Static-Dynamic Analysis of Security Metrics for Cyber-Physical Systems

Sayan Mitra (PI), Geir Dullerud (co-PI), Swarat Chaudhuri (co-PI)
University of Illinois at Urbana Champaign

NSA SoS Quarterly meeting, University of Maryland
October 29th 2014

Project team

Sayan Mitra
UIUC, ECE
Hybrid & Distributed systems

Geir Dullerud
UIUC, MechE
Control theory, hybrid systems

Swarat Chaudhuri
Rice University, CS
Programming Languages, Formal methods

Zhenqi Huang
PhD student, ECE

Yu Wang
PhD student, MechE

Project goal

Hard problem addressed: (1) Predictive security metrics and (2) scalability and composability

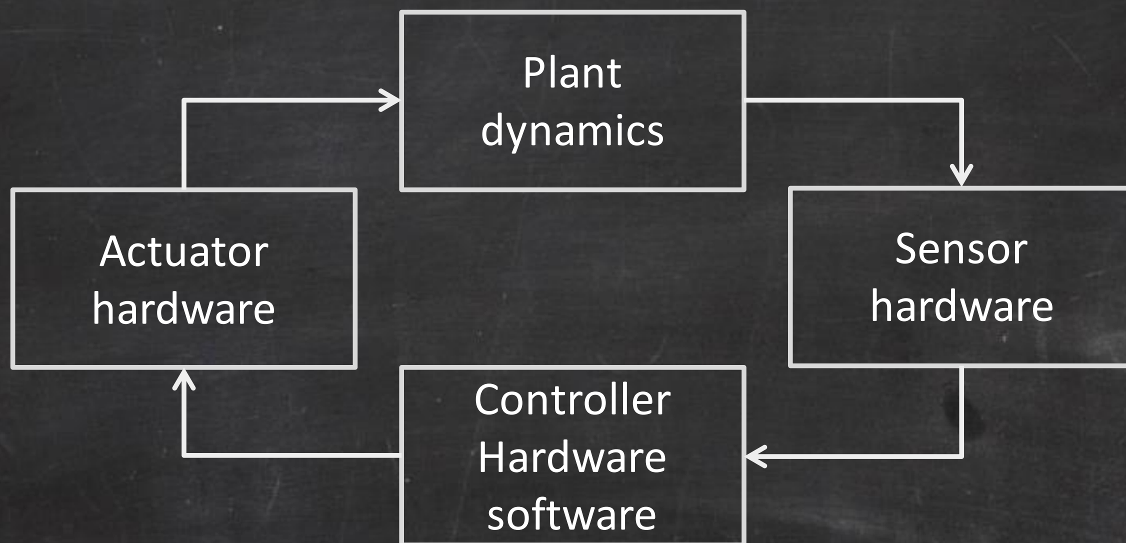
Title: Static-Dynamic Analysis of Security Metrics for Cyber-Physical Systems

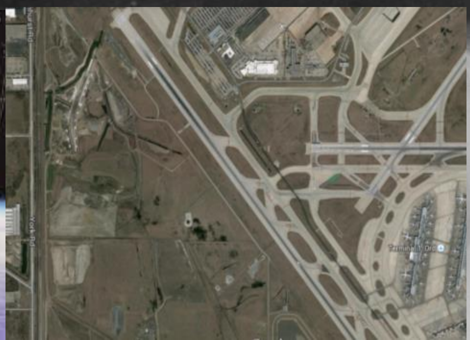
Goals:

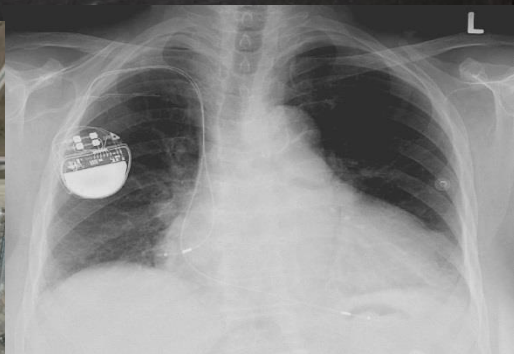
(a) Identify security metrics & adversary models

(b) develop theory, algorithms & tools for analyzing the metrics in the context of adversary models

