of H
(using discrete output only)

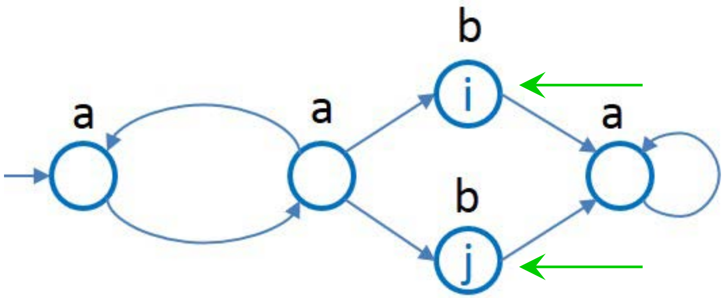
H-system		**FSM**
Current location observability	$\Delta < \infty$ =	Current location observability
$\{i\}$ – critical location observability	=	$\{i\}$ – critical location observability
Current location observability	= $\Delta = \infty$	✓ Current location observability ✓ $\{i\}$ – critical location observability $\forall i \in reach(Q_\infty)$

H-system is current location observable only if $h(i) \neq \varepsilon$, for all "persistent in time" states $i \in Q_p \cup reach(Q_\infty)$.

Current location observability
(mixed continuous and discrete information)

Question: What if the discrete output information is not sufficient to estimate the current discrete location?

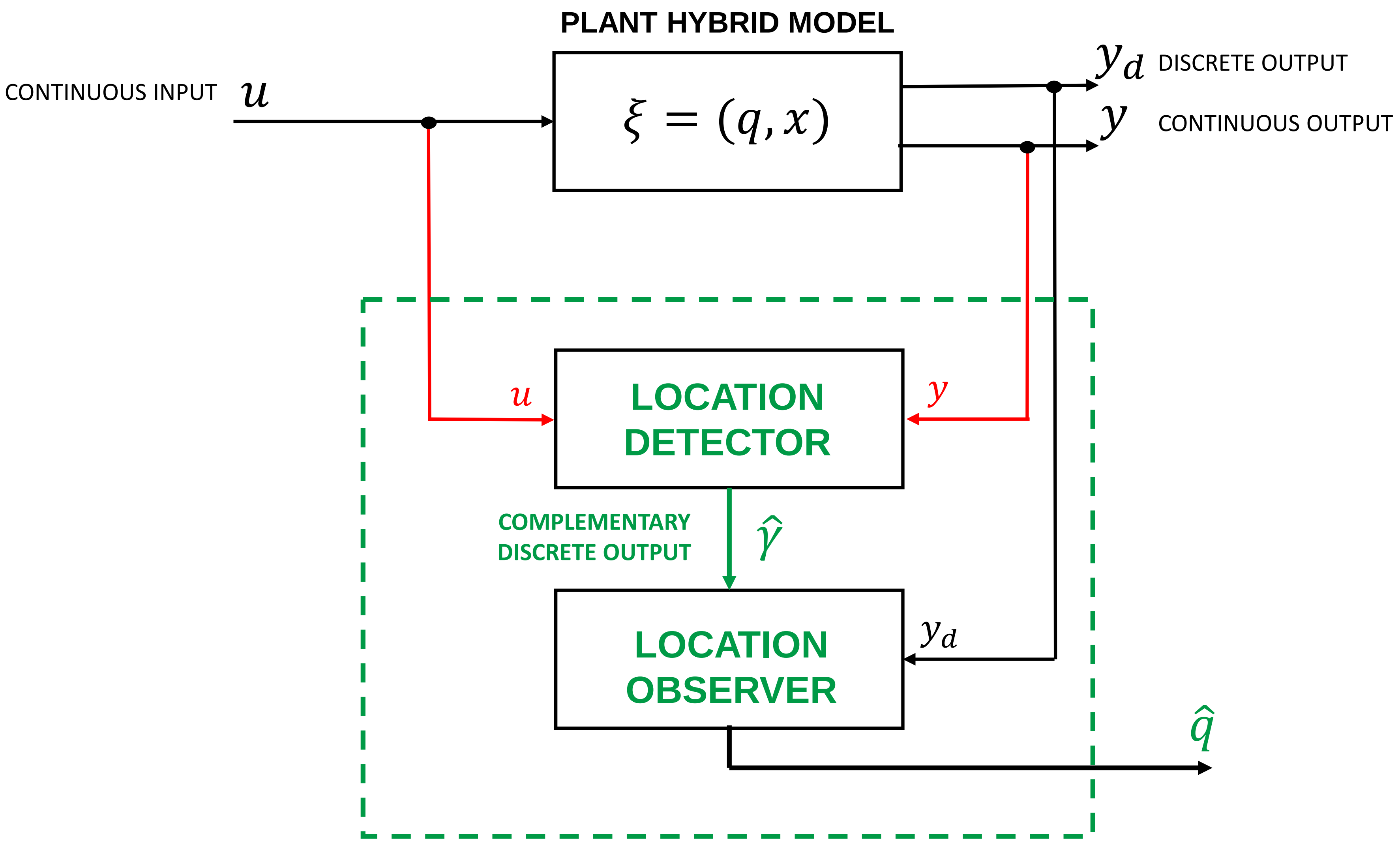
Example:



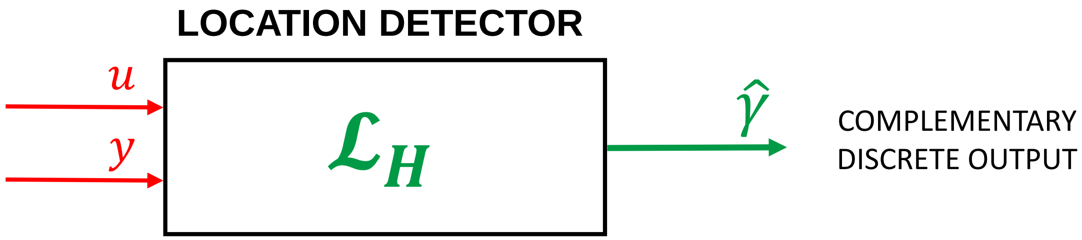
If the current output symbol is **b**, we can deduce that the current mode is either *i* or *j*. However, the modes *i* and *j* cannot be distinguished only on the basis of the discrete output information, although no state is silent.

Solution: Continuous inputs and outputs can be used to obtain some additional information that may be useful for the identification of the plant current location.

