

Computationally-Aware Cyber-Physical Systems (CACPS)

Ricardo Sanfelice (UCSC) and Jonathan Sprinkle (UofA)



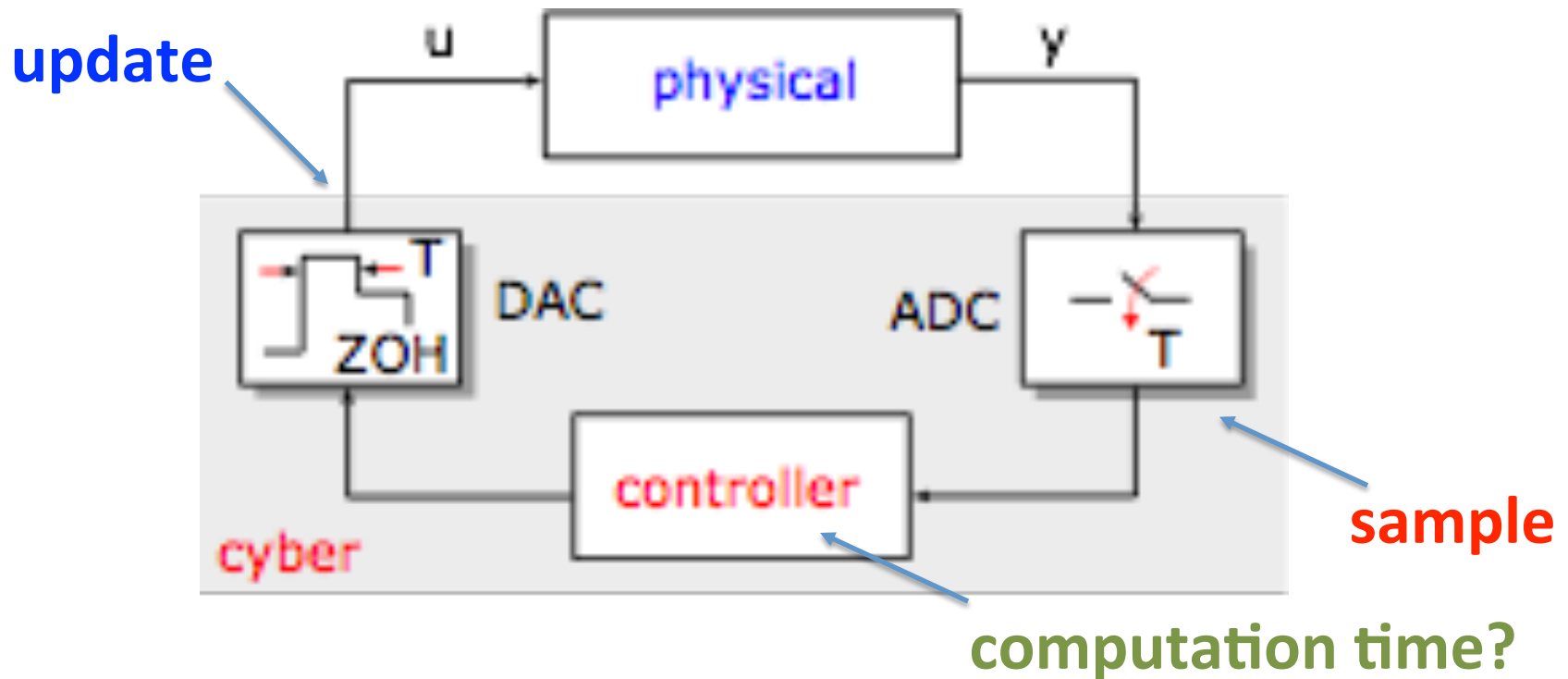
Synergy: Collaborative: Computationally Aware Cyber-Physical Systems
NSF CNS-1544395 (UCSC), NSF CNS-1544395 (UofA)



UNIVERSITY OF CALIFORNIA
SANTA CRUZ



What to do if my control algorithm **does not terminate** when an input is needed?



Feedback provides robustness to small computation time, but computing time could be large, specially in optimization-based approaches

Computationally-Aware Algorithms driven by Applications



UAVs in the National Air Space
(courtesy NASA)



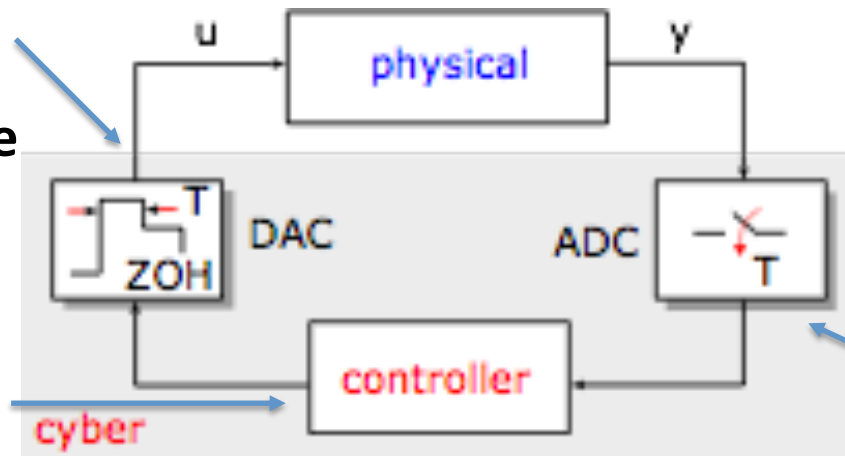
Self-driving Cars



MPC for Diesel Engines
(courtesy TOYOTA)

update

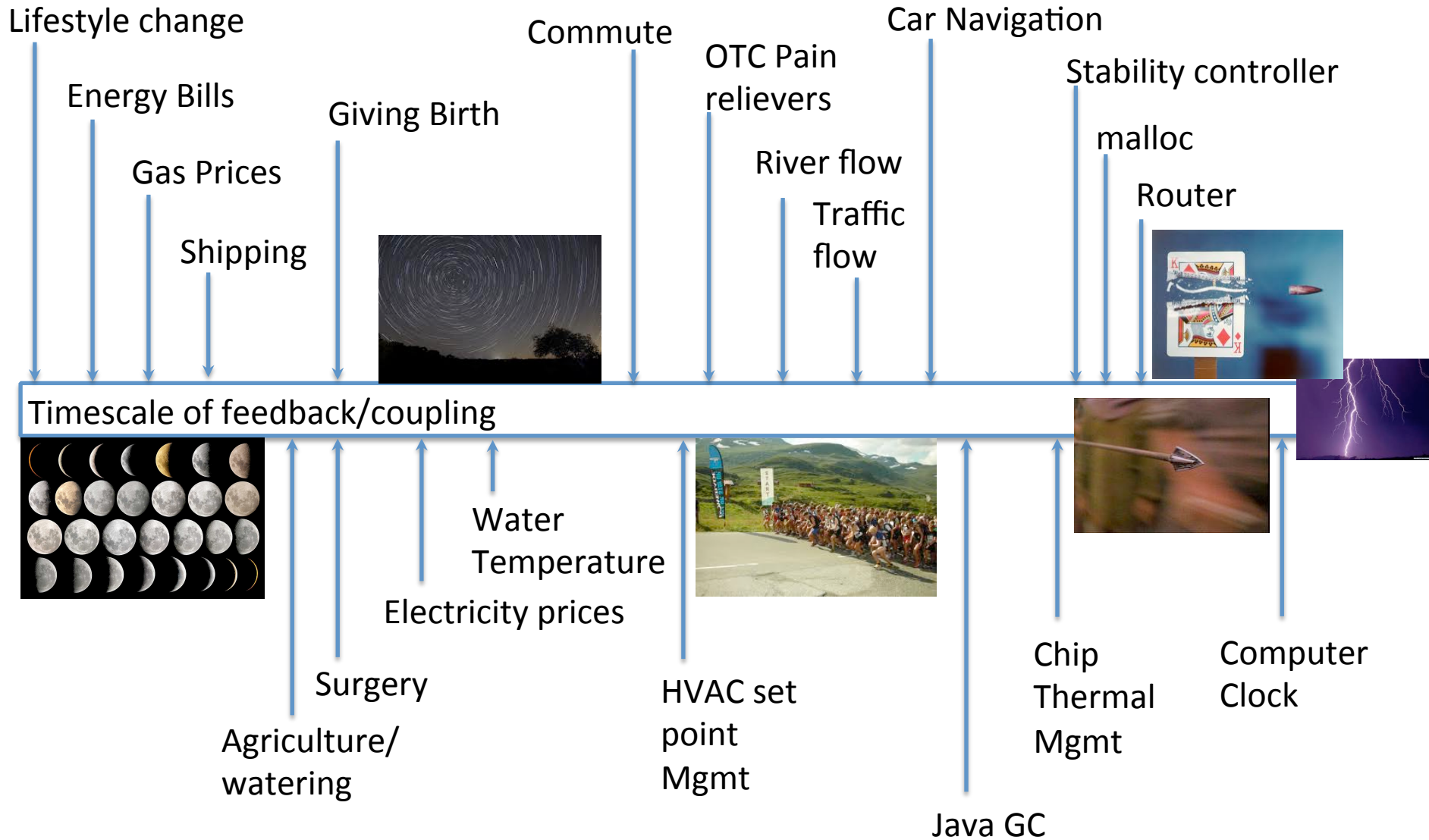
Exploit models of the physical system w/ different accuracy to reduce computation time!



sample

Trade off

Different timescales*



* Not to scale. It's not like I plotted this in MATLAB or anything...

Problem for vehicle control: comfort



<http://www.youtube.com/watch?v=ChZwuj3hCvw>

Or worse...



[Photo By Twilight Invasion](#)