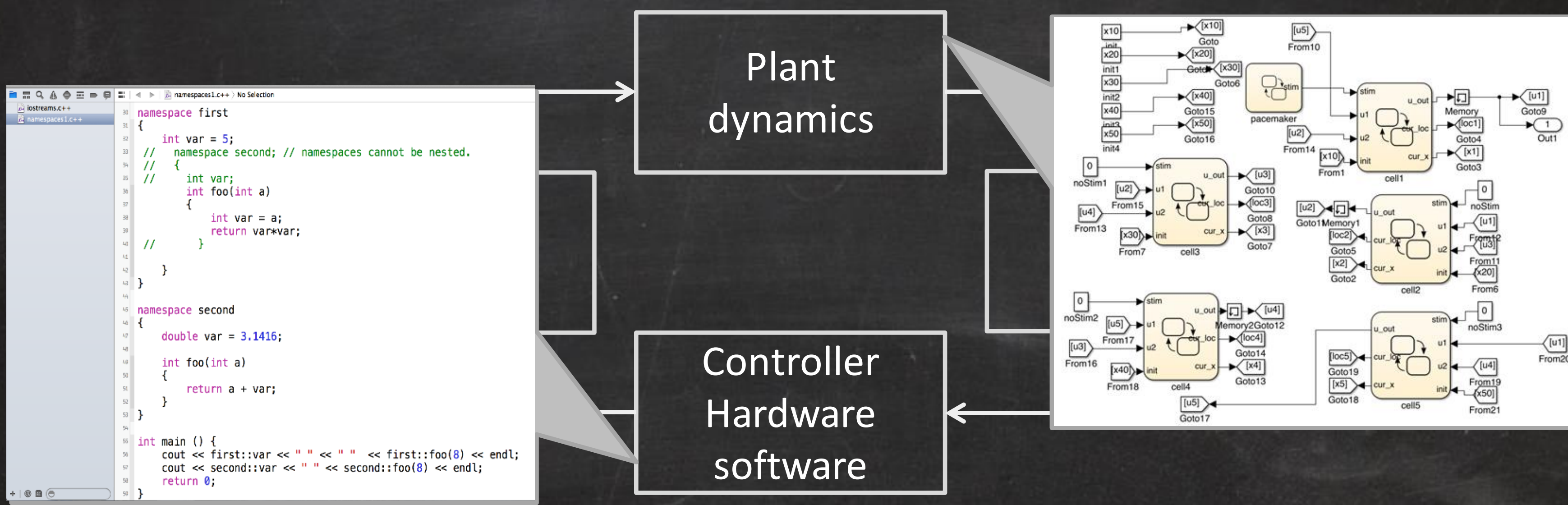
CPS & Security







Models, code, adversaries, & metrics



Hierarchy of modeling formalisms

Discrete transition systems (countable states) FSM, PDA, TMs

Communicating processes

IO automata, process algebras

Dynamical Systems

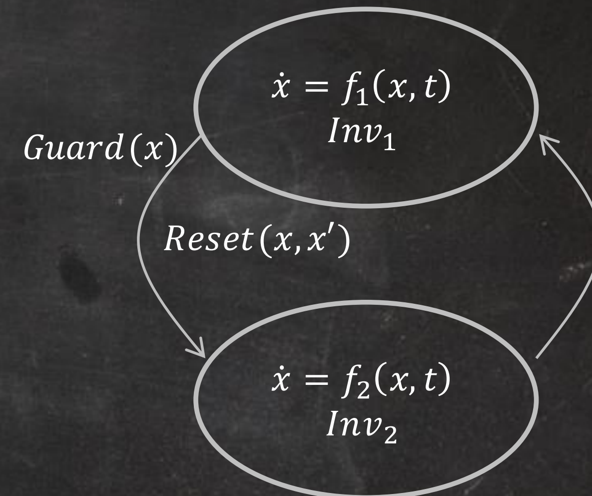
$\dot{x} = f(x, t, u)$

Nondeterministic transition systems

Switched Systems

$\dot{x} = f_{\sigma(t)}(x, t, u)$

Networked Hybrid Automata*



Metrics : Physical systems to CPS

Safety factor, Margin of safety, reserve capacity

↓

Availability, Stability envelope, safety margin, vulnerability level

Brooklyn bridge (1883)

Adversary models

access: actuator intrusion ∘ sensor jamming ∘ malicious programs

energy: opportunistic ∘ curious ∘ focused ∘ committed

Outline

- Two problems
  - Reachability for nonlinear hybrid systems
  - Cost of security in distributed control
- Two applications
  - Alerting protocol for parallel landing
  - Pacemaker with networked cardiac tissue
- Ongoing work
  - Synthesis with and for adversary

Part 1

STATIC-DYNAMIC ANALYSIS

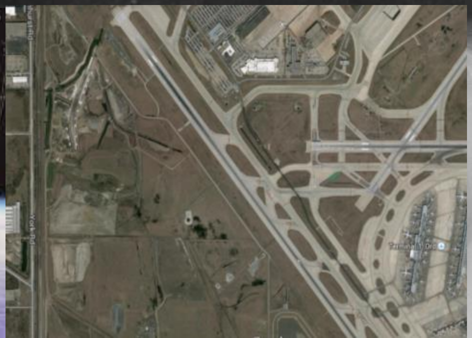
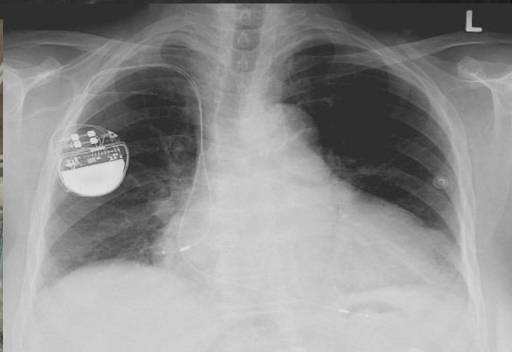
Basic analysis problem: verification

$\exists\ x_0 \in Init, u \in U, a \in A, t \in [0, T],$

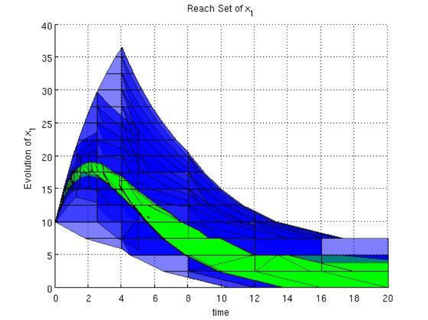
such that trajectory $\xi(x_0, a, u, t)$ violates requirements ?

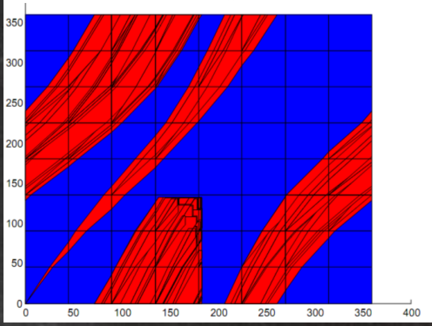
Yes (bug / security violation trace) / No (certificate)



Hybrid System Safety Verification





Early 90's: Exactly compute unbounded time reach set

Decidable for timed automata [Alur Dill 92]

Undecidable even for rectangular dynamics [Henzinger 95]

Late 90'-00': Approximate bounded time reach set

Hamilton-Jacobi-Bellman approach [Tomlin et al. 02]

Polytopes [Henzinger 97], ellipsoids [Kurzhanski] zonotopes [Girard 05], support functions [Frehse 08]

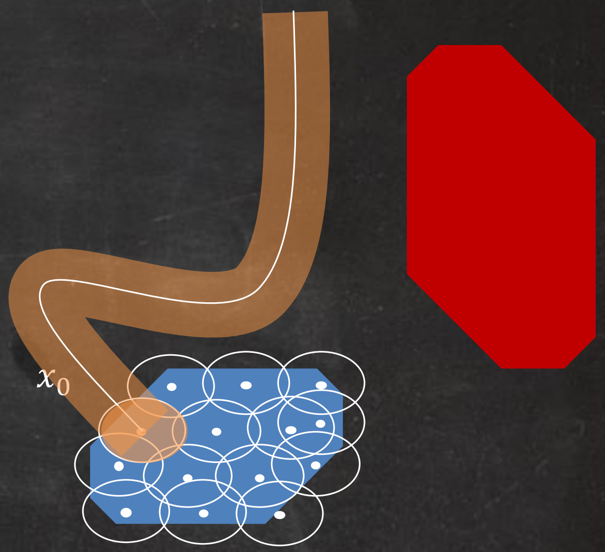
Predicate abstraction [Alur 03], CEGAR [Clarke 03] [Mitra 13]

Today: Scalability

Simulation-based methods [Julius 02] [Mitra 10-13][Donze 07]