Location detector



Location detector design



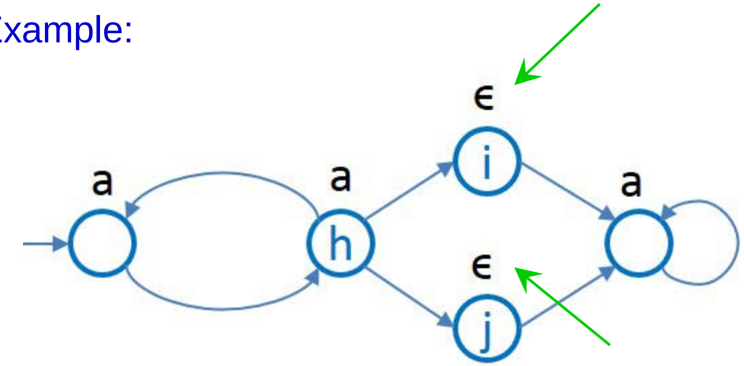
Theorem. The FSM M is **current location observable** if and only if for every persistent state $i \in Q_p$ of M:

1) ~~$h(i) \neq \varepsilon$;~~

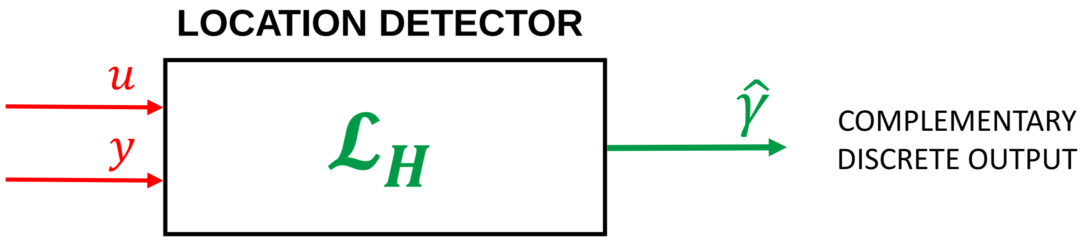
There exists persistent state of M having unobservable output.
$\mathcal{L}_H$ has to produce an output event γ

2) there exists a singleton state $\{i\}$ in the observer O_M and it is the only persistent state of O_M containing i.

Example:



Location detector design



Theorem. The FSM M is **current location observable** if and only if for every persistent state $i \in Q_p$ of M:

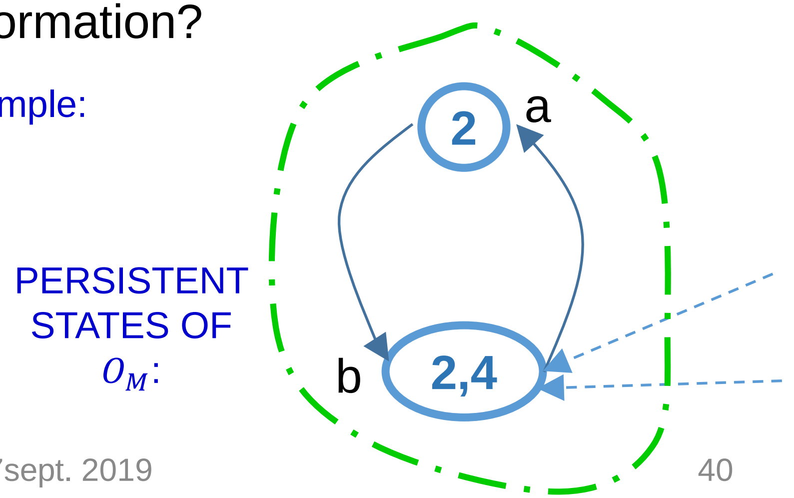
1) $h(i) \neq \varepsilon$;

2) ~~there exists a singleton state $\{i\}$ in the observer O_M and it is the only persistent state of O_M containing i.~~

There exist persistent states of M that are not *distinguishable* by using only discrete output information.

Question: Is it possible to *distinguish* those states by using continuous information?

Example:

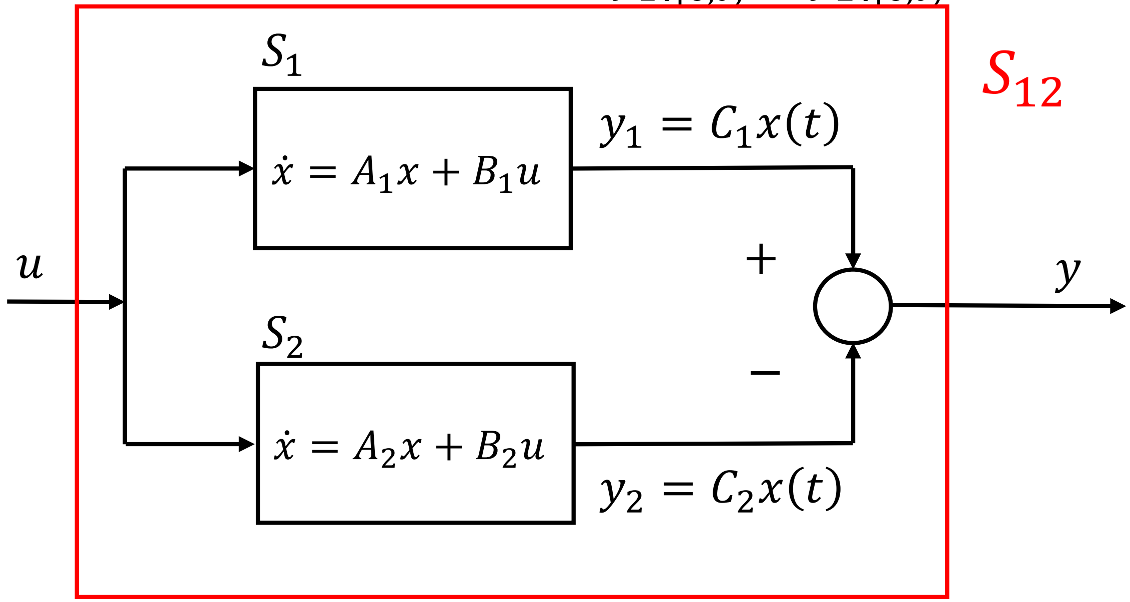


Input-generic distinguishability

Goal: Determine the current **discrete state** of a linear H-system by using only the **continuous output** information.