Consider a moving ground vehicle



State update



State update

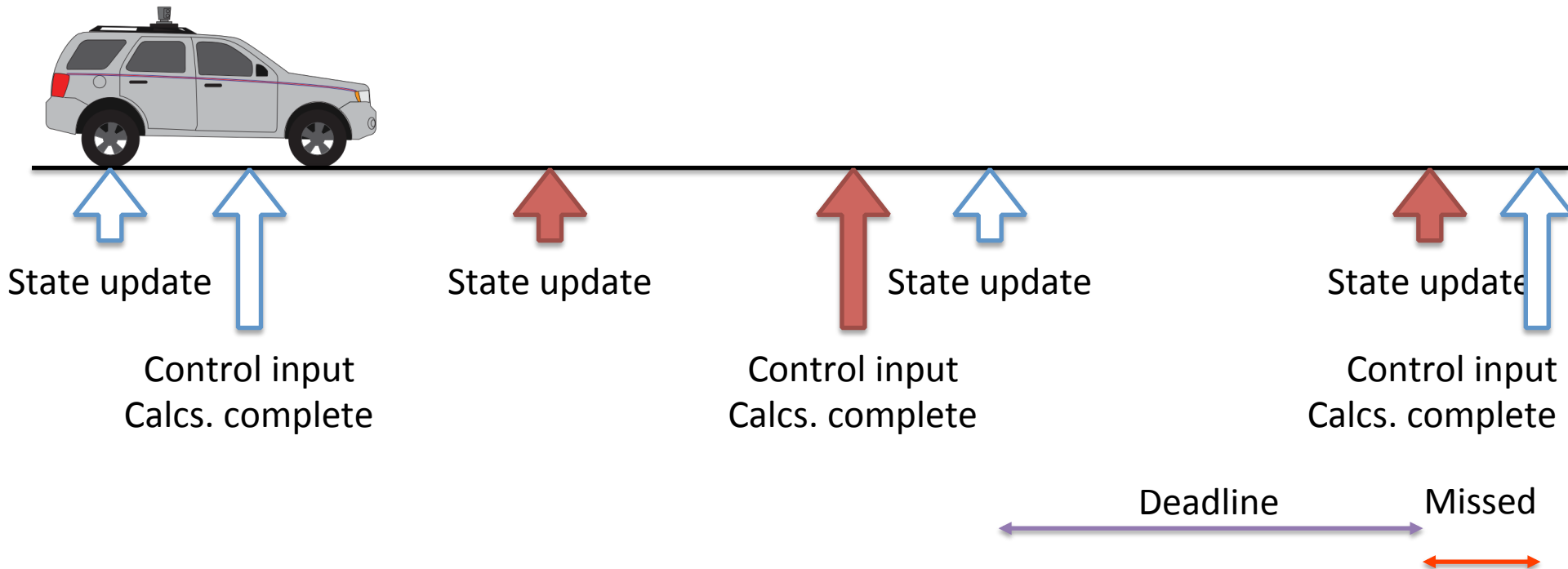


State update

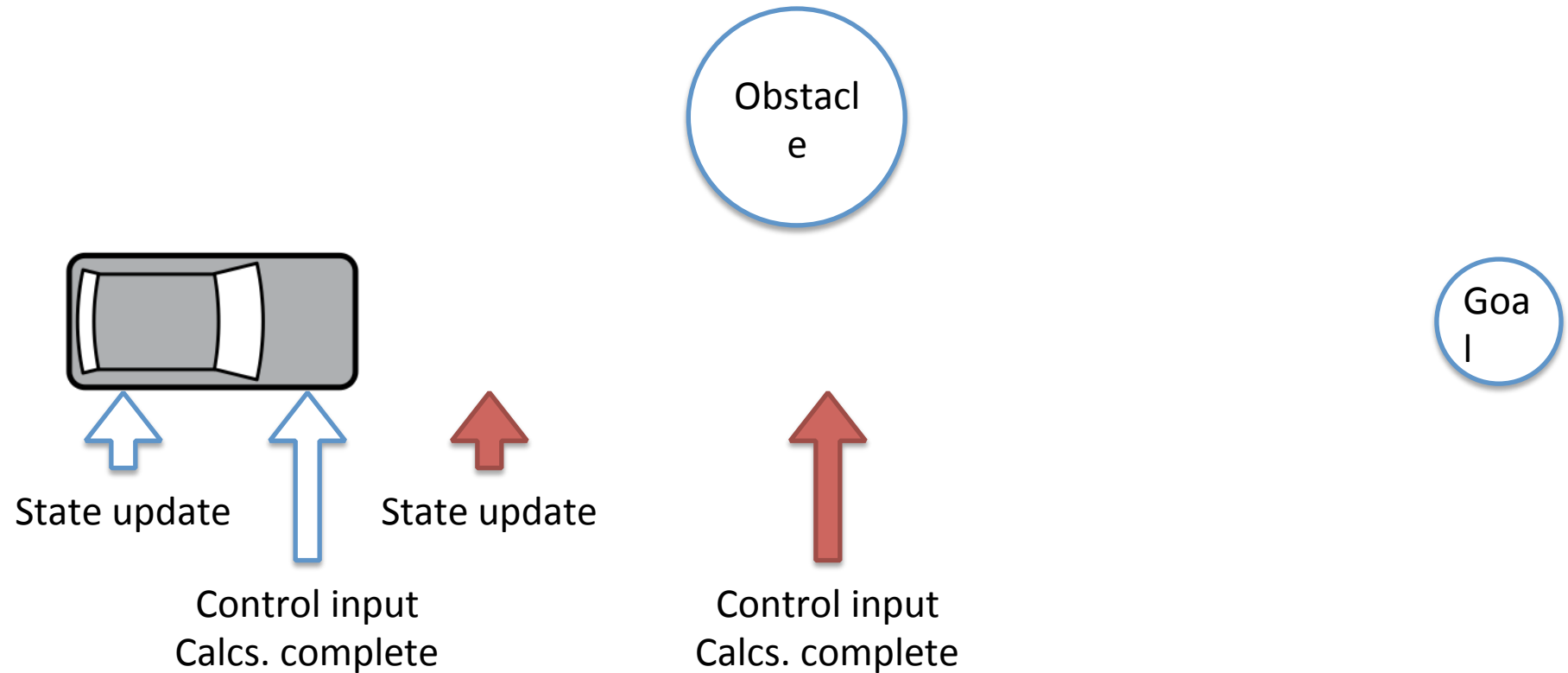


State update

When computing a control decision...



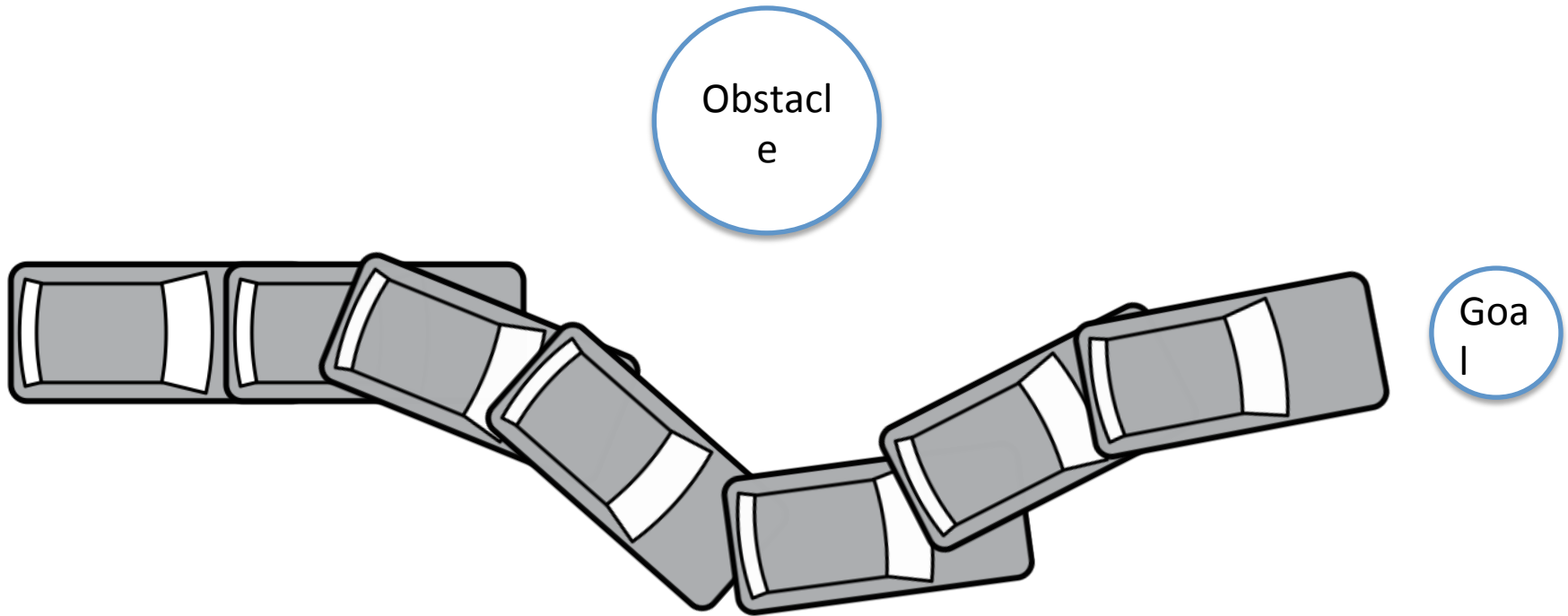
What if an obstacle moves while we compute?



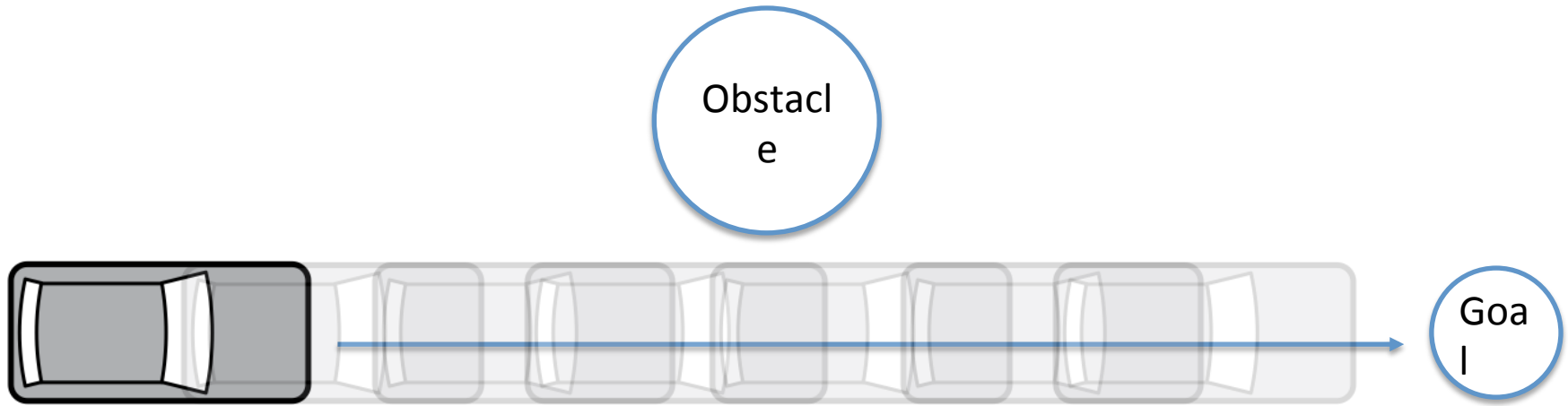
Obvious design goal: Reduce the period for state update/control input as much as possible.

This is a motivating animation, not a kinematic simulation.

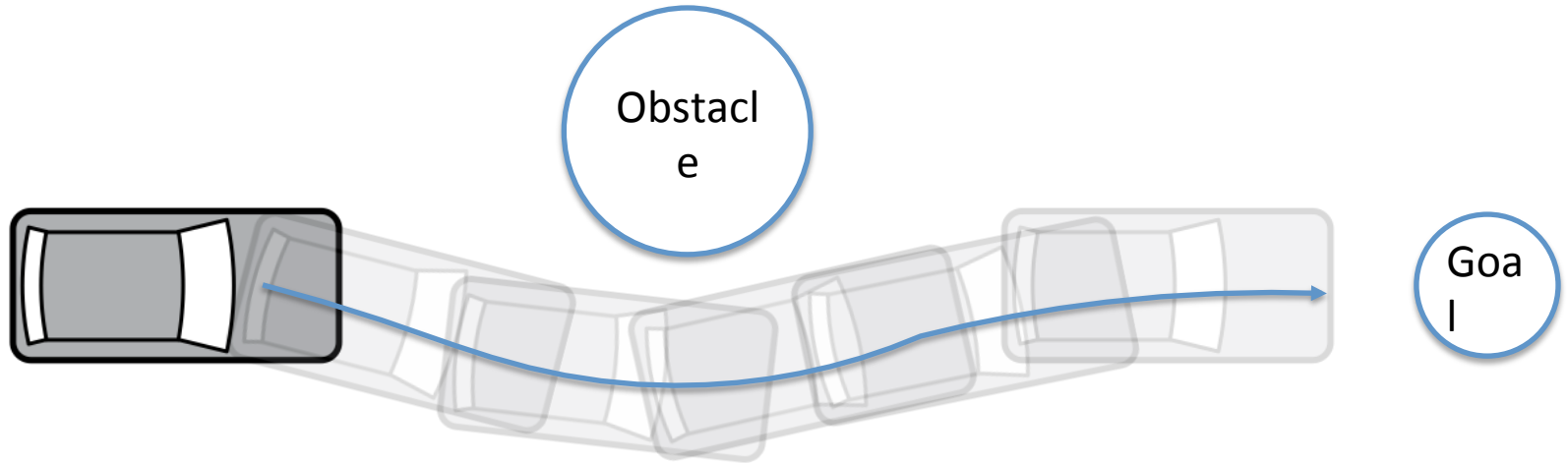
Model-Predictive Control: Plan Trajectories at Runtime