A simple strategy

- Given start S and target T
- Compute finite cover of initial set
- Simulate from the center x_0 of each cover
- **Bloat** simulation so that bloated tube contains all trajectories from the cover
- Union = over-approximation of reach set
- Check intersection/containment with T
- Refine



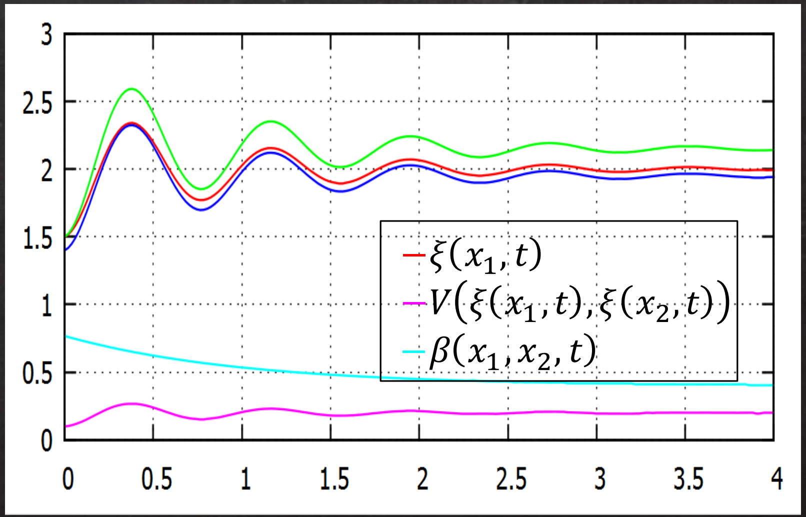
- How much to bloat?
- How to handle mode switches?

Discrepancy (Annotations in the spirit of loop invariants)

Definition. $\beta: \mathbb{R}^{2n} \times \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ defines a discrepancy of the system if for any two states x_1 and $x_2 \in X$, For any t,

1. $|\xi(x_1, t) - \xi(x_2, t)| \leq \beta(x_1, x_2, t)$ and
2. $\beta \rightarrow 0$ as $x_1 \rightarrow x_2$

```
x := 0
invariant x ≤ 10
until x ≥ 10
do
        x := x + 1
od
```



Lipschitz Constant

If L is a Lipschitz constant for *f(x,t)* then
$|\xi(x_1, t) - \xi(x_2, t)| \leq e^{Lt}|x_1 - x_2|$

Theorem [Lohmiller & Slotine '98]. A positive definite matrix M is a **contraction metric** if there is a constant $b_M > 0$ such that the Jacobian *J* of *f satisfies:*

$$J^T M + M\, J + b_M M \preccurlyeq 0.$$

If *M* is a contraction metric then $\exists k, \delta > 0$ such that $|\xi(x, t) -$

Hybrid Systems: Invariants

Track & propagate may and $must$ fragments of reachtube

$$\boldsymbol{tagRegion(R,P)} = \begin{cases} must & R \subseteq P \\ may & R \cap P \neq \emptyset \\ not & R \cap P = \emptyset \end{cases}$$

$\boldsymbol{invariantPrefix(\psi,S)} =$

$\langle R_0, tag_0, \ldots, R_m, tag_m \rangle$, such that either

$tag_i = must$ if all the $R_j' s$ before it are must

$tag_i = may$ if all the $R_j' s$ before it are at least may and at least one of them is not must

Sound & Relatively Complete

Theorem. (Soundness). If Algorithm returns safe or a counter-example, then A is indeed safe or has a counter-example.

Definition Given HA $A = \langle V, Loc, A, D, T \rangle$, an **$\epsilon$-perturbation** of A is a new HA A' that is identical except, $\Theta' = B_\epsilon(\Theta)$, $\forall\, \ell \in Loc, Inv' = B_\epsilon(Inv)$ (b) a ∈ A, $Guard_a = B_\epsilon(Guard_a)$.

A is **robustly meets U** iff $\exists \epsilon > 0$, such that A' meets U_ϵ upto time bound T, and transition bound N. **Robustly violates** iff $\exists\ \epsilon < 0$ such that A' is violates U_ϵ.

Theorem. (Relative Completeness) Algorithm always terminates whenever the A is either robustly meets or violates the requirement.