Definition: Two linear systems S_1 and S_2 are **input generic distinguishable** if, given an arbitrarily small $t > 0$, for all $(x_1(0)\ ,\ x_2(0))$ and for a generic input $u \in \mathcal{U}$,

$$y_1|_{[0,t)} \neq y_2|_{[0,t)}.$$



$A_i \in \mathbb{R}^{n\times n} \quad i = 1,2$

$B_i \in \mathbb{R}^{n\times m} \quad i = 1,2$

$C_i \in \mathbb{R}^{p\times n} \quad i = 1,2$

$$A_{12} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \qquad B_{12} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \qquad C_{12} = \begin{bmatrix} C_1 & -C_2 \end{bmatrix}$$

Sparse attacks

- Physical process modeled as a linear dynamic system:

$$x(t+1) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + e(t)$$

with $t \in \mathbb{N}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$, where $e_i(t) \neq 0$ (some sensors are attacked)

Sparse attacks *[Fawzi and Tabuada, 2014]*:

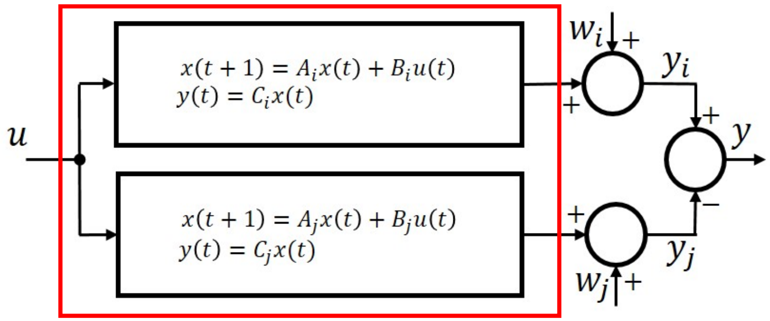
- e_i *(t)* **can be arbitrary** (no stochastic model, no boundedness,…)
- set of attacked sensors is **fixed**, but unknown
- the attacker has only access to a subset of sensors (whose cardinality is at most equal to σ)

$$[e(0)\,|\,e(1)\,|\,e(2)\,|\,e(3)] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ * & * & * & * \\ 0 & 0 & 0 & 0 \\ * & * & * & * \end{bmatrix}$$

Notation:

- $e(t) \in \mathbb{S}_\sigma^p \qquad \boldsymbol{\sigma} = \|e(t)\|_0 < \boldsymbol{p}$
- $e|_{[0,3]} \in \mathbb{CS}_\sigma^{4p}$

Secure distinguishability



$$x(t+1) = A_q x(t) + B_q u(t) \qquad q = i, j$$
$$y_q(t) = C_q x(t) + w_q(t)$$

$w_q(t) \in \mathbb{S}_\sigma^p$: sparse attack

$w_q(t)|_{[0,\tau-1]} \in \mathbb{CS}_s^{p\tau}$: collecting τ samples

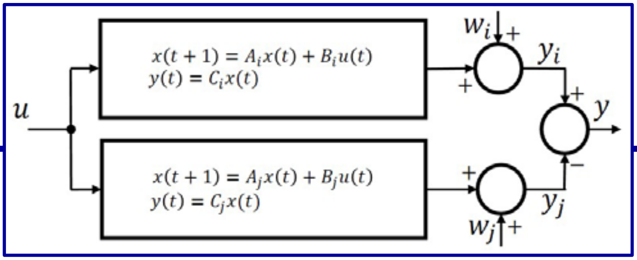
$$A_{ij} = \begin{bmatrix} A_i & 0 \\ 0 & A_j \end{bmatrix} \qquad B_{ij} = \begin{bmatrix} B_i \\ B_j \end{bmatrix} \qquad C_{ij} = \begin{bmatrix} C_i & -C_j \end{bmatrix}$$

Definition: S_i and S_j are **$\sigma 0$ −securely distinguishable** (w.r.t. generic inputs and for all σ −sparse attacks on sensors) if there exists $\tau \in \mathbb{N}$ s. t.

$$y_i|_{[0,\tau-1]} \neq y_j|_{[0,\tau-1]}$$

for any pair of intial states x_{0i} and x_{0j} , for any pair of σ −sparse attack vectors $w_i(t)|_{[0,\tau-1]} \in \mathbb{CS}_\sigma^{p\tau}$ and $w_j(t)|_{[0,\tau-1]} \in \mathbb{CS}_\sigma^{p\tau}$, and for any generic input sequence $u|_{[0,\tau-1)}$, and $u \in \mathcal{U}$.

Secure distinguishability

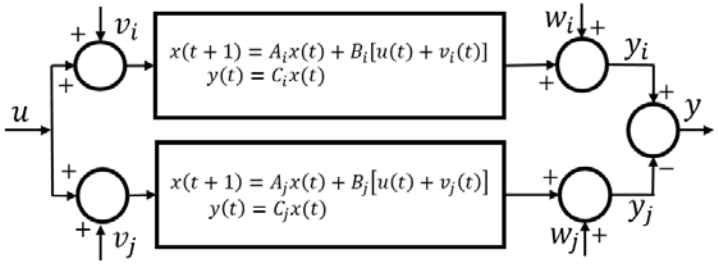


$$M_{ij} = \begin{bmatrix} C_{ij}B_{ij} & 0 & \dots & 0 \\ C_{ij}A_{ij}B_{ij} & C_{ij}B_{ij} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ C_{ij}A_{ij}^{2n-2}B_{ij} & C_{ij}A_{ij}^{2n-3}B_{ij} & \dots & C_{ij}B_{ij} \end{bmatrix} \qquad \mathcal{O}_{ij} = \begin{bmatrix} C_{ij} \\ C_{ij}A_{ij} \\ \vdots \\ C_{ij}A_{ij}^{2n-1} \end{bmatrix} = [\mathcal{O}_i \quad -\mathcal{O}_i]$$

Given the set $\Gamma \subset \{1, \dots, p\}, |\Gamma| \leq 2\sigma$, let $M_{ij,\Gamma}$ be the matrix obtained by the triples $(A_i, B_i, \bar{C}_{i,\Gamma})$ and $(A_j, B_j, \bar{C}_{j,\Gamma})$, where $\bar{C}_{i,\Gamma}$ is the matrix obtained from C_i by removing the rows contained in Γ.

Theorem: S_i and S_j are **$\sigma 0$ −securely distinguishable** if and only if for any set Γ with $\Gamma \subset \{1, \dots, p\}, |\Gamma| \leq 2\sigma$, the matrix $M_{ij,\Gamma} \neq \mathbf{0}$.

Secure distinguishability



$$x(t+1) = A_q x(t) + B_q[u(t) + v_q(t)] \quad q = i, j$$
$$y_q(t) = C_q x(t) + w_q(t)$$

$$w_q(t) \in \mathbb{S}_\sigma^p \text{ , } v_q(t) \in \mathbb{S}_\rho^m$$

$$A_{ij} = \begin{bmatrix} A_i & 0 \\ 0 & A_j \end{bmatrix} \qquad B_{ij} = \begin{bmatrix} B_i \\ B_j \end{bmatrix} \qquad C_{ij} = [C_i \quad -C_j]$$

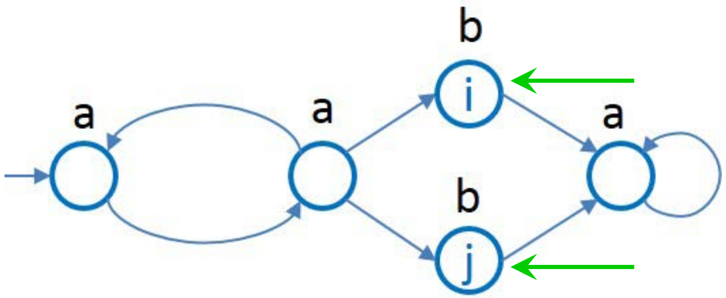
Definition: S_i and S_j are ***$\sigma\rho-$securely distinguishable*** (w.r.t. generic inputs, generic $\rho-$sparse attacks on actuators, and for all $\sigma-$sparse attacks on sensors) if there exists $\tau \in \mathbb{N}$ s. t.