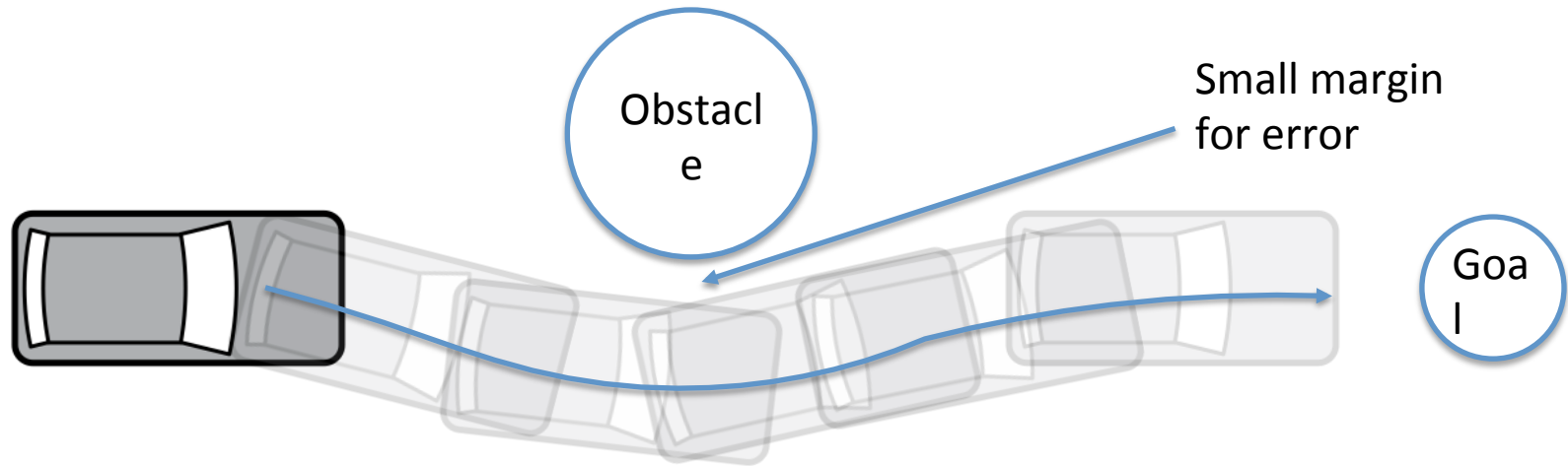
Model-Predictive Control: Plan Trajectories at Runtime



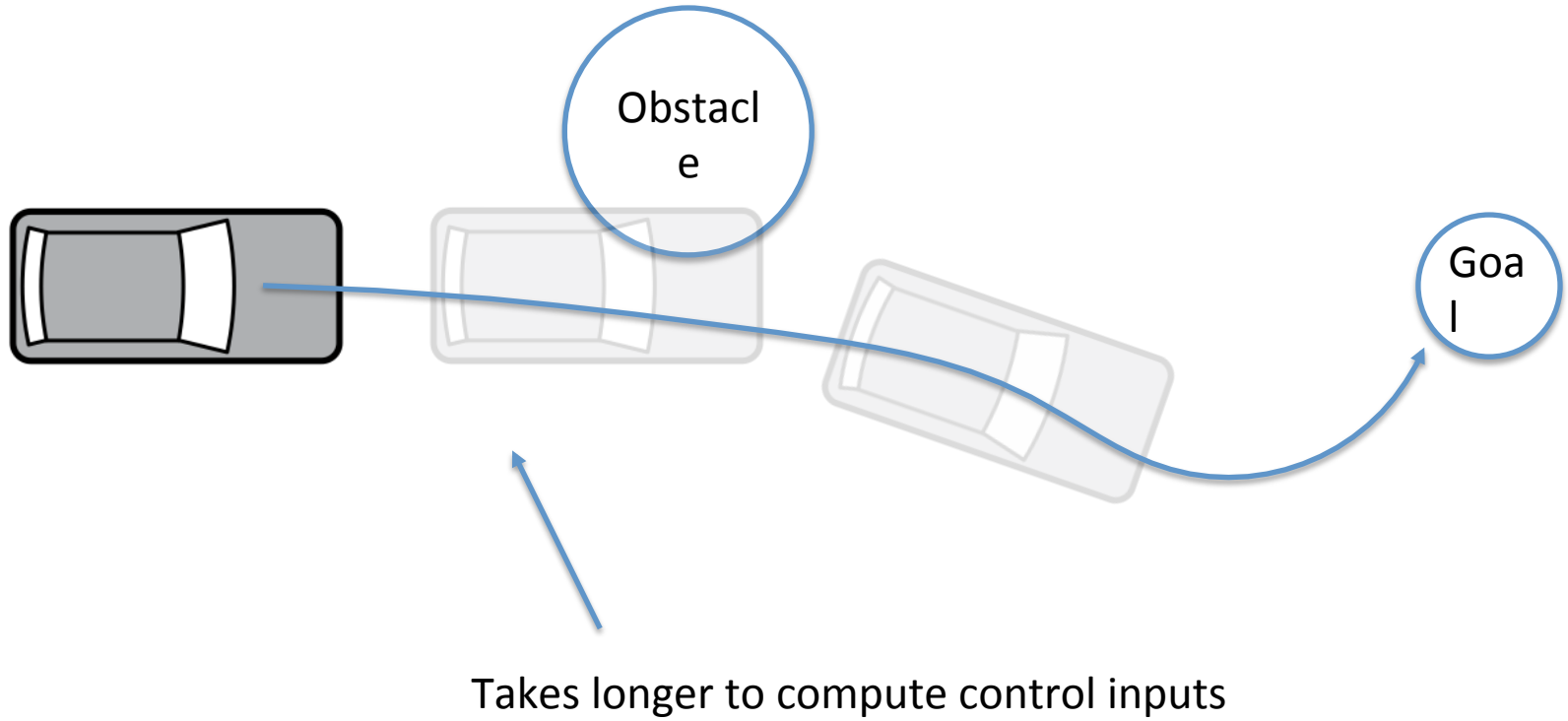
Model-Predictive Control: Plan Trajectories at Runtime



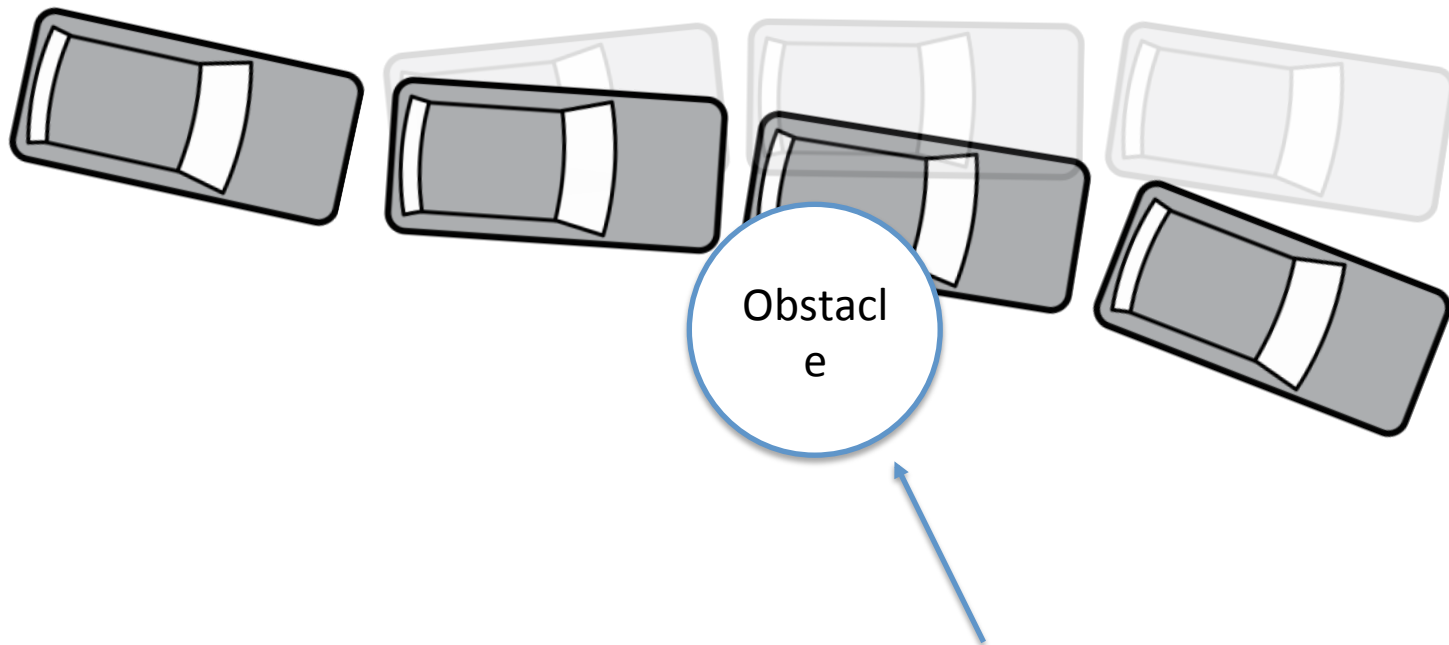
Model-Predictive Control: Plan Trajectories at Runtime



Easy...I'll use an accurate vehicle model to predict the trajectory and avoid the obstacle.

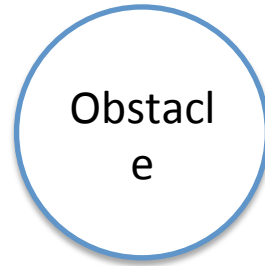
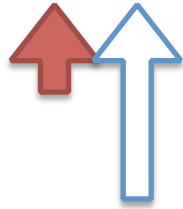
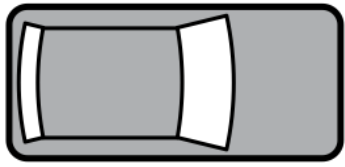


Easy...I'll pick a simpler vehicle model to make it more likely to return control inputs in time



That control input does not mean what you think it means.

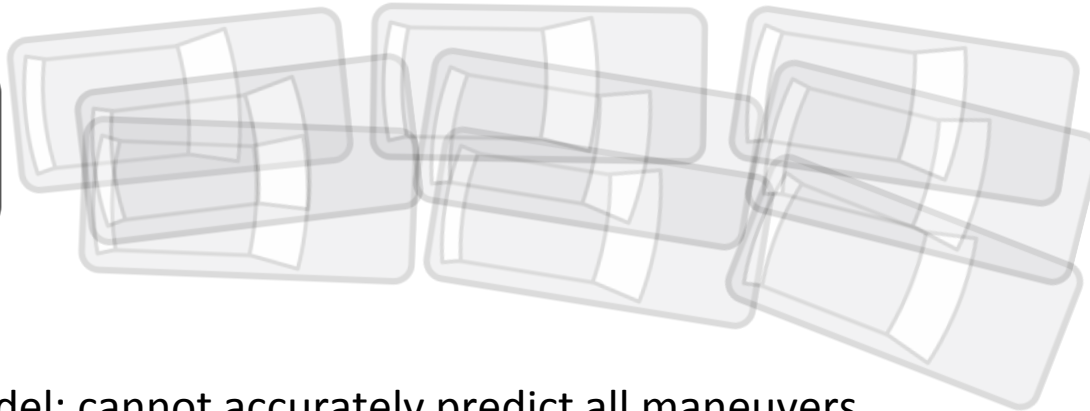
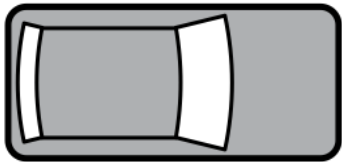
Competing constraints



High accuracy model: takes longer to optimize.



High vehicle speed: cannot tolerate slow return time.



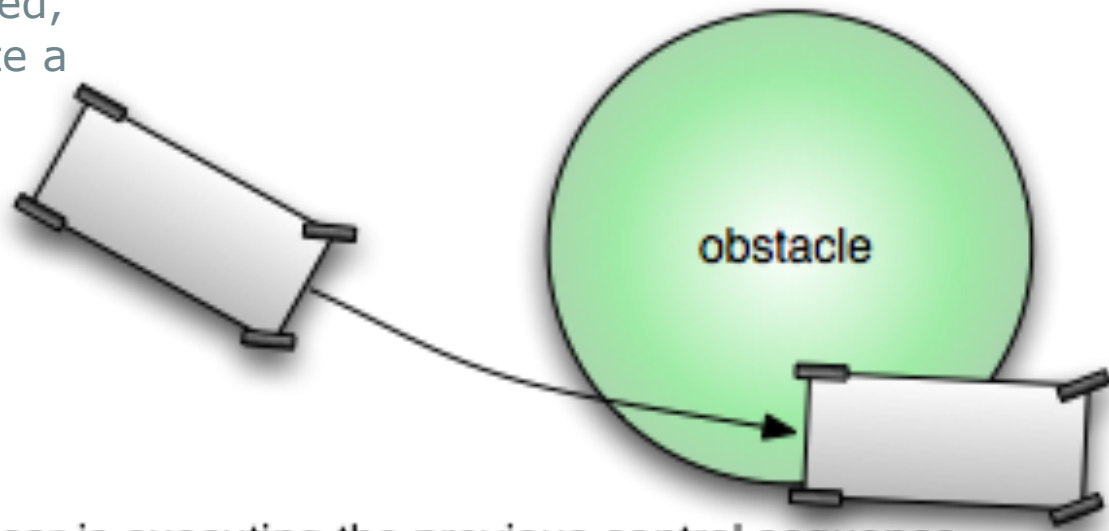
Low accuracy model: cannot accurately predict all maneuvers.