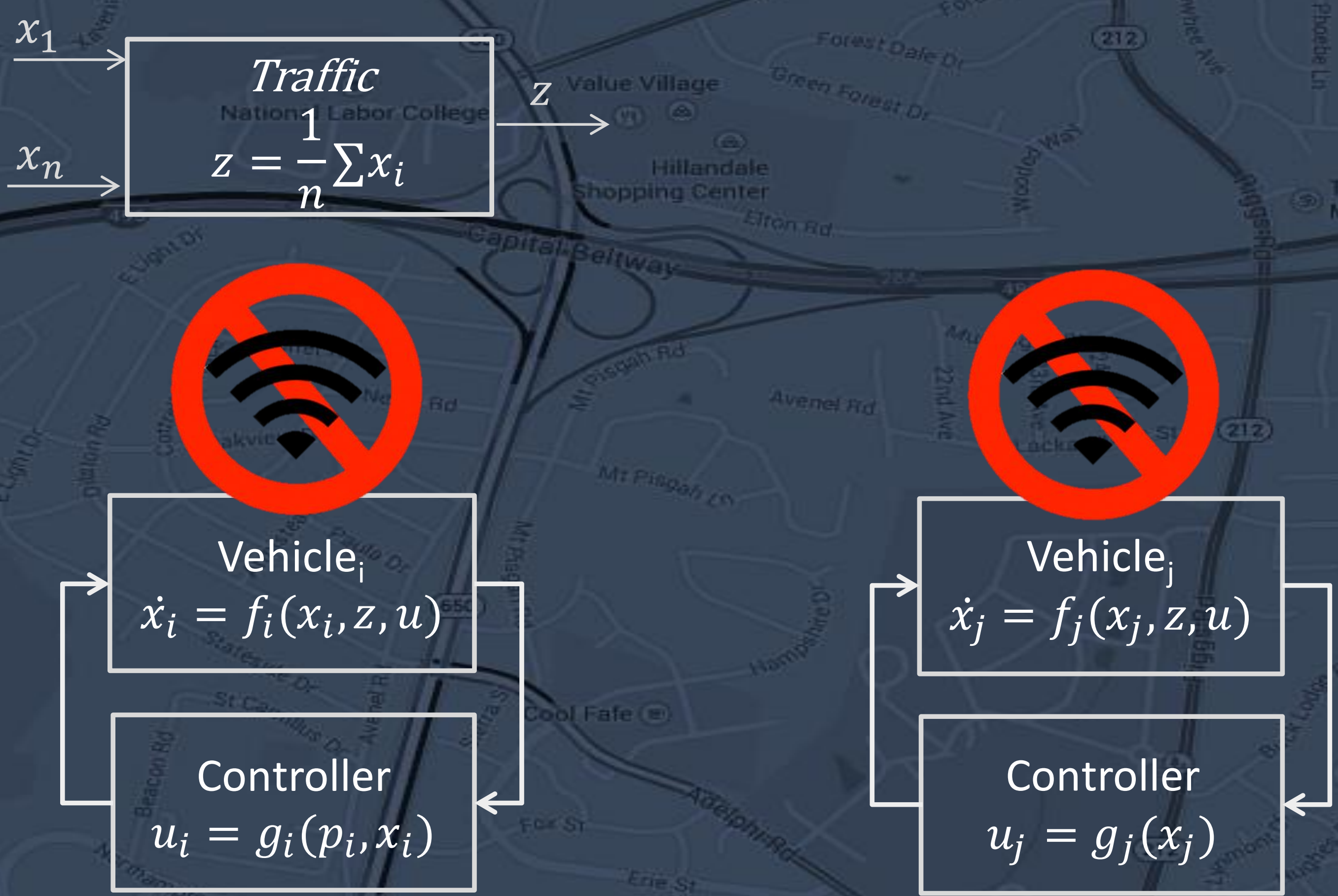
Part II

COST OF PRIVACY IN CONTROL

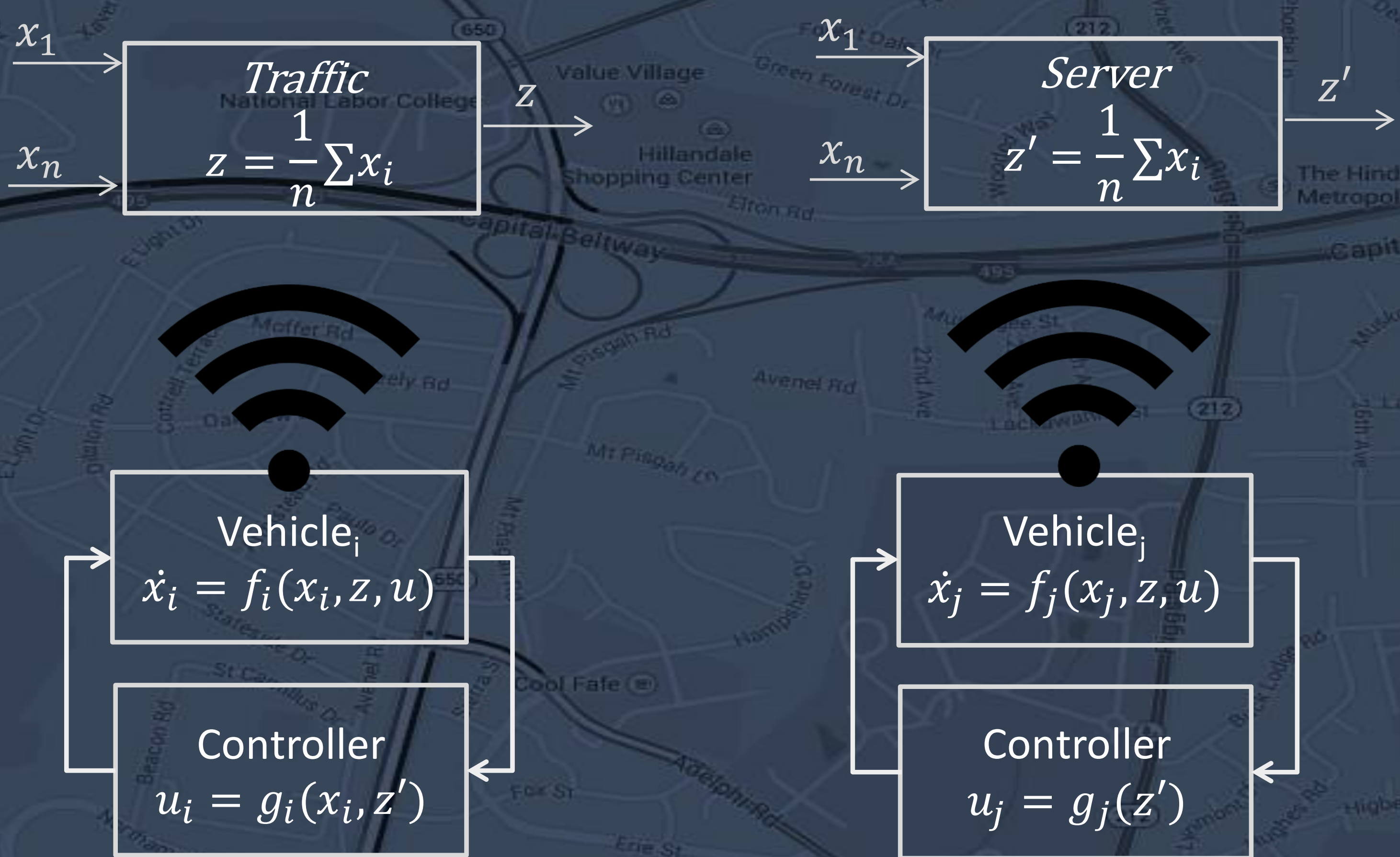
Huang ∘ Wang ∘ Mitra ∘ Dullerud

[CCS WPES 2012] [HiCons 2014] [CDC 2014] [ICDCN 2015]

Controlling Agents in a Shared Environment



Controlling Agents in a Shared Environment



Control while Protecting Sensitive Data

Obs: observation stream of the system bounded by time T, the broadcast positions.