$$y_i|_{[0,\tau-1]} \neq y_j|_{[0,\tau-1]}$$

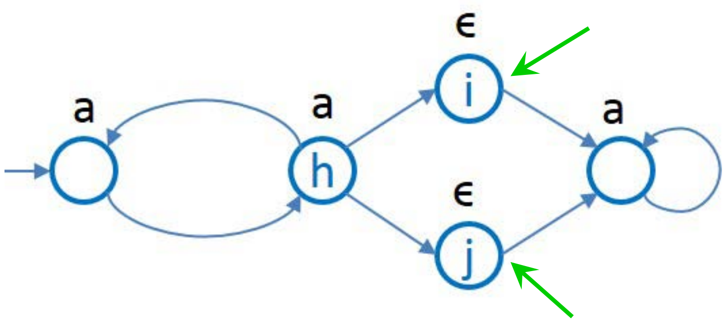
for any pair of intial states x_{0i} and x_{0j}, for any pair of $\sigma-$sparse attack vectors $w_i(t)|_{[0,\tau-1]} \in \mathbb{CS}_\sigma^{p\tau}$ and $w_j(t)|_{[0,\tau-1]} \in \mathbb{CS}_\sigma^{p\tau}$, and for any generic $(u, v_i, v_j) \in \mathcal{U} \times \mathbb{S}_\rho^m \times \mathbb{S}_\rho^m$.

Location detector design

Examples:



Distinguishability of (S_i, S_j) allows distinguishing mode i and mode j, despite the same output symbol



Distinguishability of (S_i, S_j), (S_h, S_i) and (S_h, S_j) ensures current location observability even though the persistent states i and j are silent

Current location observability
(mixed continuous and discrete information)

When only discrete output information is used, current location observability of H can be checked on the FSM M.

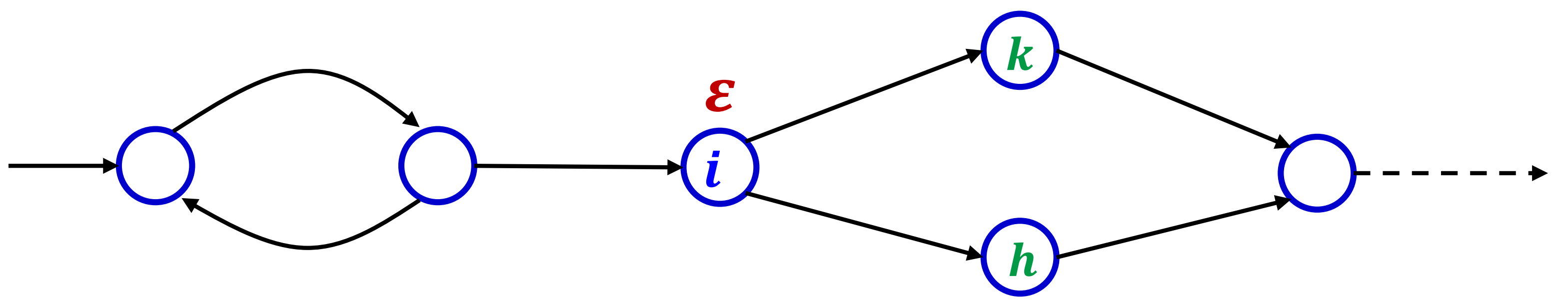
How to check current location observability of H when continuous output information is used?

H is transformed into an **«equivalent» hybrid system H'** with **purely discrete output** information and with **no silent states** by translating the continuous output information into discrete output signals.

Current location observability
(mixed continuous and discrete information)

1. If $i \in Q_p$ is a persistent state, then either it is **not silent** ($h(i) \neq \varepsilon$) or the pair of dynamical systems (S_i, S_j) is **distinguishable** for any other state j such that j belongs to *succ*(i).

Example:

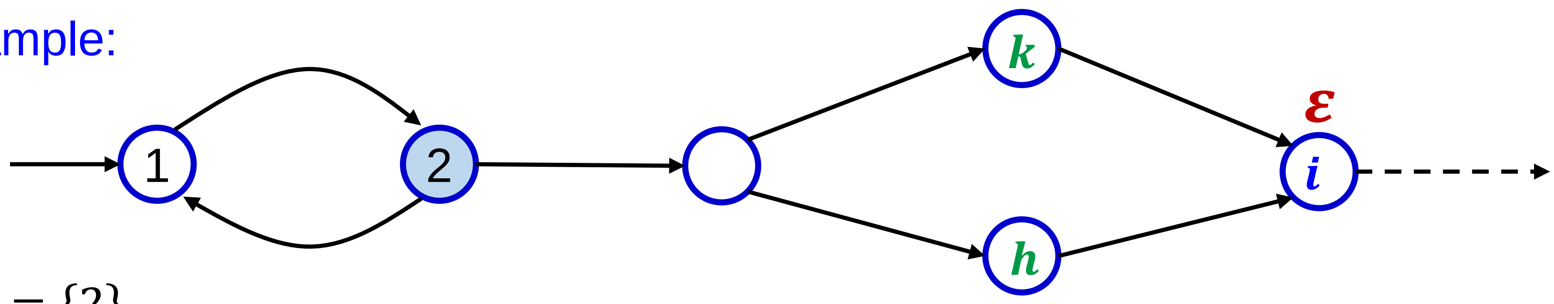


State $\boldsymbol{i}$ is a persistent state and it is silent, thus distinguishability of pairs $(\boldsymbol{S_i}, \boldsymbol{S_k})$ and $(\boldsymbol{S_i}, \boldsymbol{S_h})$ is necessary

Current location observability
(mixed continuous and discrete information)

2. If $i \in reach(Q_\infty) \backslash Q_0$, then either it is **not silent** $(h(i) \neq \varepsilon)$ or the pair of dynamical systems (S_j, S_i) is **distinguishable** for any other state j predecessor of i.