Goal: Computationally-Aware CPS

1. Consider the time required to perform the computation.
2. Switch between controllers using accuracy and time as switching criteria.
3. Explore conditions for stability and convergence.

Problems: MPC return time

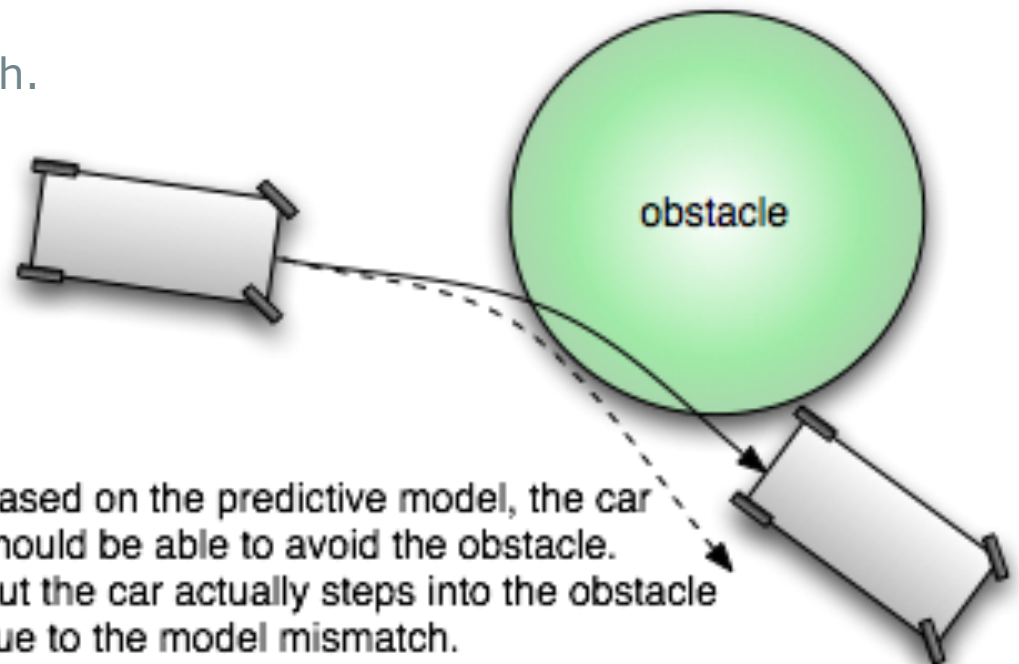
- A complex model may introduce predictive accuracy,
- however, this increases the computational burden.
- Especially under high speed, the system cannot tolerate a slow return rate.



The car is executing the previous control sequence while waiting for the new control sequence

Problems: control accuracy

- The return time problem can be addressed via model reduction,
 - potential drawback is higher model mismatch.
- Model mismatch can also introduce problems.
 - Wrong prediction
 - Infeasible trajectories



Model Predictive Control

$$\xi_{k,t+1} = \hat{f}_q(\xi_{k,t}, u_{k,t})$$
$$t \in \{k, k+1, \dots, k+N-1\}$$

MPC solves the optimization problem $\mathbf{P}^q(\xi_k)$ at time k by using the model \hat{f}_q . We denote the input sequence $\{u_{k,k}^q, u_{k,k+1}^q, \dots, u_{k,k+N-1}^q\}$ by U_k^q , and formulate the following problem:

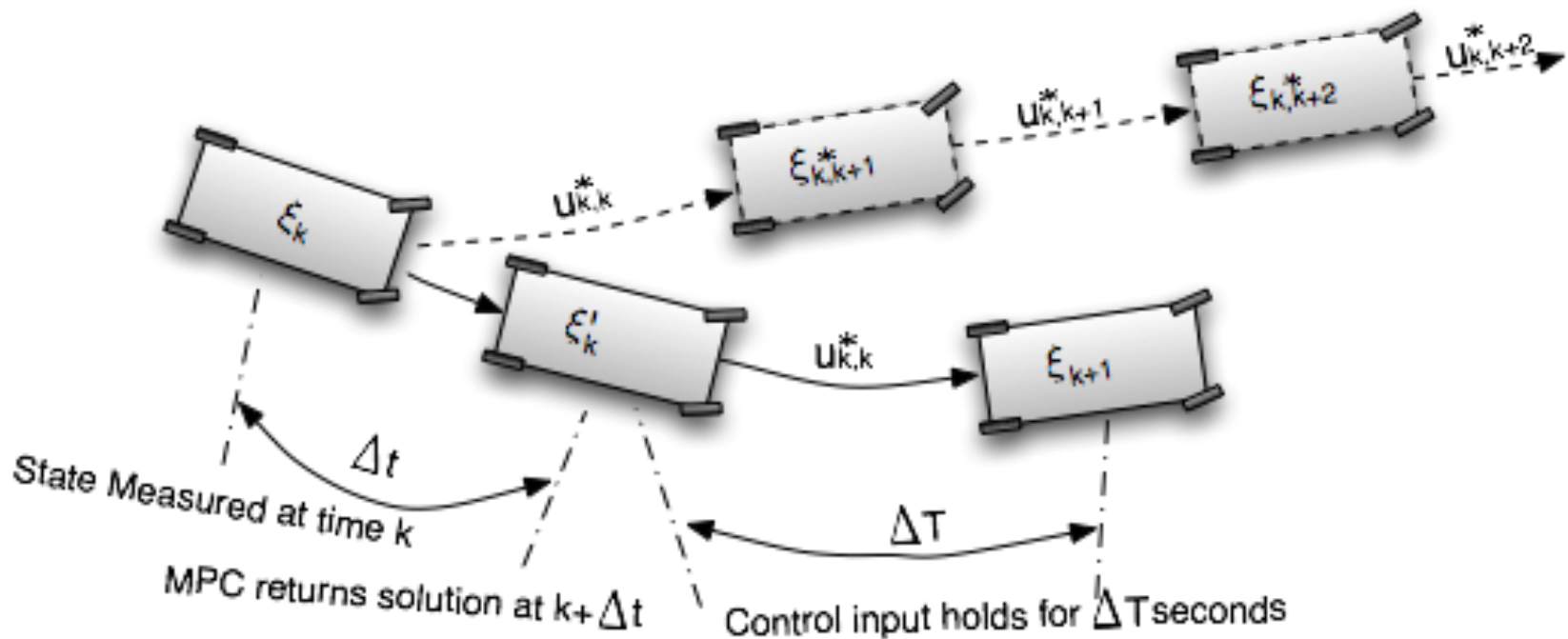
$$\mathbf{P}^q(\xi_k) : \underset{U_k^q}{\operatorname{argmin}} \{J_N(\xi_k, U_k^q) : U_k^q \subset \mathbb{R}^m\}$$
$$J_N(\xi_k, U_k^q) = \sum_{t=k}^{k+N-1} \ell(\xi_{k,t}^q, u_{k,t}^q) + F(\xi_{k,k+N}^q)$$
$$U_k^{q*} = \{u_{k,k}^{q*}, u_{k,k+1}^{q*}, \dots, u_{k,k+N-1}^{q*}\}$$

Problem Modeling

$$\xi'_k \approx \xi_k + f(\xi_k, u_{k-1,k}^*) \Delta t$$

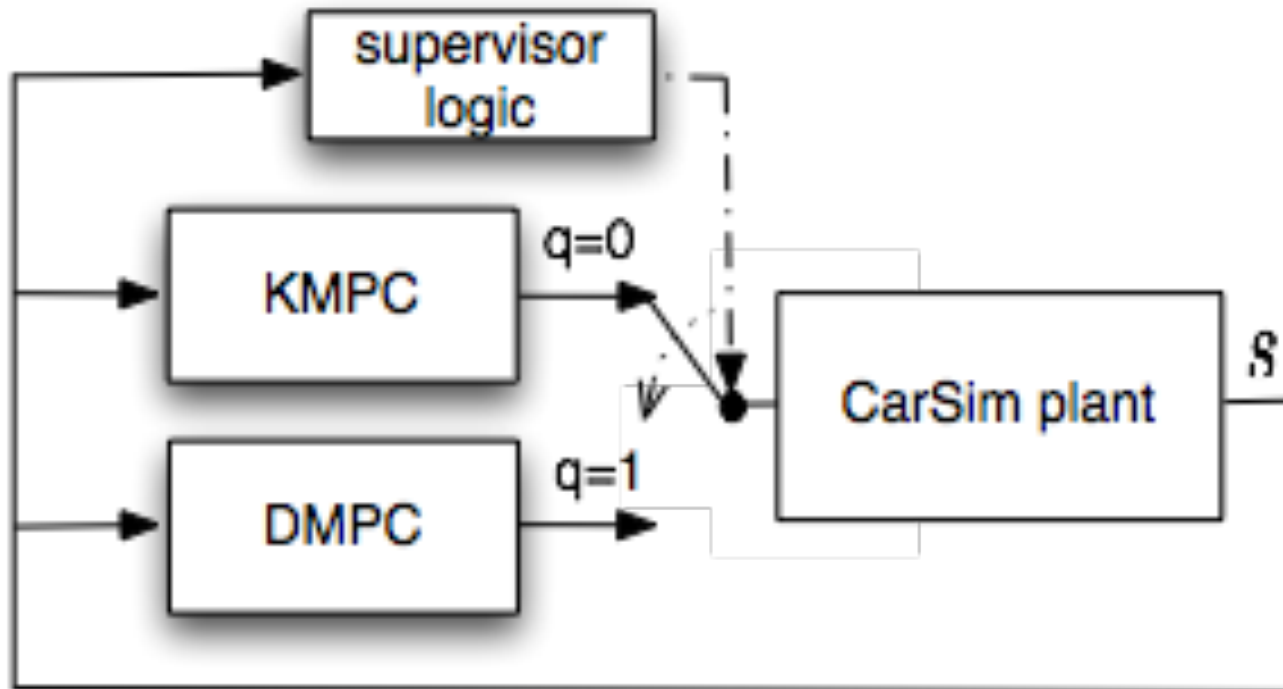
$$\xi_{k+1} = \hat{f}(\xi'_k, \kappa_q(\xi_k)) = \hat{f}_q(\xi'_k, \kappa_q(\xi_k)) + \hat{\Gamma}_q(\xi'_k)$$

$$\approx \hat{f}_q(\xi_k, \kappa_q(\xi_k)) + \hat{\Gamma}_q(\xi_k) + \left(\frac{\partial \hat{f}_q(\xi_k, \kappa_q(\xi_k))}{\partial \xi} + \frac{\partial \hat{\Gamma}_q(\xi_k, \kappa_q(\xi_k))}{\partial \xi} \right) f(\xi_k, \kappa_q(\xi_k)) \Delta t_q(\xi_k)$$



Problem Statement

Suppose at time $k \in \{0, 1, \dots\}$, the vehicle state ξ_k is observed for an optimization problem indexed by the predictive model in use (i.e., $\mathbf{P}^q(\xi_k)$), and that two alternative predictive models are available.