Sensitive data: $g = \{g_1, \ldots, g_n\}$

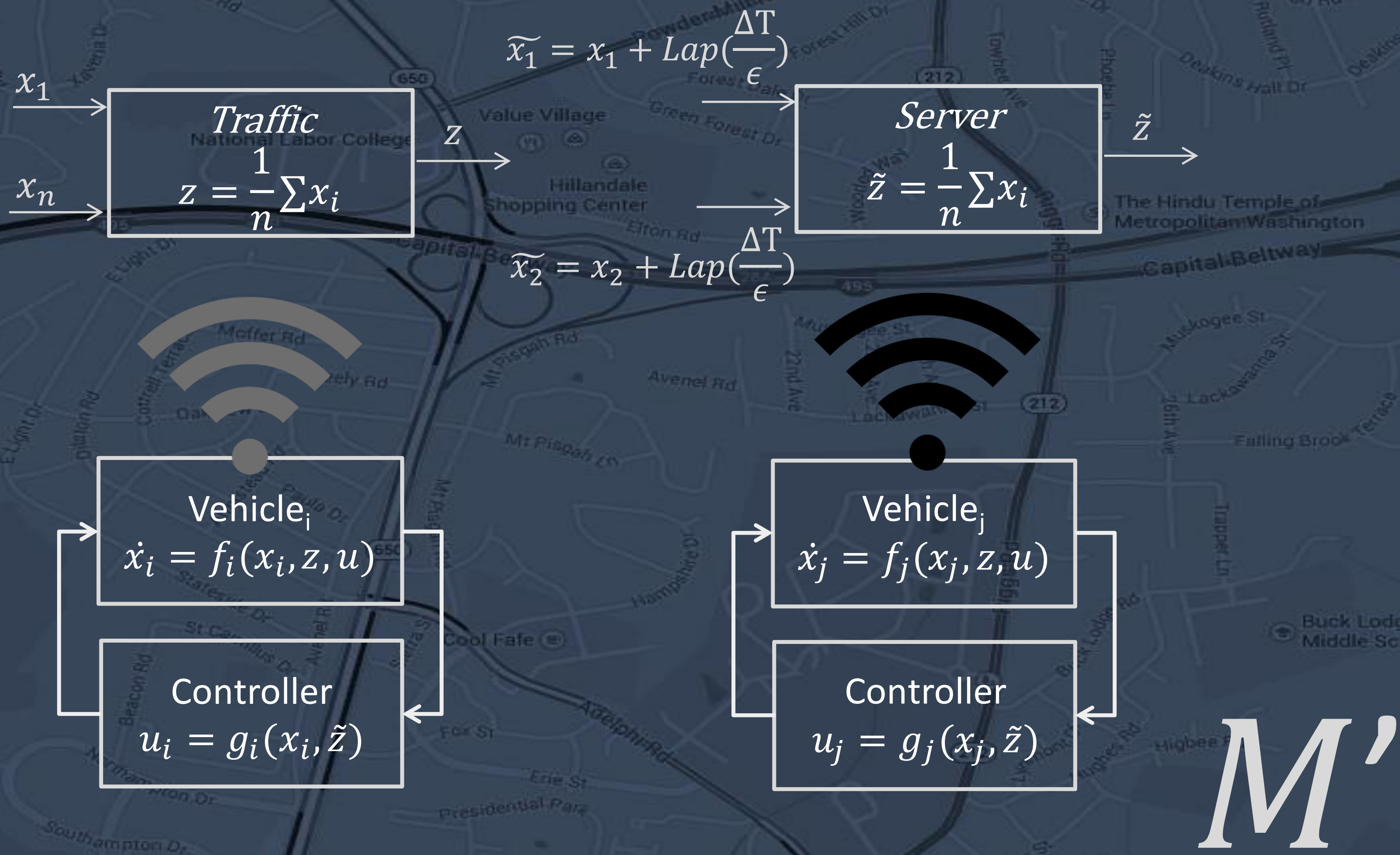
g and g' be two sequences of controllers that are identical except g_i and $g_i{}'$. The system is **differentially private** iff

$$\frac{P[g\ leads\ to\ Obs]}{P[g'\ leads\ to\ Obs]} \leq e^{|g_i - g_i'|}$$

Cost of privacy: $\sup_{g,i} E[Cost(g, M^*) - Cost(g', M')]$

What is the cost of Privacy in distributed control?

DP Control



x_1
x_n
Traffic
$z = \frac{1}{n}\sum x_i$
z
$\widetilde{x_1} = x_1 + Lap(\frac{\Delta T}{\epsilon})$
$\widetilde{x_2} = x_2 + Lap(\frac{\Delta T}{\epsilon})$
Server
$\tilde{z} = \frac{1}{n}\sum x_i$
$\tilde{z}$
Vehicle$_i$
$\dot{x}_i = f_i(x_i, z, u)$
Controller
$u_i = g_i(x_i, \tilde{z})$
Vehicle$_j$
$\dot{x}_j = f_j(x_j, z, u)$
Controller
$u_j = g_j(x_j, \tilde{z})$
M’

Control while Protecting Sensitive Data

Obs: observation stream of the system bounded by time T, the broadcast positions.

Privacy: g and g' be two sequences of controllers that are identical except g_i and $g_i{}'$. The system preserves differentially private iff

$$\frac{P[g\ leads\ to\ Obs]}{P[g'leads\ to\ Obs]} \leq e^{|g_i - g_i'|}$$

Cost of privacy: $\sup_{g} E[Cost(g, M) - Cost(g', M)]$

Theorem. COP = $O(\frac{T^3}{N^2\epsilon^2})$ for stable linear systems [HiCons 2014]

Cost reasonable for short-lived agents and large number of agents

Adversary estimates the initial system state from observations. $\tilde{X}(t) = E[X(0) \mid Z(0), Z(1), \ldots, Z(t)]$. Accuracy at time t ∈ N is measured by $H(\tilde{X}(t))$. Lower-bound on H for any ϵ-DP one shot query [CDC 2014].

TWO APPLICATIONS OF STATIC-DYNAMIC ANALYSIS

Duggirala ∘ Wang ∘ Mitra ∘ Munoz ∘ Viswanathan (FM 2014)

Huang ∘ Fan ∘ Meracre ∘ Mitra ∘ Kiwatkowska (CAV 2014)

SAPA-ALAS Parallel Landing Protocol

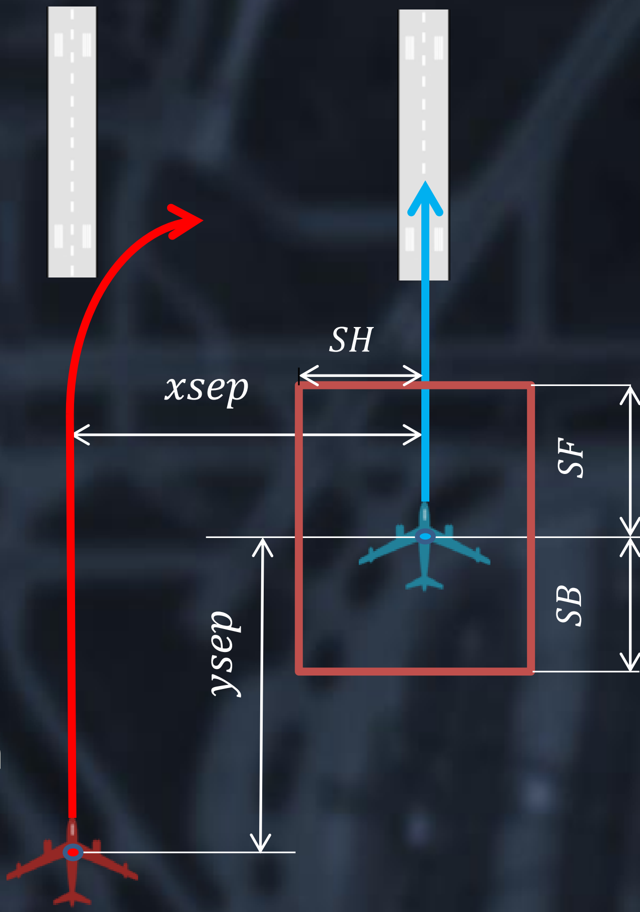
Ownship and *Intruder* approaching parallel runways with small separation