Example:

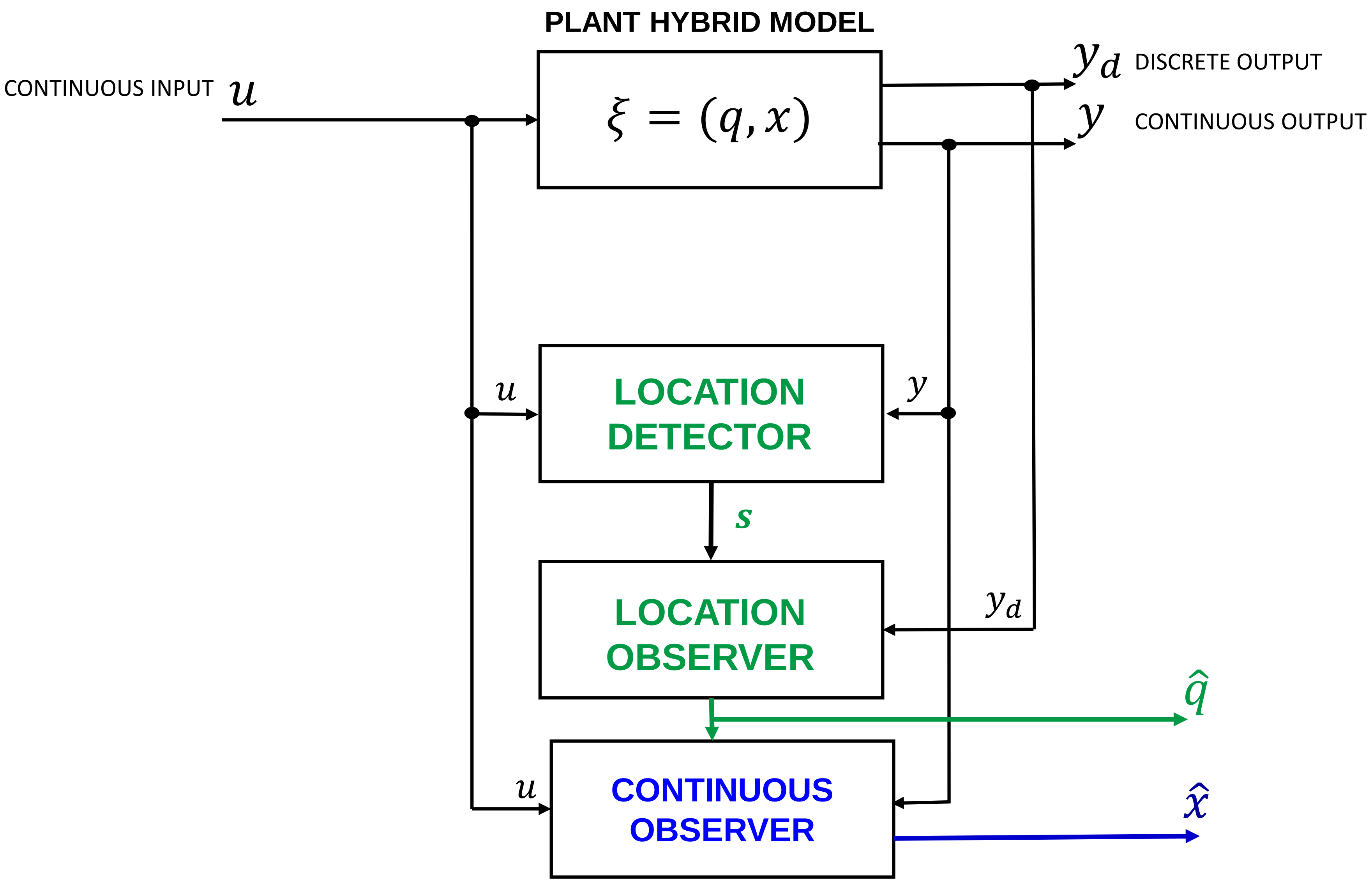


$Q_\infty = \{2\}$

State $\boldsymbol{i}$ is a persistent state and it is silent, thus distinguishability of pairs $(\boldsymbol{S_i}, \boldsymbol{S_k})$ and $(\boldsymbol{S_i}, \boldsymbol{S_h})$ is necessary

3. If step 1 and step 2 are possible, H is current location observable if H' (with purely discrete output and no silent states) is current location observable, and this can be checked on the FSM associated to H'.

Hybrid observer design



Diagnosability of M

$M = (Q, Q_0, Y, h, E)$ **Critical set:** $\Omega \subset Q$

Ω −diagnosability describes the possibility of inferring that **the state belongs to Ω**, on the basis of the output execution

For any infinite state execution $\sigma \in \mathcal{X}$ two cases are possible:

i. $\sigma(k) \notin \Omega, \forall\ k \in \mathbb{Z}$

ii. $\sigma(k) \in \Omega$, for some $k \in \mathbb{Z}$ (crossing event)

If (ii) holds, let k_σ be the minimum value of k such that $\sigma(k) \in \Omega$,otherwise $k_\sigma = \infty$

Parametrical Ω –Diagnosability

Definition: M is **parametrically *Ω* –diagnosable** if there exist $\tau \in \mathbb{Z}$, $\delta \in \mathbb{Z}$, and $\mathrm{T} \in \mathbb{Z} \cup \{\infty\}$ such that for any string $\sigma \in \mathcal{X}$ with **finite** k_σ, whenever $\sigma(k) \in \Omega$ and $k \in [max\{k_\sigma, (\tau+1)\}, k_\sigma + T]$, it follows that for any string $\hat{\sigma} \in \boldsymbol{y}^{-1}(y(\sigma|_{[1,k+\boldsymbol{\delta}]}))$, $\hat{\sigma}(l) \in \Omega$ for some $l \in [max\{1, (k-\gamma_1)\}, k+\gamma_2]$ and for some $\gamma_1, \gamma_2 \in \mathbb{Z}, \gamma_2 \leq \delta$.

- $\gamma = max\{\gamma_1, \gamma_2\}$: uncertainty radius in the reconstruction of the step at which the crossing event occurred
- $\delta \in \mathbb{Z}$: delay of the crossing event detection
- $\tau \in \mathbb{Z}$: initial time interval in which the crossing event is not required to be detected
- $\mathrm{T} \in \mathbb{Z} \cup \{\infty\}$: time interval in which the occurrence of the crossing event must be detected