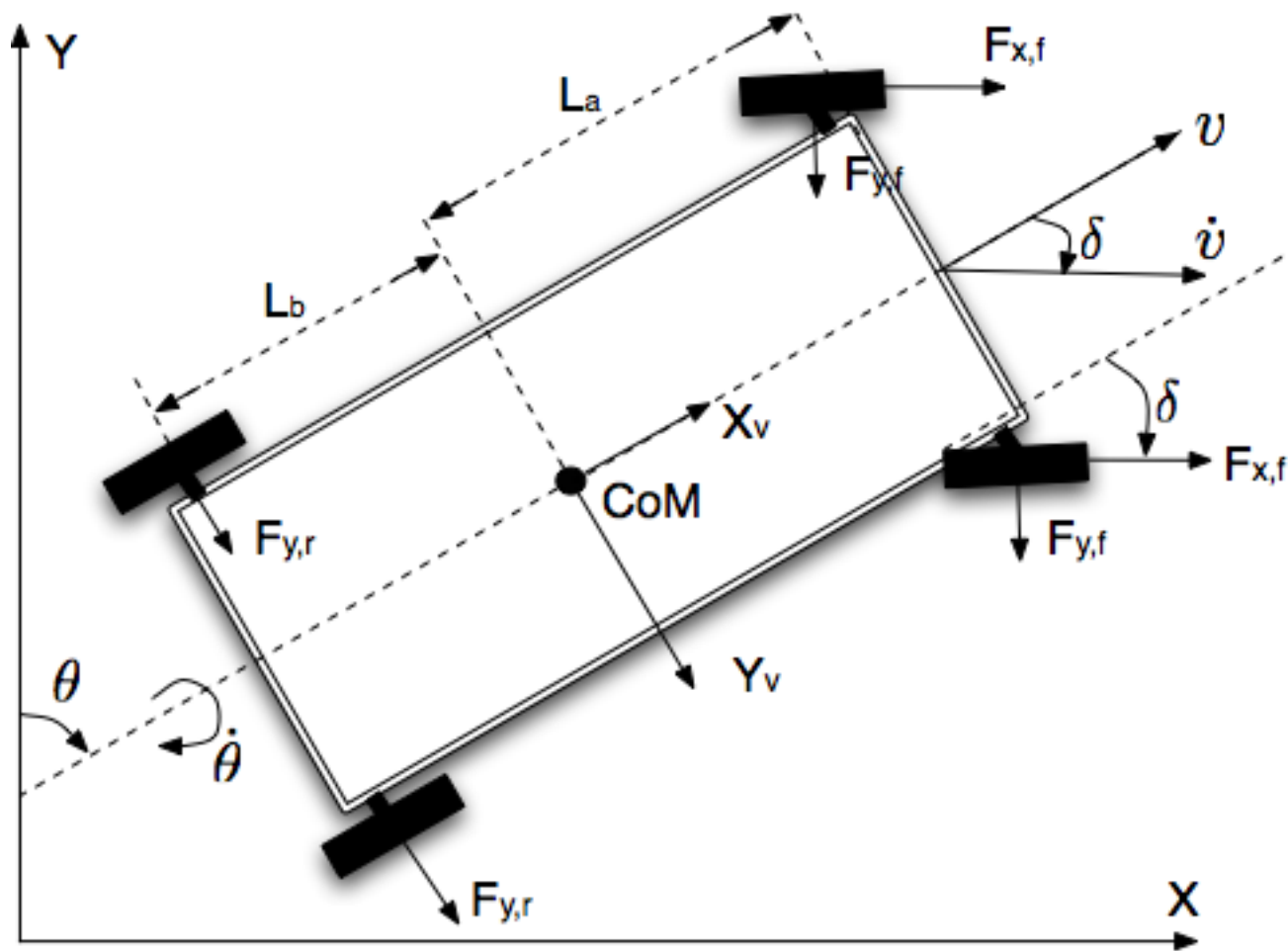
Problem: select the predictive model q such that the divergence of the state at ξ_k from the plant's state with the same inputs is minimized.

Kinematic/Dynamical Models

$$\dot{\xi} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \frac{v \tan \delta}{L} \end{bmatrix}$$

$$\dot{\xi} = \begin{bmatrix} v \sin(\theta) \\ v \cos(\theta) \\ \cos(\delta)a - \frac{2}{m}F_{y,f} \sin(\delta) \\ \varphi \\ \frac{1}{J} (L_a (ma \sin(\delta) + 2F_{y,f} \cos(\delta)) - 2L_b F_{y,r}) \\ \omega \end{bmatrix}$$

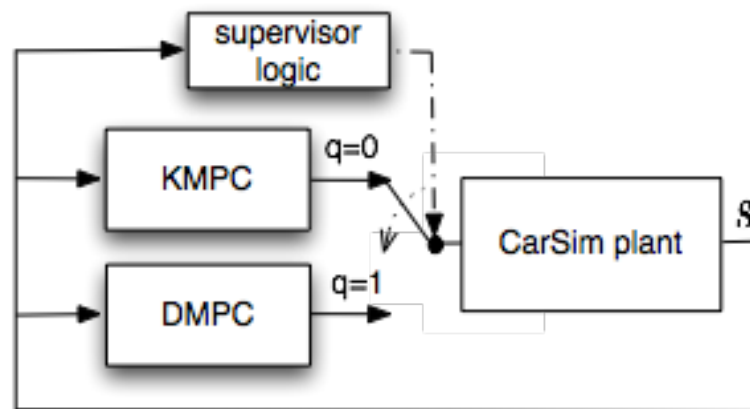
$$\dot{\xi} = \begin{bmatrix} v \sin(\theta) \\ v \cos(\theta) \\ \cos(\delta)a - \frac{2}{m}F_{y,f} \sin(\delta) \\ \varphi \\ \frac{1}{J} (L_a (ma \sin(\delta) + 2F_{y,f} \cos(\delta)) - 2L_b F_{y,r}) \\ \omega \end{bmatrix}$$



Hybrid MPC Design

$$\dot{\xi}^{q=0} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \frac{v \tan \delta}{L} \end{bmatrix} \quad \dot{\xi}^{q=1} = \begin{bmatrix} v \sin(\theta) \\ v \cos(\theta) \\ \cos(\delta)a - \frac{2}{m}F_{y,f} \sin(\delta) \\ \varphi \\ \frac{1}{J} (L_a (ma \sin(\delta) + 2F_{y,f} \cos(\delta)) - 2L_b F_{y,r}) \\ \omega \end{bmatrix}$$

$$q = \underset{q}{\operatorname{argmin}} \left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$



Hybrid MPC Design

$$q = \operatorname{argmin}_q \left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$

To make this decision, we need to know two things:

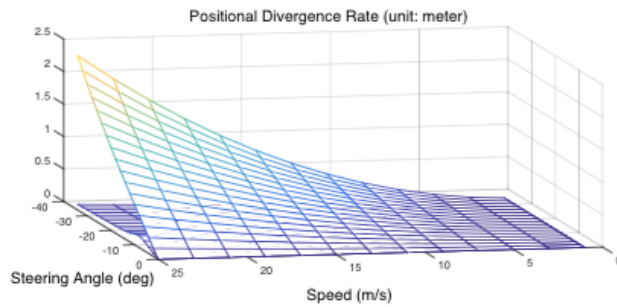
(1) Model mismatch

(2) Return time for MPC for this model

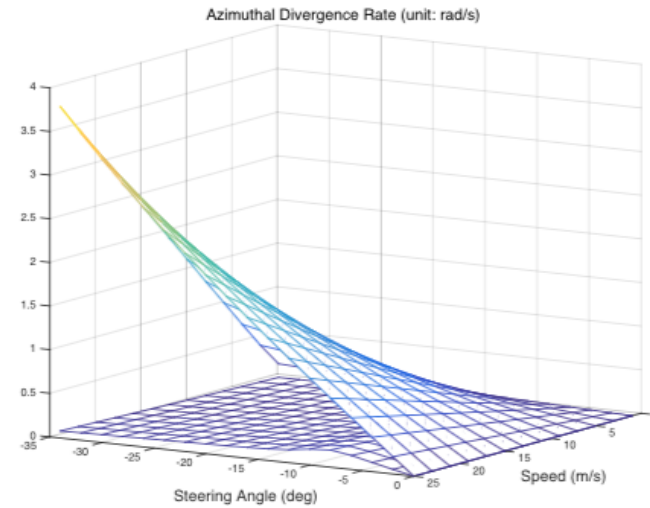
$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\| \leq \left\| \hat{\Gamma}_q(\xi_k) \right\| + \left\| I + \frac{\partial f(\xi_k, \kappa_q(\xi_k))}{\partial \xi} \Delta T \right\| \left\| f(\xi_k, \kappa_q(\xi_k)) \right\| \Delta t_q(\xi_k)$$

$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\| \approx \left\| \hat{\Gamma}_q(\xi_k) \right\| + v \Delta t_q(\xi_k) \sqrt{1 + \left(\frac{\tan \delta}{L} v \Delta T \right)^2}$$

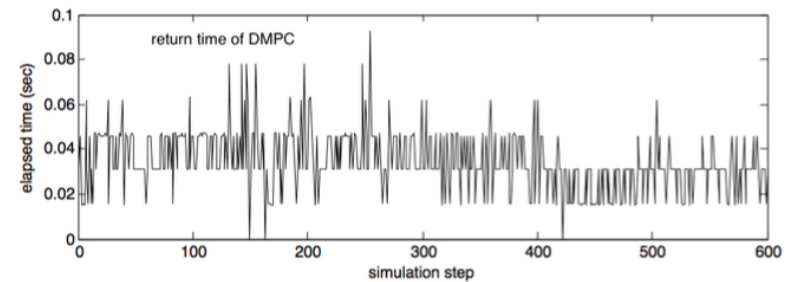
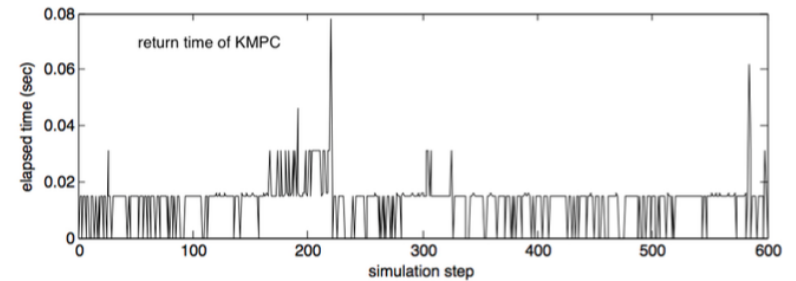
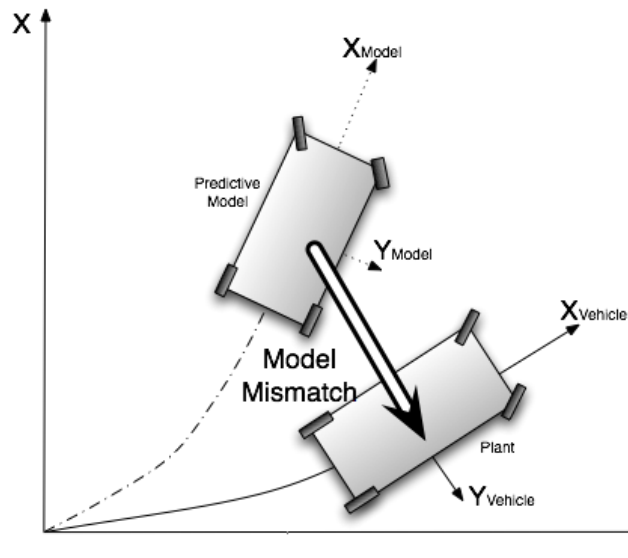
Model Mismatch & Return Time