ALAS (at ownship) protocol is supposed to raise an alarm if within T time units the *Intruder* can violate safe separation based on 3 different projections

Verify Alert$\preccurlyeq_b$Unsafe for different scenarios

Scenario 1. With xsep [.11,.12] Nm ysep [.1,.21] Nm, $\phi = 30^o$ $\phi_{max} = 45^o$ vy_o= 136 Nmph, vy_i = 155 Nmph

$Alert \prec_b Unsafe$ is **satisfied** by Reachtube ψ if $\forall I_2 \in Must(Unsafe) \cup May(Unsafe)$ there exists $I_1 \in Must(Alert)$ such that $I_1 < I_2 - b$



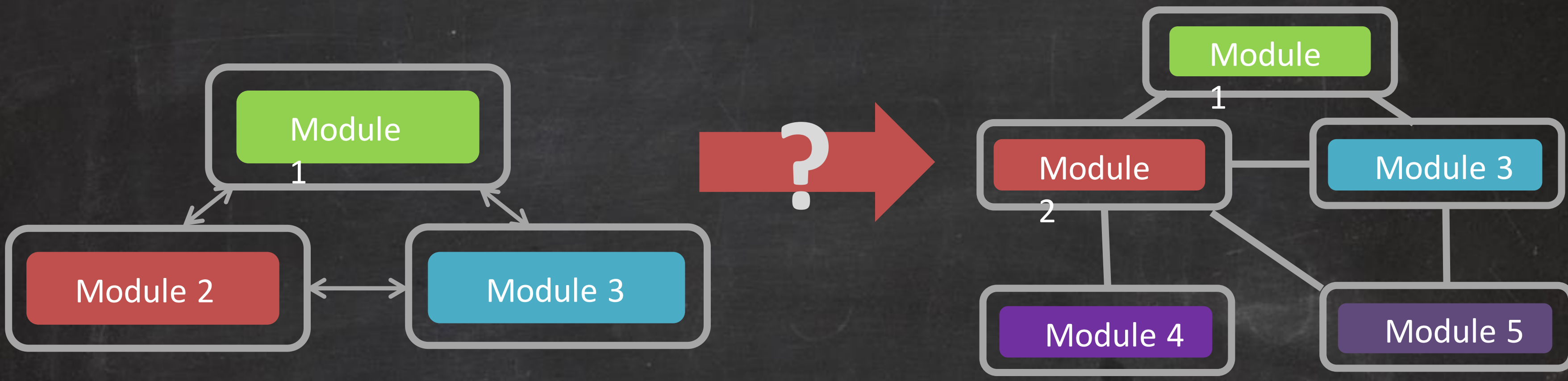
Real-time Alerting Protocol

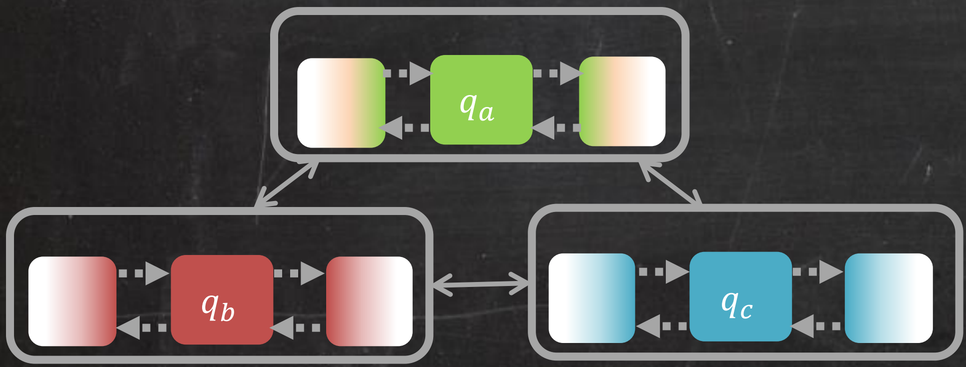
Scenario	Alert $\preccurlyeq_4$ Unsafe	Running time (mins:sec)	Alert $\preccurlyeq_?$ Unsafe
6	False	3:27	2.16
7	True	1:13	–
8	True	2:21	–
6.1	False	7:18	1.54
7.1	True	2:34	–
8.1	True	4:55	–
9	False	2:18	1.8
10	False	3:04	2.4
9.1	False	4:30	1.8
10.1	False	6:11	2.4

Sound & robustly completeness

C2E2 verifies interesting scenarios in reasonable time; shows that false alarms are possible; found scenarios where alarm may be missed

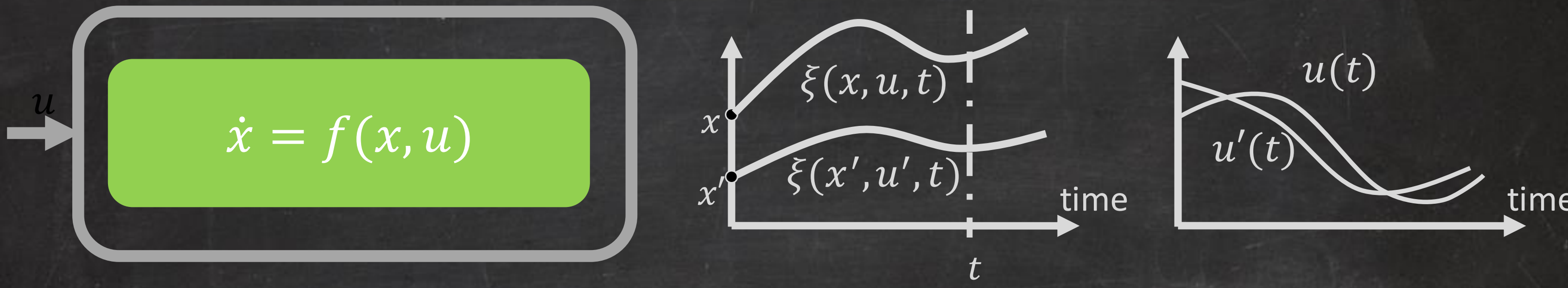
Scalability through Compositionality





$$\begin{aligned}\dot{x}_1 &= f_a(x_1, x_2, x_3)\\ \dot{x}_2 &= f_b(x_2, x_1, x_3)\\ \dot{x}_3 &= f_c(x_3, x_1, x_2)\end{aligned} \times L^N$$