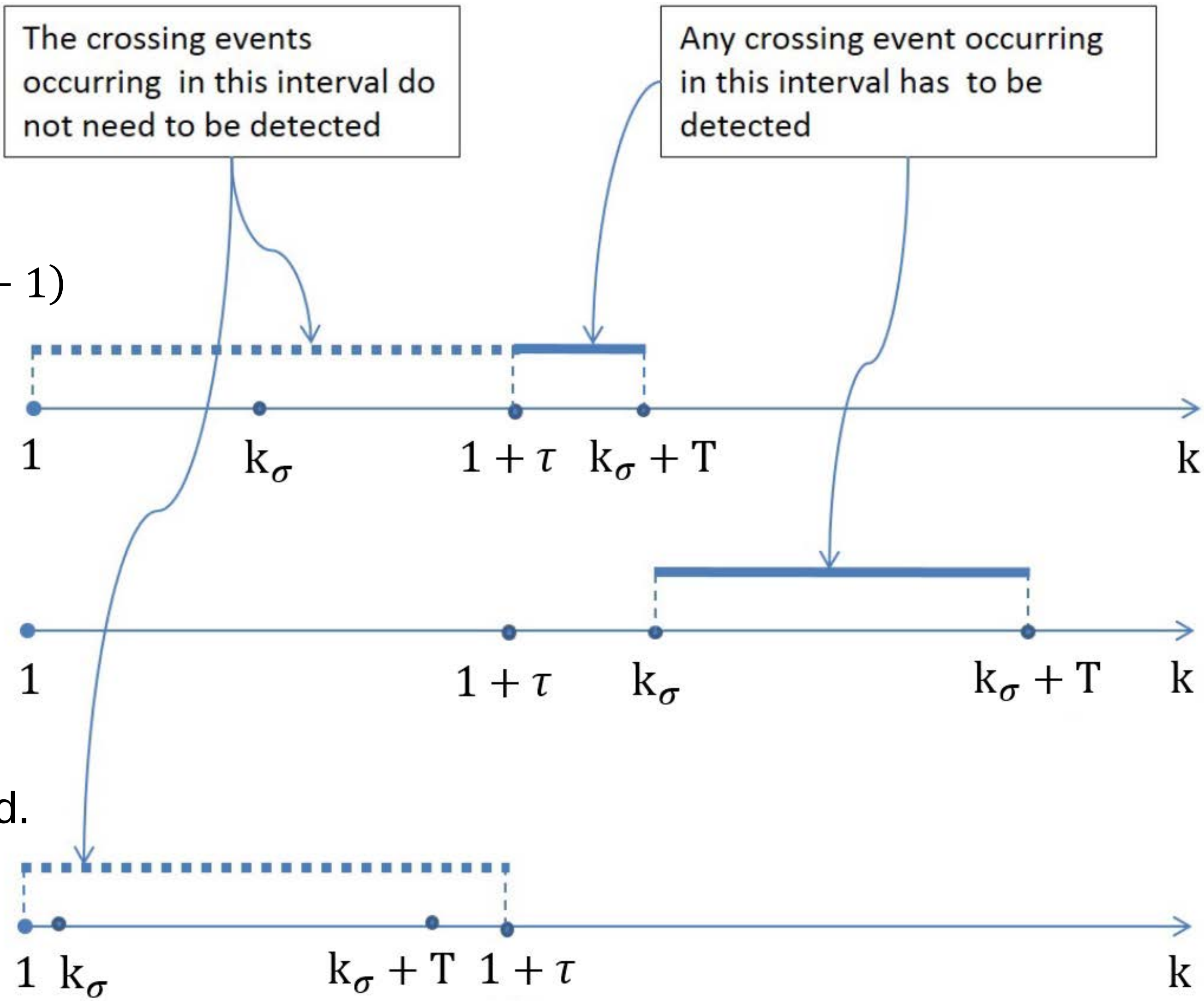
Parametrical Ω –Diagnosability

Parameters τ, T, δ, γ

1. $max\{k_\sigma, (\tau+1)\} = (\tau+1)$

2. $max\{k_\sigma, (\tau+1)\} = k_\sigma$

3. No detection is required.



Parametrical Ω –Diag: Special cases

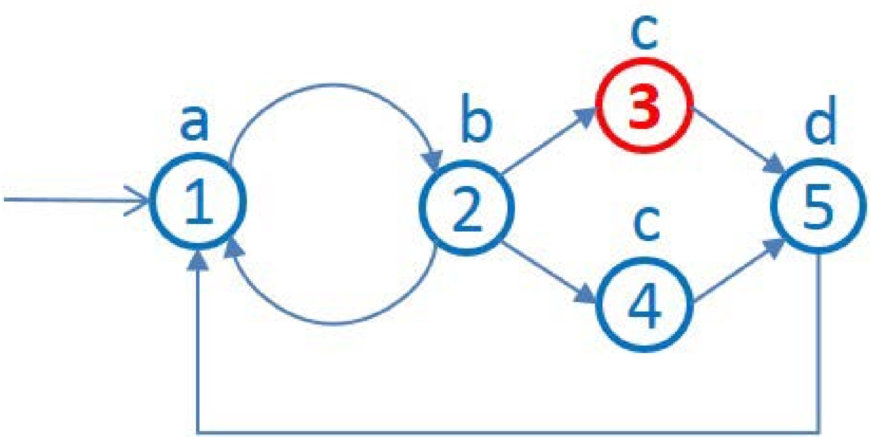
- **Ω –current state observability**
  - time interval within which the occurrence of the crossing event must be detected: $T = \infty$
  - initial time interval where the crossing event is not required to be detected: $\tau > 0$
  - delay of the crossing event detection: $\delta = 0$

- **critical Ω –observability**
  - time interval within which the occurrence of the crossing event must be detected: $T = \infty$
  - initial time interval where the crossing event is not required to be detected: $\tau = 0$
  - delay of the crossing event detection: $\delta = 0$

- **Ω** –initial state observability. $T = 0, \tau = 0, \delta \geq 0, \Omega \subset Q_0, \gamma_1 = \gamma_2 = 0$
  The crossing event is detected the first time it occurs, with delay $\delta \geq 0$

- **Ω –diagnosability**. $T = 0, \tau = 0$. If $\delta = 0$, **Ω –observability**

Parametrical Ω –Diagnosability



$\Omega = \{3\}$ M is not {3}-diag!

$\Omega = \{2\}$ M is {2}-diag!

- {3}-diagnosability: For any τ there exists an execution that crosses for the first time after the interval τ, and it is not possible to detect the set Ω nor immediately neither with a delay, or uncertainty

Checking Ω –Diagnosability

- The set-membership formalism and the derived algorithms are very simple and intuitive, and allow checking the diagnosability properties without constructing an observer.

- We can check diagnosability of a critical event, such as a faulty event, and at the same time compute
  - delay of the diagnosis with respect to the occurrence of the event,
  - the uncertainty about the time at which that event occurred,
  - the duration of a possible initial transient where the diagnosis is not possible or not required.

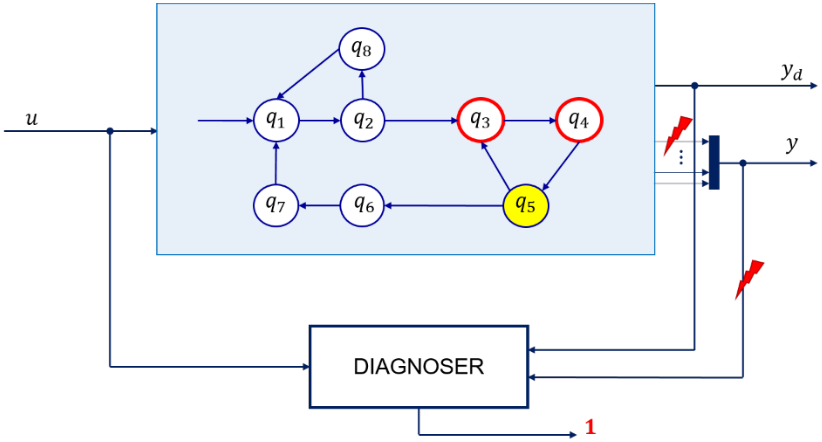
[De Santis, Di Benedetto, 2017]

Secure diagnosability of hybrid systems

Definition: A linear hybrid system is σ −securely Ω − diagnosable if there exists $T \in \mathbb{N}$ and a function $\mathcal{D}: \left(\mathcal{U} \times \mathcal{Y} \times \mathbb{S}_{\sigma}^{p}\right) \rightarrow \{0,1\}$, called diagnoser, s.t.