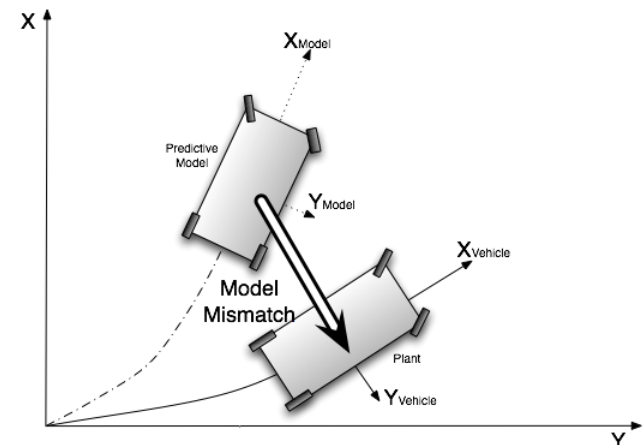
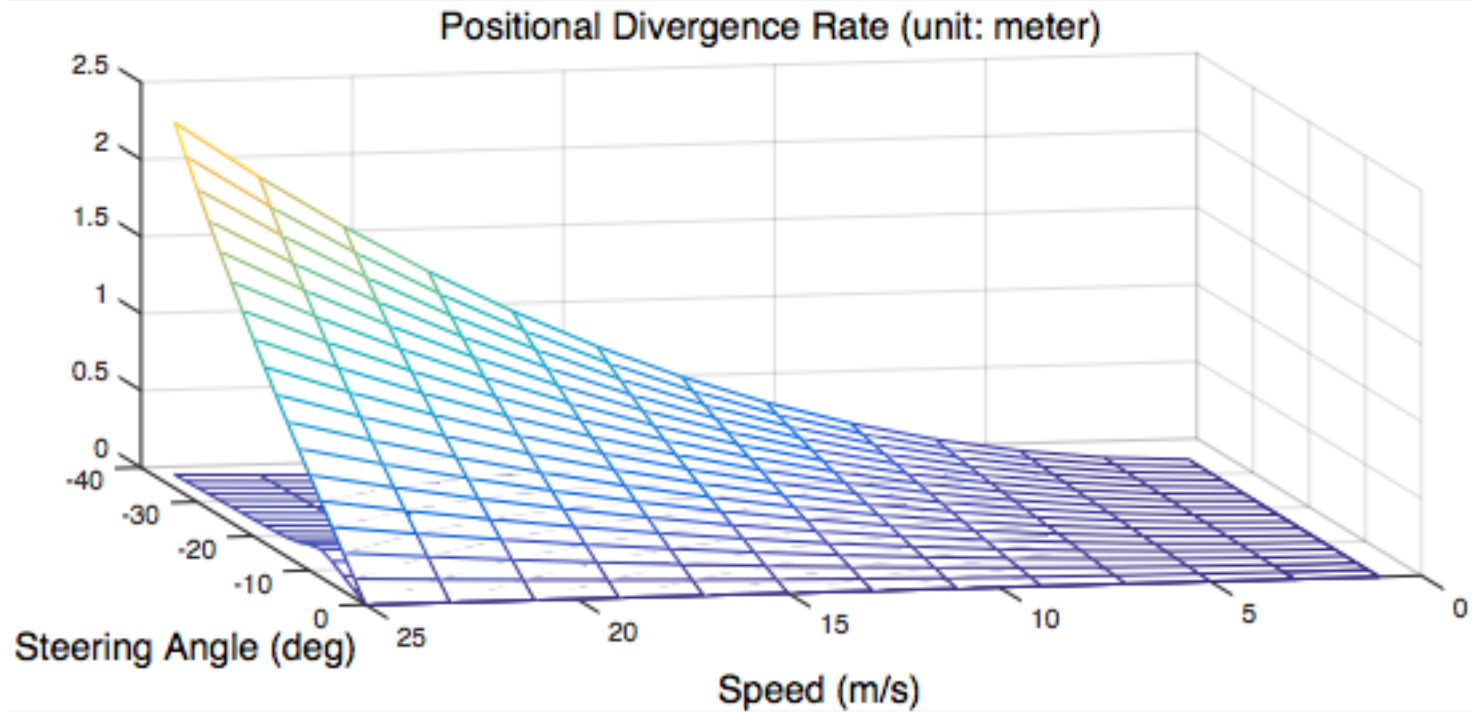
(a)



(b)

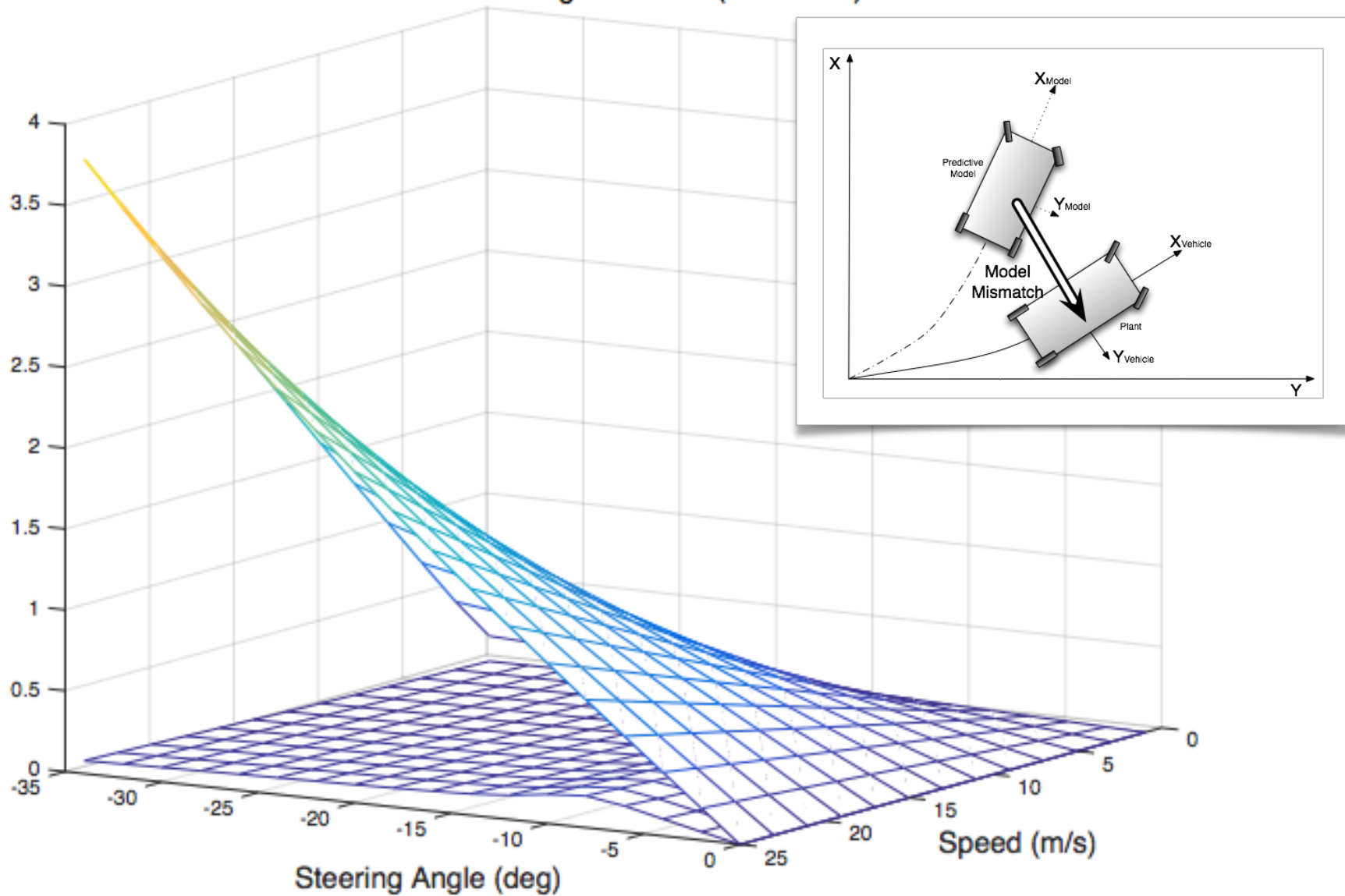


Model Mismatch

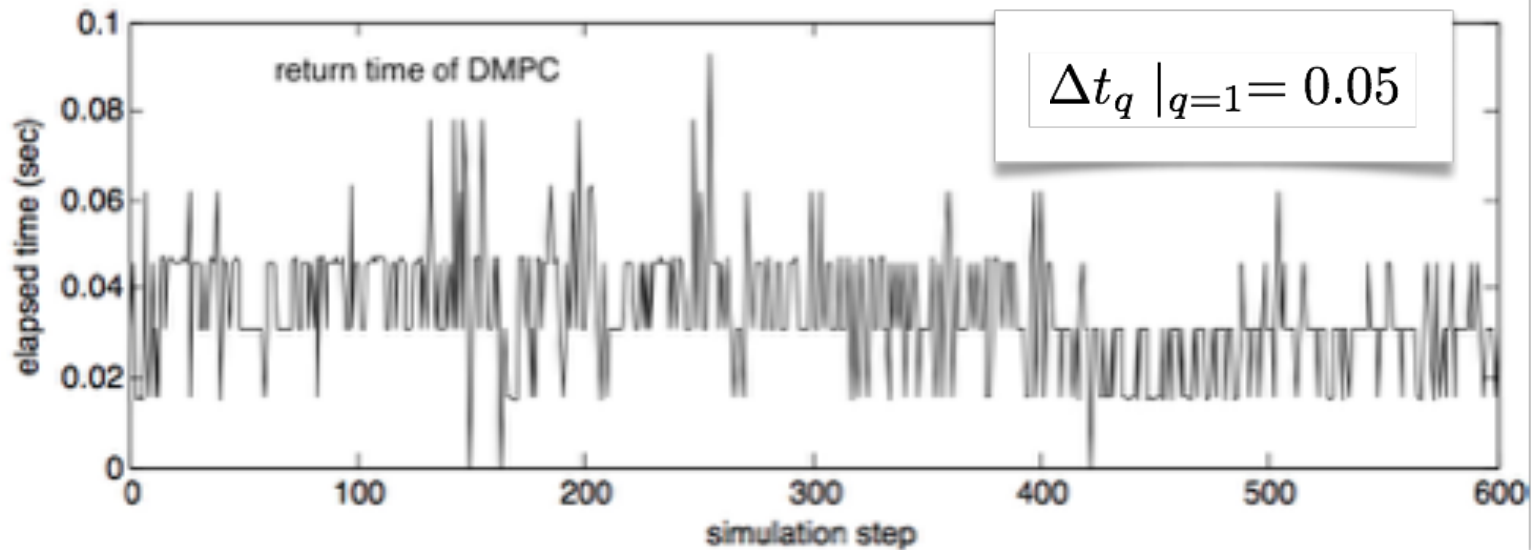
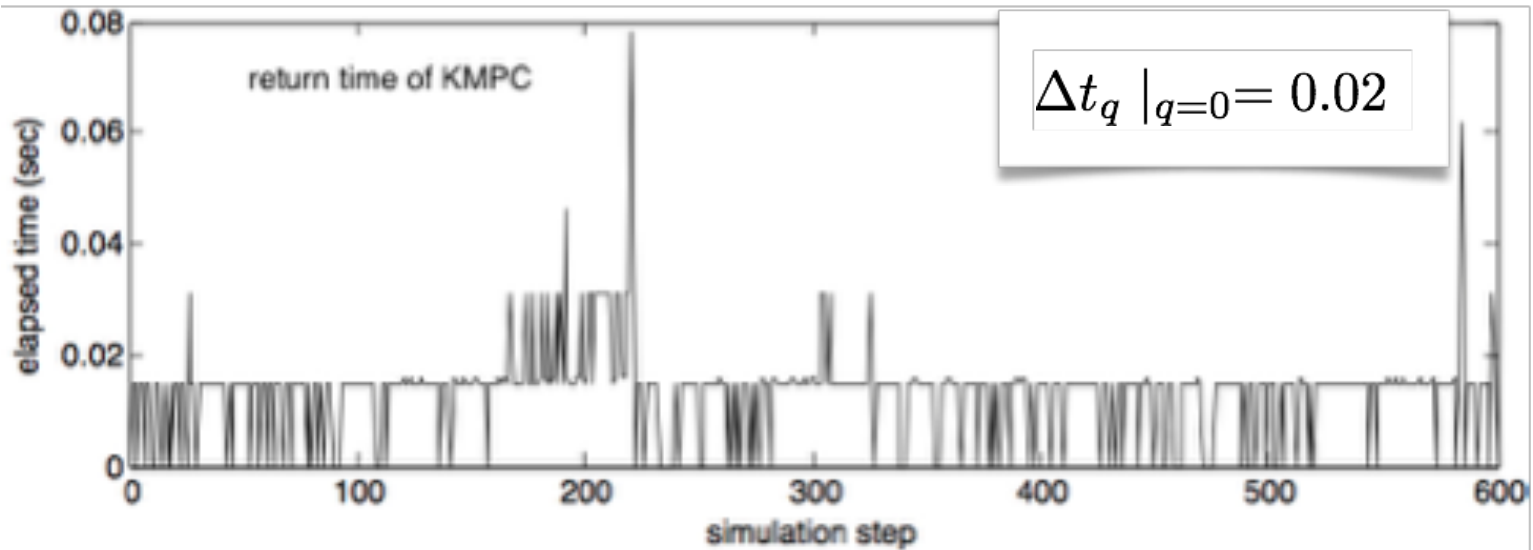