Input-to-State (IS) Discrepancy



Definition. **IS discrepancy** is defined by β and γ such that for any initial states x, x' and any inputs u, u',

$$|\xi(x,u,t) - \xi(x',u',t)| \leq \beta(x,x',t) + \int_0^t \gamma(|u(s) - u'(s)|)ds$$

$\beta \to 0$ as $x \to x'$, and $\gamma \to 0$ as $u \to u'$

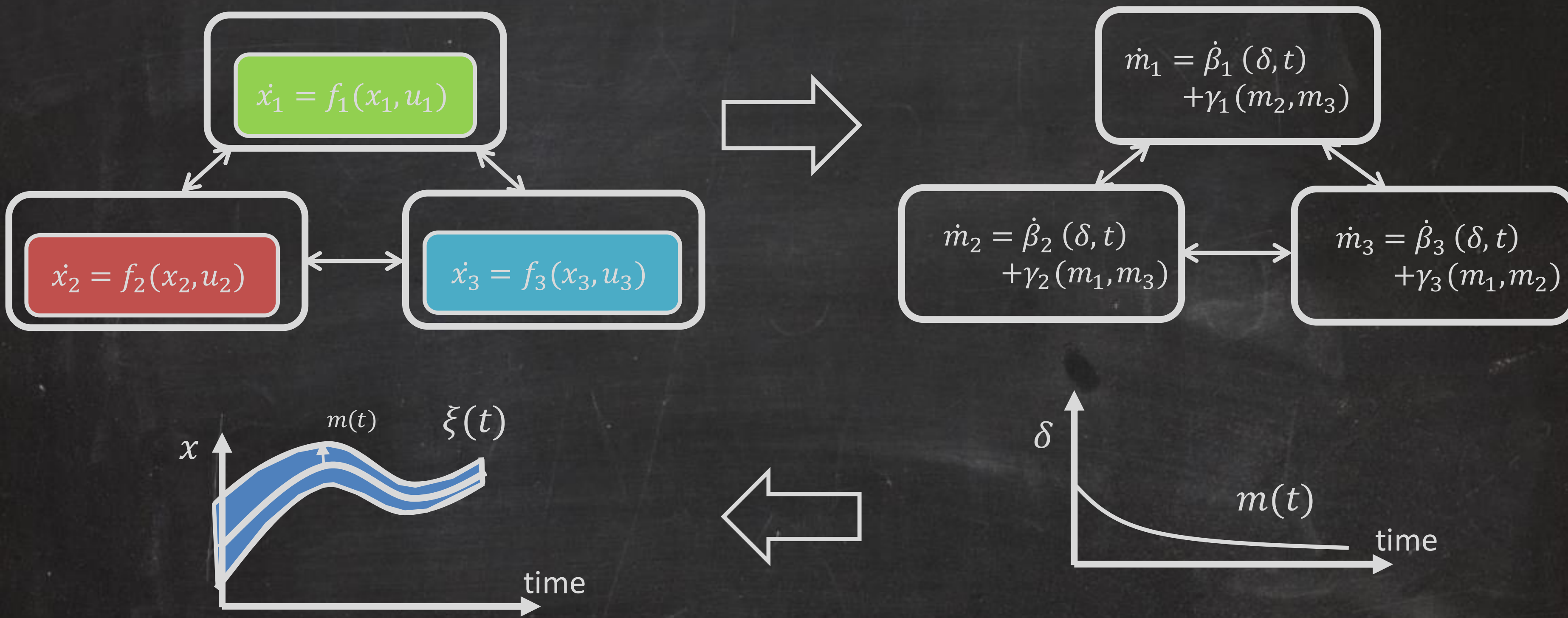
Reduced System $M(\delta_1, \delta_2, V_1, V_2)$

$$\dot{x} = f_M(x)$$

$$x = \langle m_1, m_2, clk \rangle$$

$$\begin{bmatrix} \dot{m}_1 \\ m_2 \\ clk \end{bmatrix} = f_M(x) = \begin{bmatrix} \dot{\beta}_1(\delta_1, clk) + \gamma_1(m_2) \\ \dot{\beta}_2(\delta_2, clk) + \gamma_2(m_1) \\ 1 \end{bmatrix}$$

Bloating with Reduced Model



The bloated tube contains all trajectories start from the δ-ball of x.

The over-approximation can be computed arbitrarily precise.

Reduced M gives effective Discrepancy of A

Theorem. For any $\delta = \langle \delta_1, \delta_2 \rangle, V = \langle V_1, V_2 \rangle$ and T
$Reach_A(B_\delta(x), T) \subseteq \bigcup_{t \leq T} B^V_{\mu(t)}(\xi(x,t))$

Theorem. For any $\epsilon > 0$ there exists $\delta = \langle \delta_1, \delta_2 \rangle$ such that
$\bigcup_{t \leq T} B^V_{\mu(t)}(\xi(x,t)) \subseteq B_\epsilon(Reach_A(B_\delta(x), T)$

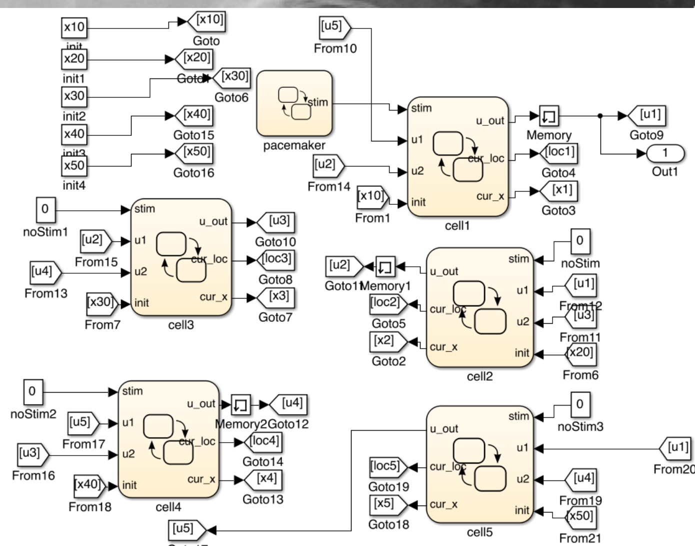
Here $\mu(t)$ is the solution of $M(\delta_1, \delta_2, V_1, V_2)$.