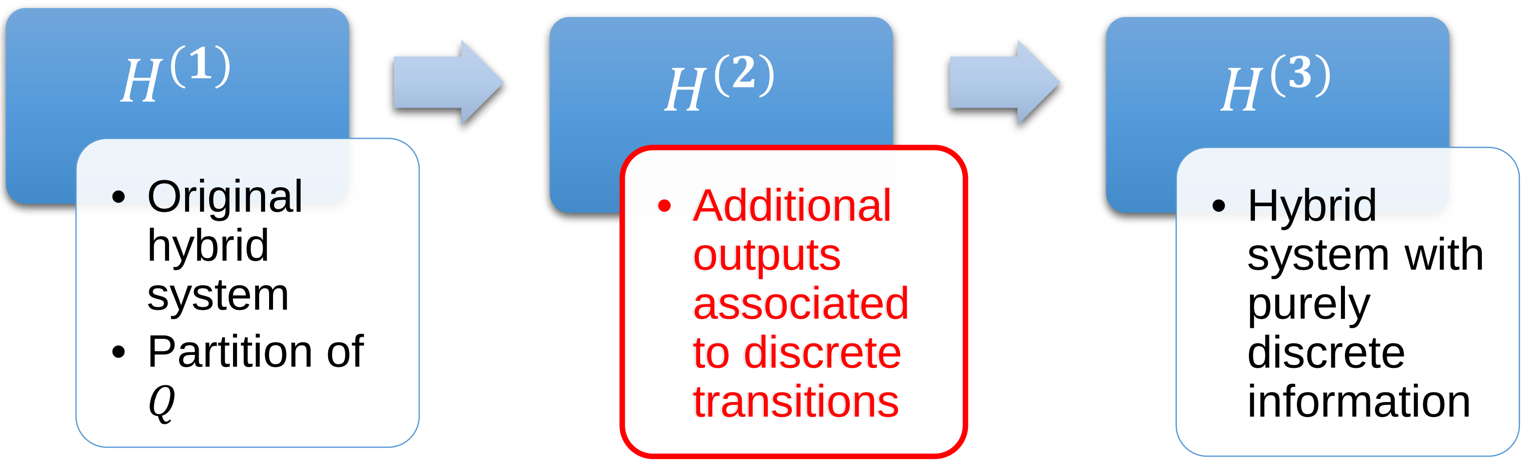
i. if $\xi(\hat{t}) \in \Omega \wedge (\hat{t} = 0 \vee (\xi(t) \notin \Omega,\ \forall\, t \in [0, \hat{t} - 1], \hat{t} > 0))$ then $\mathcal{D}\left(u|_{[0,\hat{t}+T-1]}, \eta|_{[0,\hat{t}+T]}\right) = 1$, with $\eta|_{[0,\hat{t}+T]} = (y_d|_{[0,\hat{t}+T]}, y|_{[0,\hat{t}+T]} + w|_{[0,\hat{t}+T]})$, for any generic input sequence $u|_{[0,\hat{t}+T-1]}$, with $u \in \mathcal{U}$, and for any attack sequence $w|_{[0,\hat{t}+T]} \in \mathbb{CS}_{\sigma}^{(\hat{t}+T)p}$

ii. if for any generic input sequence $u|_{[0,t-1\,]}$, with $u \in \mathcal{U}$, and for any attack sequence $w|_{[0,t\,]} \in \mathbb{CS}_{\sigma}^{tp}$, $\mathcal{D}\left(u|_{[0,t-1]}, \eta|_{[0,t\,]}\right) = 1$ and $\left(t = 0 \vee \left(\mathcal{D}\left(u|_{[0,t'-1\,]}, \eta|_{[0,t'\,]}\right) = 0, \forall\, t' \in [0, t-1], t > 0\right)\right)$ then $\xi(\hat{t}) \in \Omega$, for some $\hat{t} \in [\max\{0, t - T\}, t]$.



Abstracting procedure

If with $\Omega = \mathrm{Q_C} \times \mathbb{R}^n$, and discrete information is not sufficient to identify the discrete state, continuous output information is needed.



The abstracting procedure leads to a hybrid system with purely discrete information, that is equivalent to $H^{(1)}$ with respect to the secure diagnosability property.

Abstracting procedure

$H^{(1)}$

- Original hybrid system: partition of Q

$H^{(2)}$

- Additional outputs associated to discrete transitions

$H^{(3)}$

- Hybrid system with purely discrete information

Theorem: Let the linear hybrid system $H^{(1)}$ be given, with $\delta(q) \geq \delta_{min}$, $\Delta(q) \neq \infty$, $\forall q \in Q$. If $H^{(3)}$ is $\boldsymbol{Q_C}$ −diagnosable, then $H^{(1)}$ is σ −securely Ω −diagnosable with $\Omega = \mathbf{Q_C} \times \mathbb{R}^n$.

Approximate diagnosability

Let $F \subseteq X$ be a set of faulty states, $\rho \geq 0$ a desired accuracy, $\Omega = \mathrm{Q_C} \times F$

- If one is able to construct a symbolic metric system approximating a continuous or hybrid control system Σ (with an infinite number of states) in the sense of approximate simulation, we can check approximate diagnosability of Σ on the symbolic system

- Symbolic models approximating continuous or hybrid control systems are extensively investigated. Papers working with approximate simulation that fit the framework of our contribution:

[Pola et al., TAC-16; Pola et al., Autom-08]

[Zamani et al., TAC-12], for possibly unstable nonlinear systems

[Girard et al., TAC-10], for incrementally stable switched systems

[Pola & Di Benedetto, TAC-14], for piecewise affine systems

Outline

- Introduction
  - Cyber-Physical Systems (CPS)
  - Security for CPS
- Modeling CPS as hybrid systems
- Secure state estimation for hybrid systems
  - Observability and diagnosability
  - Secure mode distinguishability
  - Secure diagnosability
  - Approximate diagnosability
- Conclusions and future work

Conclusions and ongoing work

- Secure state estimation problem for hybrid systems
- Predictability for hybrid systems
- Malicious attacks on both continuous and discrete output information
- More general representation of attacks
- Application of the results