Model Mismatch

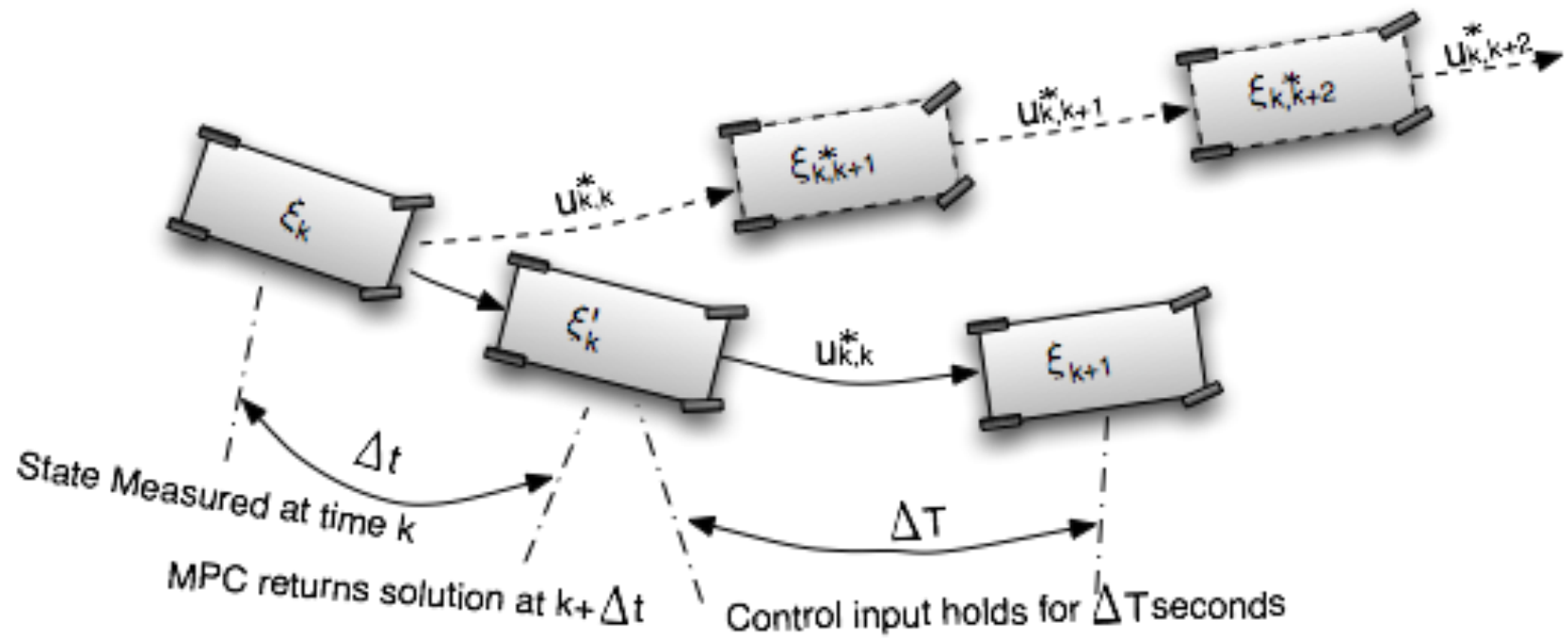
Azimuthal Divergence Rate (unit: rad/s)



Time to return



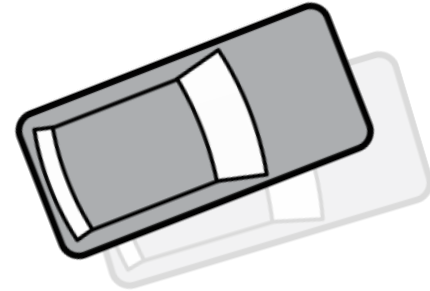
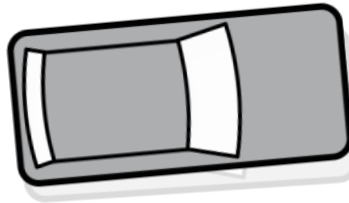
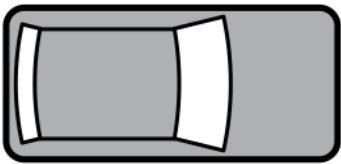
Uncontrollable divergence



$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$

Uncontrollable divergence: DMPC

$\Delta t_q = 0.05$ (DMPC value)

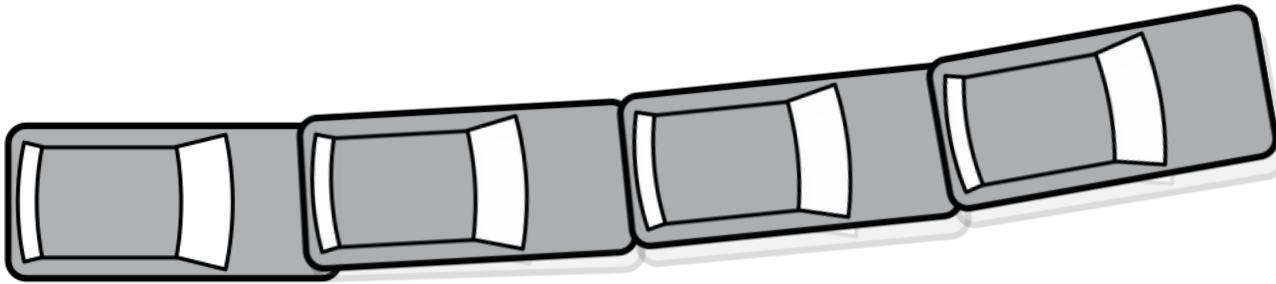


$$\dot{\xi}^{q=1} = \begin{bmatrix} v \sin(\theta) \\ v \cos(\theta) \\ \cos(\delta)a - \frac{2}{m}F_{y,f} \sin(\delta) \\ \varphi \\ \frac{1}{J} (L_a (ma \sin(\delta) + 2F_{y,f} \cos(\delta)) - 2L_b F_{y,r}) \\ \omega \end{bmatrix}$$

$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$

Uncontrollable divergence: KMPC

$\Delta t_q = 0.02$ (KMPC value...)

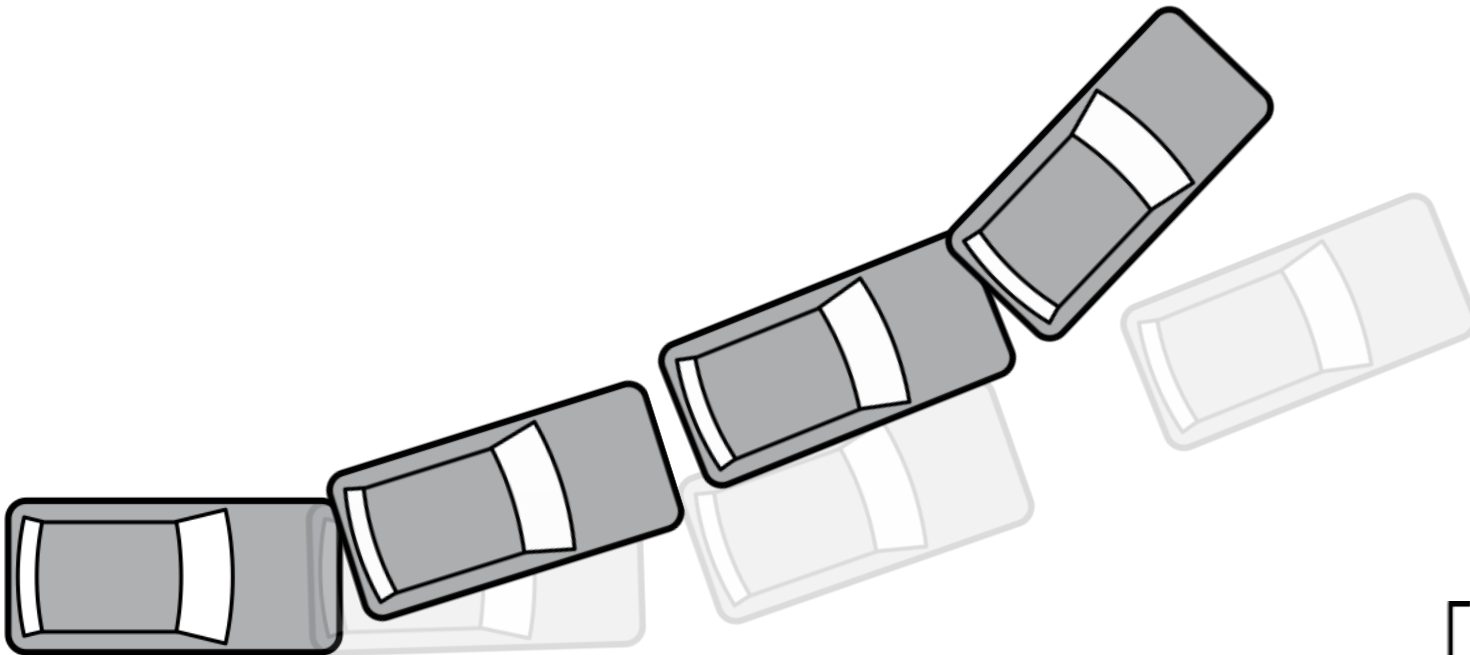


$$\dot{\xi}^{q=0} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \frac{v \tan \delta}{L} \end{bmatrix}$$

$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$

Uncontrollable divergence: KMPC with large steering

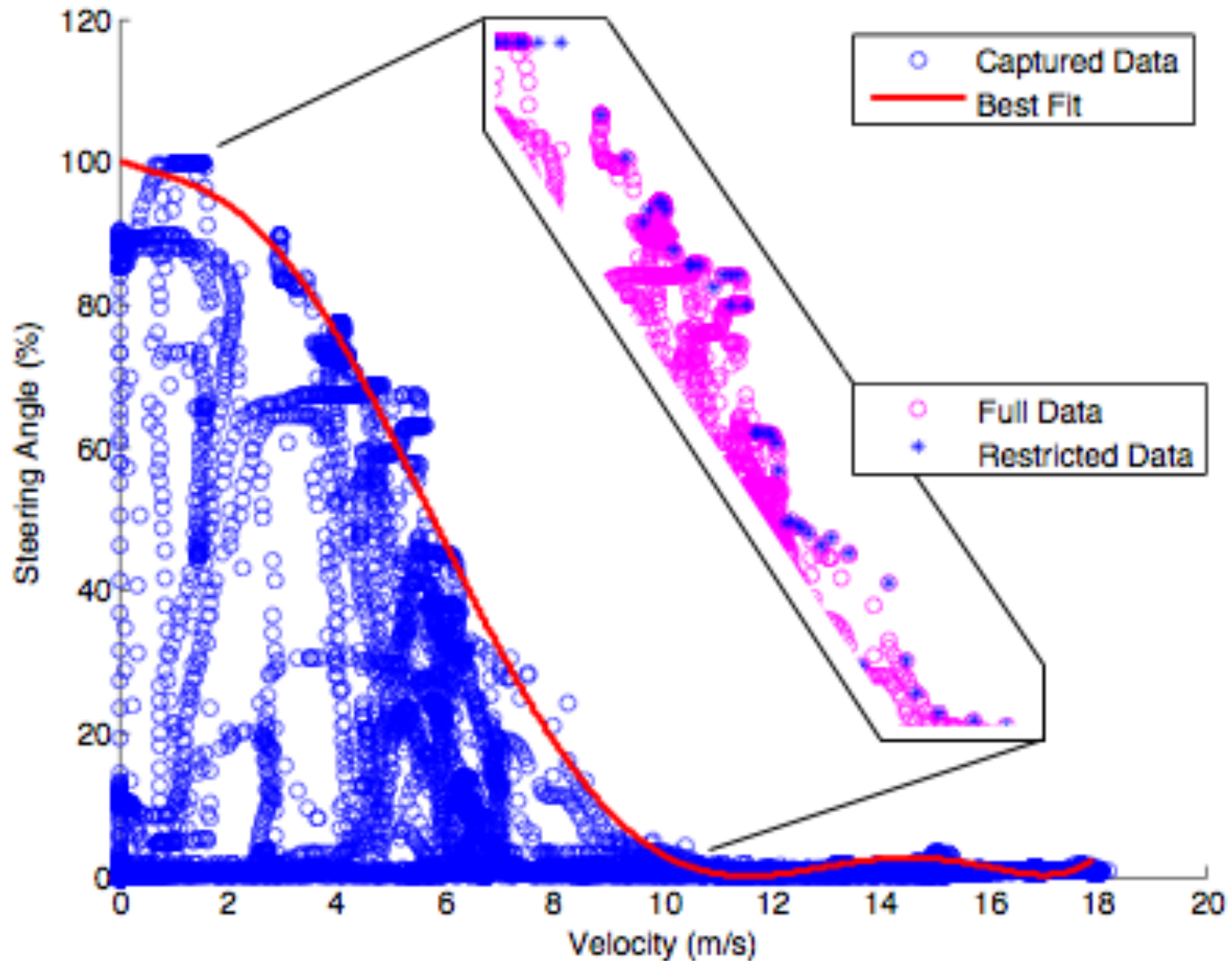
$\Delta t_q = 0.02$ (KMPC value...)