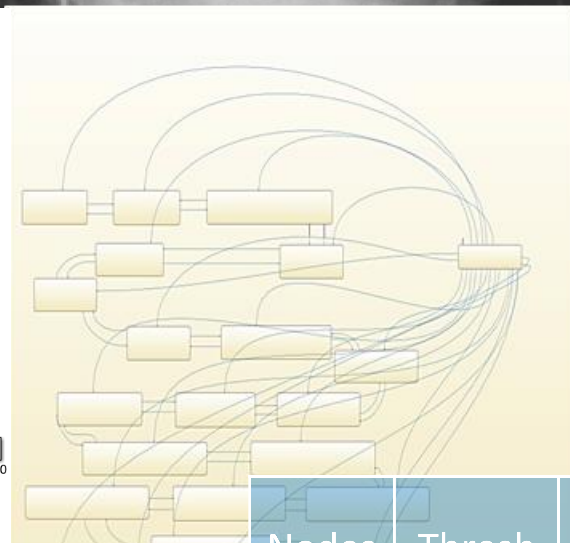
Huang et al. HSCC 2014, CAV 2014

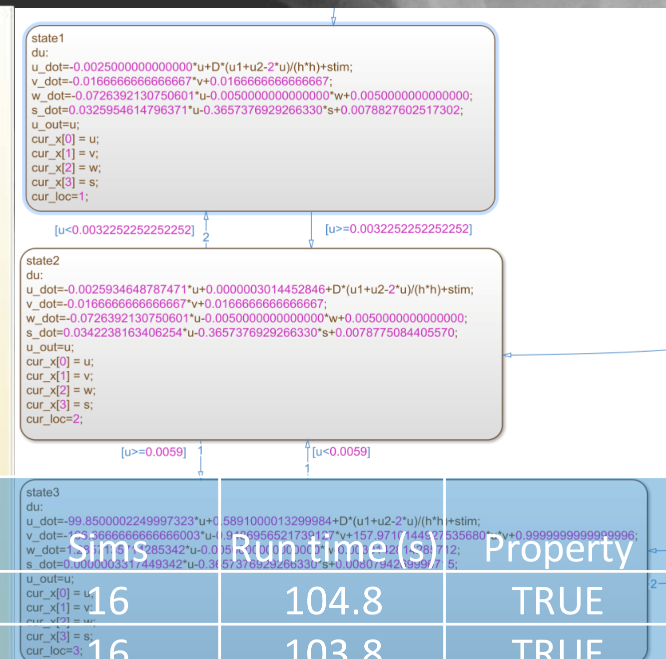
Pacemaker + Cardiac Network

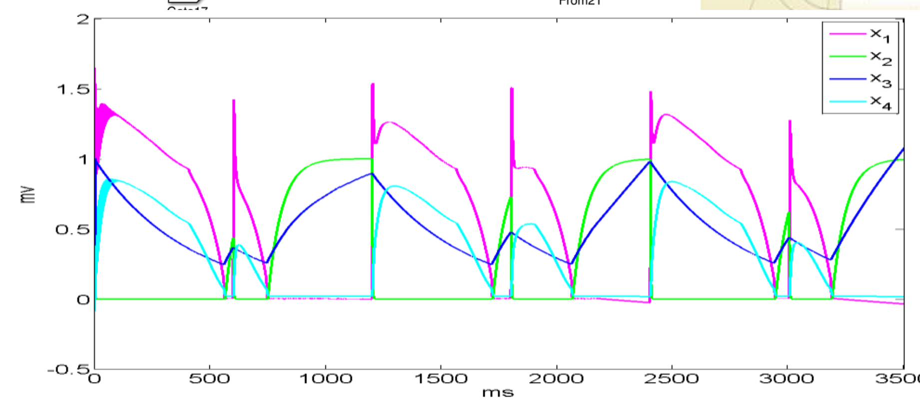
Action potential remains in specific range

No alternation of action potentials







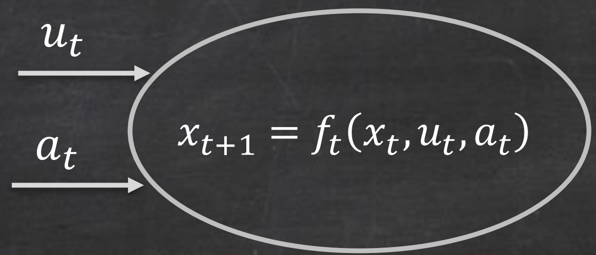


Nodes	Thresh	Sims	Run time (s)	Property
3	2	16	104.8	TRUE
3	1.65	16	103.8	TRUE
5	2	3	208	TRUE
5	1.65	5	281.6	TRUE
5	1.5	NA	63.4	FALSE
8	2	3	240.1	TRUE
8	1.65	73	2376.5	TRUE

PART IV

ONGOING WORK

Adversarial synthesis problem



Given system A, $\exists\, u \in Ctr, \forall\, x_0 \in Init, a \in Adv$:

$$\left.\begin{array}{l} \forall t\ \ \xi(x_0, u, a, t) \in Safe \\ \quad \xi(x_0, u, a, T) \in Goal \end{array}\right\}$$ requirements are met ?

Adv: $\sum |a_i|^2 \leq b$: intrusion budget constraints

Ctr: $\sum c_i u_i \leq k$: actuation constraints

Decomposition with Leverage

$Reach(x_0, u, Adv, t) = Reach(x_0, u, O, t) \oplus L(x_0, u, t)$ **---Leverage**

For each $t \leq H$, compute $Safe_t \oplus L(t) = Safe$ & $Goal_t \oplus L(t) = Goal$

Check $\exists u \in Ctrl : \forall t, x_0 \in Init,\ Reach(Init, u, 0, t) \subseteq Safe_t$?

For linear dynamics and L2-budget $\mathrm{L}(x_0, u, t)$ can be computed exactly

We can find b_{crit} that makes control impossible

Classify initial states based on vulnerability

Summary

- Static-Dynamic Analysis = sound and relatively complete algorithm for analysis of nonlinear – nondeterministic models
  - Tool support (C2E2, try it: http://publish.illinois.edu/c2e2-tool/)
  - Compositional analysis
- Symbolic simulation of adversary-free system + over-approximation of leverage
  - Synthesize controllers and attack strategies
  - Measure vulnerability of states w.r.t. attacks