Some references

- F. Pasqualetti, F. Dorfler, and F. Bullo, “Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems”, IEEE Control Systems, vol. 35, no. 1, pp. 110–127, Feb. 2015.
- G. Fiore, A. Iovine, E. De Santis, M.D. Di Benedetto, “Secure state estimation for DC microgrids control”. IEEE Conference on Automation Science and Engineering (CASE) 2017
- A. Iovine, S. B. Siad, G. Damm, E. De Santis and M. D. Di Benedetto, "Nonlinear Control of a DC MicroGrid for the Integration of Photovoltaic Panels," in *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 2, pp. 524-535, April 2017.
- H. Fawzi, P. Tabuada and S. Diggavi, "Secure Estimation and Control for Cyber-Physical Systems Under Adversarial Attacks," in IEEE Transactions on Automatic Control, vol. 59, no. 6, pp. 1454-1467, June 2014.
- Y. Shoukry and P. Tabuada, “Event-triggered state observers for sparse sensor noise/attacks”, IEEE Transactions on Automatic Control, vol. 61, no. 8, pp. 2079– 2091, Aug. 2016.
- Q. Hu, Y. H. Chang, and C. J. Tomlin, “Secure estimation for Unmanned Aerial Vehicles against adversarial cyber-attacks”, 30th Congress of the International Council of the Aeronautical Sciences (ICAS), Sep. 2016.
- M. S. Chong, M. Wakaiki, and J. P. Hespanha, “Observability of linear systems under adversarial attacks”, in American Control Conference (ACC), 2015, Jul. 2015, pp. 2439–2444.
- M. Pajic, I. Lee, and G. J. Pappas, “Attack-resilient state estimation for noisy dynamical systems”, IEEE Transactions on Control of Network Systems, vol. 4, no. 1, pp. 82–92, Mar. 2017.
- Y. Shoukry, P. Nuzzo, A. Puggelli, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, “Secure state estimation for Cyber Physical Systems under sensor attacks: A satisfiability modulo theory approach”, IEEE Transactions on Automatic Control, vol. PP, no. 99, pp. 1–1, 2017.
- Y. Zacchia Lun, A. D'Innocenzo, F. Smarra, I. Malavolta, M. D. Di Benedetto: State of the Art of Cyber-Physical Systems Security: an Automatic Control perspective Journal of Systems and Software, Journal of Systems and Software, 2019 DOI:10.1016/j.jss.2018.12.006
- J. Zaytoon and S. Lafortune. Overview of fault diagnosis methods for discrete event systems. Annual Reviews in Control, 37(2):308 – 320, 2013; and references therein.

Some references

- M. Sayed-Mouchaweh. Discrete Event Systems: Diagnosis and Diagnosability. Springer Science & Business Media, 2014
- F. Lin. Diagnosability of discrete event systems and its applications. Discrete Event Dynamic Systems, 4(2):197–212, 1994.
- L. Ye and P. Dague. An optimized algorithm of general distributed diagnosability analysis for modular structures. IEEE Transactions on Automatic Control, 62(4):1768–1780, April 2017.
- S. Narasimhan and G. Biswas. Model-based diagnosis of hybrid systems. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 37(3):348–361, May 2007.
- M. Bayoudh and L. Travé-Massuyès. Diagnosability analysis of hybrid systems cast in a discrete-event framework. Discrete Event Dynamic Systems, 24(3):309–338, 2014.
- A. Grastien, L. Travé-Massuyès, and V. Puig. Solving diagnosability of hybrid systems via abstraction and discrete event techniques, 20th World Congress of the International Federation of Automatic Control (IFAC), 2017
- O. Diene, E. R. Silva, and M. V. Moreira. Analysis and verification of the diagnosability of hybrid systems. 53rd IEEE Conference on Decision and Control, pages 1–6, Dec 2014.
- D. Luenberger, An introduction to observers, IEEE Transactions on Automatic Control, Dec 1971, Volume: 16, Issue: 6, pp. 596- 602. [21]
- E. Sontag, On the Observability of Polynomial Systems, I: Finite-Time Problems, SIAM Journal on Control and Opt., Volume 17 Issue 1, pp. 139-151, 1979.
- P. Ramadge, Observability of discrete-event systems, CDC 1986.
- P. Caines et al., Current-state tree, CDC 1988.
- C.M. Ozveren, A.S. Willsky, Observability of discrete event dynamic systems, IEEE Trans. Automatic Control, 1990
- A. Bemporad, G. Ferrari-Trecate, M. Morari, Observability and controllability of piecewise affine and hybrid systems. IEEE Transactions on Automatic Control, 2000.
- R. Vidal, A. Chiuso, S. Soatto, and S. Sastry. Observability of linear hybrid systems. In A. Pnueli and O. Maler, editors, Hybrid Systems: Computation and Control, 2003.
- P. Collins and J.H. van Schuppen. Observability of piecewise-affine hybrid systems., Hybrid Systems: Computation and Control, 2004.

Some references

- E. De Santis and M. D. Di Benedetto. Observability and diagnosability of finite state systems: a unifying framework. Automatica, 81:115–122, 2017.
- E. De Santis, M.D. Di Benedetto, Observability of hybrid dynamical systems, Foundations and Trends in Systems and Control, vol. 3, n.4, pp. 363-540, 2016.
-
- G. Fiore, E. De Santis, M.D. Di Benedetto, Secure Diagnosability of Hybrid Dynamical Systems, Chapter 7 in "Diagnosability, Security and Safety of Hybrid Dynamic and Cyber-Physical Systems", Springer International Publishing GA2018, M. Sayed-Mouchaweh Ed., pp. 175-200, 2018.
- G. Fiore, E. De Santis, M.D. Di Benedetto, “Secure mode distinguishability for switching systems subject to sparse attacks”, 20th World Congress of the International Federation of Automatic Control (IFAC), Toulouse, France, 2017.
- A.Balluchi, L.Benvenuti, M.D.Di Benedetto, A.L.Sangiovanni-Vincentelli: Dynamical observers for hybrid systems: Theory and Application to an Automotive Control Problem, Automatica, vol. 49, n. 4, 2013, pp. 915 - 925.
- A.Balluchi, L.Benvenuti, M.D.Di Benedetto, A.L.Sangiovanni-Vincentelli: Design of Observers for Hybrid Systems. HSCC02, Claire J. Tomlin and Mark R. Greenstreet, Eds., vol. 2289 of Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg New York, 2002, pp. 76-89.
- G. Pola, E. De Santis, M.D. Di Benedetto: Approximate diagnosability of metric transition systems, 15th International Conference on Software Engineering and Formal Methods, September 4-8, 2017, Trento (Italy), A. Cimatti and M. Sirjani Eds. Lecture Notes in Computer Science, Springer Verlag, vol. no. 10469, pp. 269-283
- G. Pola, E. De Santis, M.D. Di Benedetto: Approximate Diagnosis of Metric Systems, Control Systems Letters, IEEE L-CSS, 2(1): 115-120, January 2018.

Thank you!