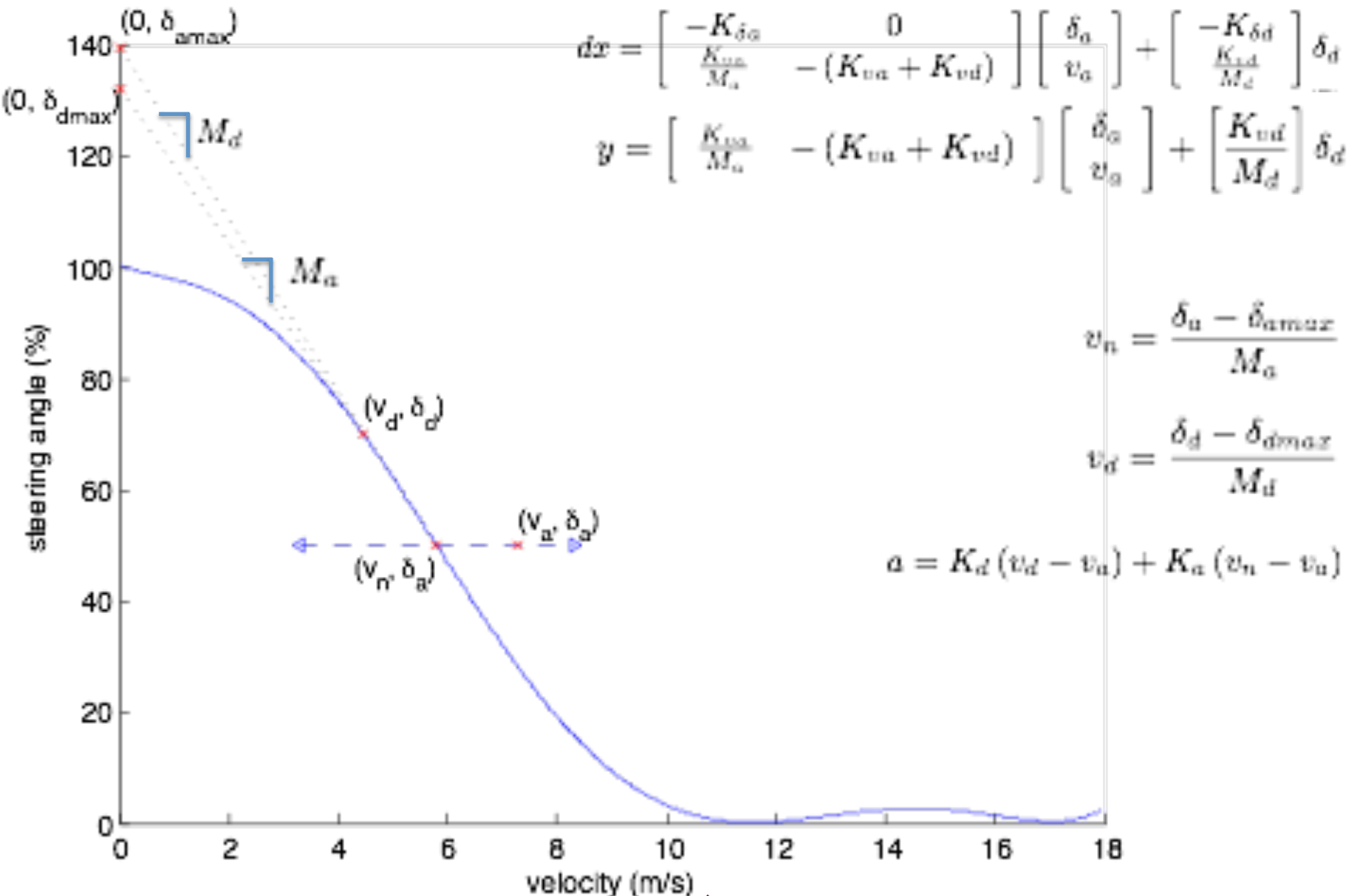
$$\dot{\xi}^{q=0} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \frac{v \tan \delta}{L} \end{bmatrix}$$

$$\left\| \xi_{k+1} - \xi_{k,k+1}^{q*} \right\|$$

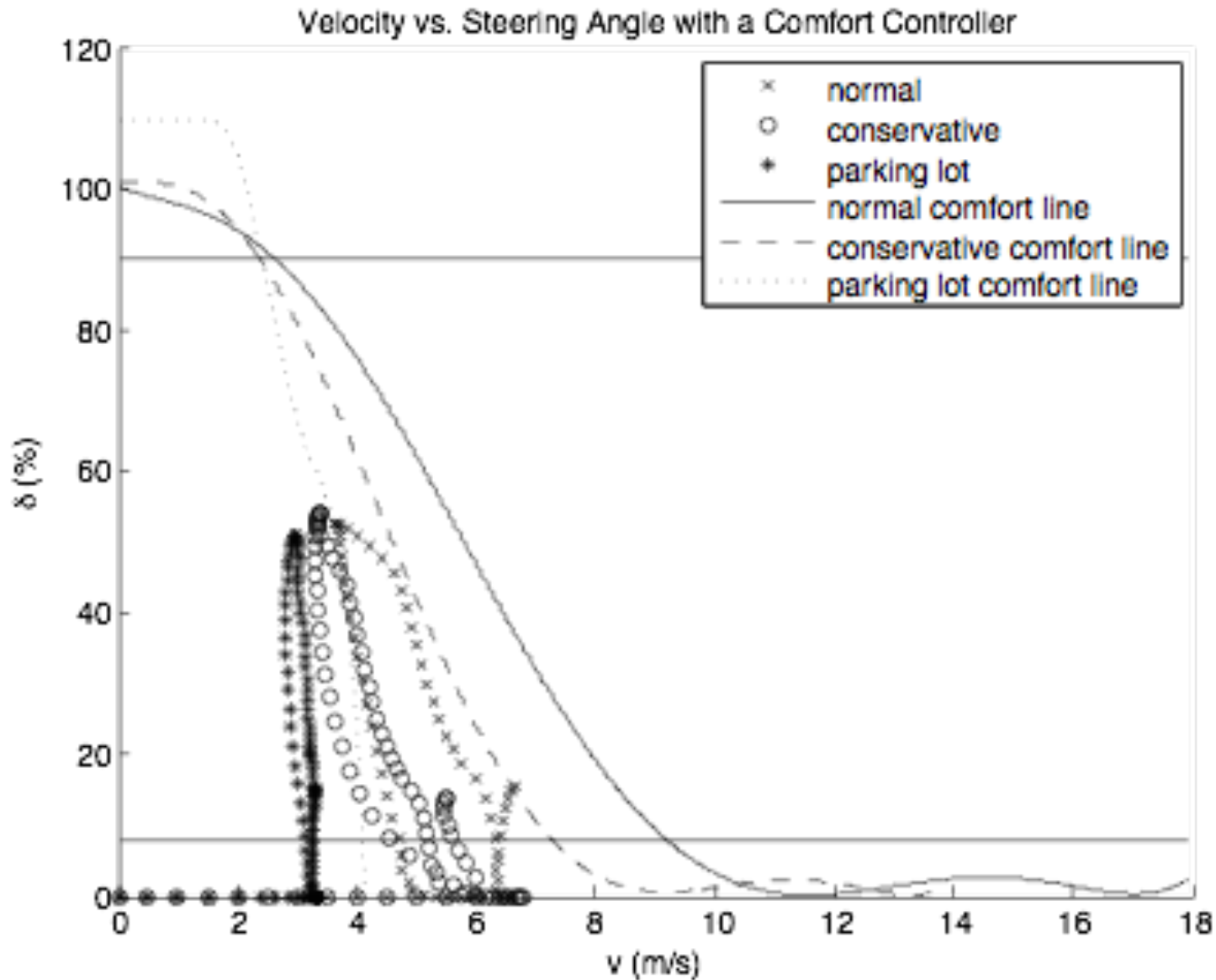
Fortunate Previous Result from [1]



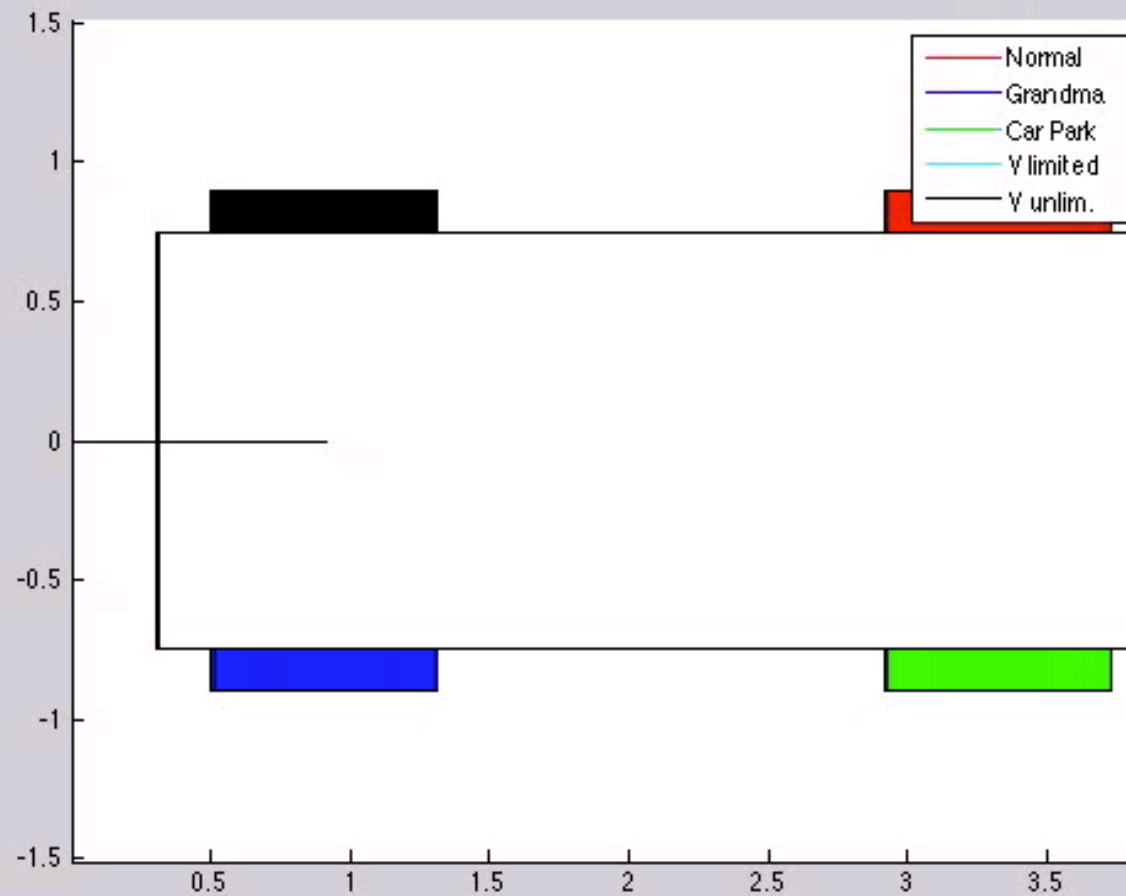
Take the fit data and utilize linearization techniques

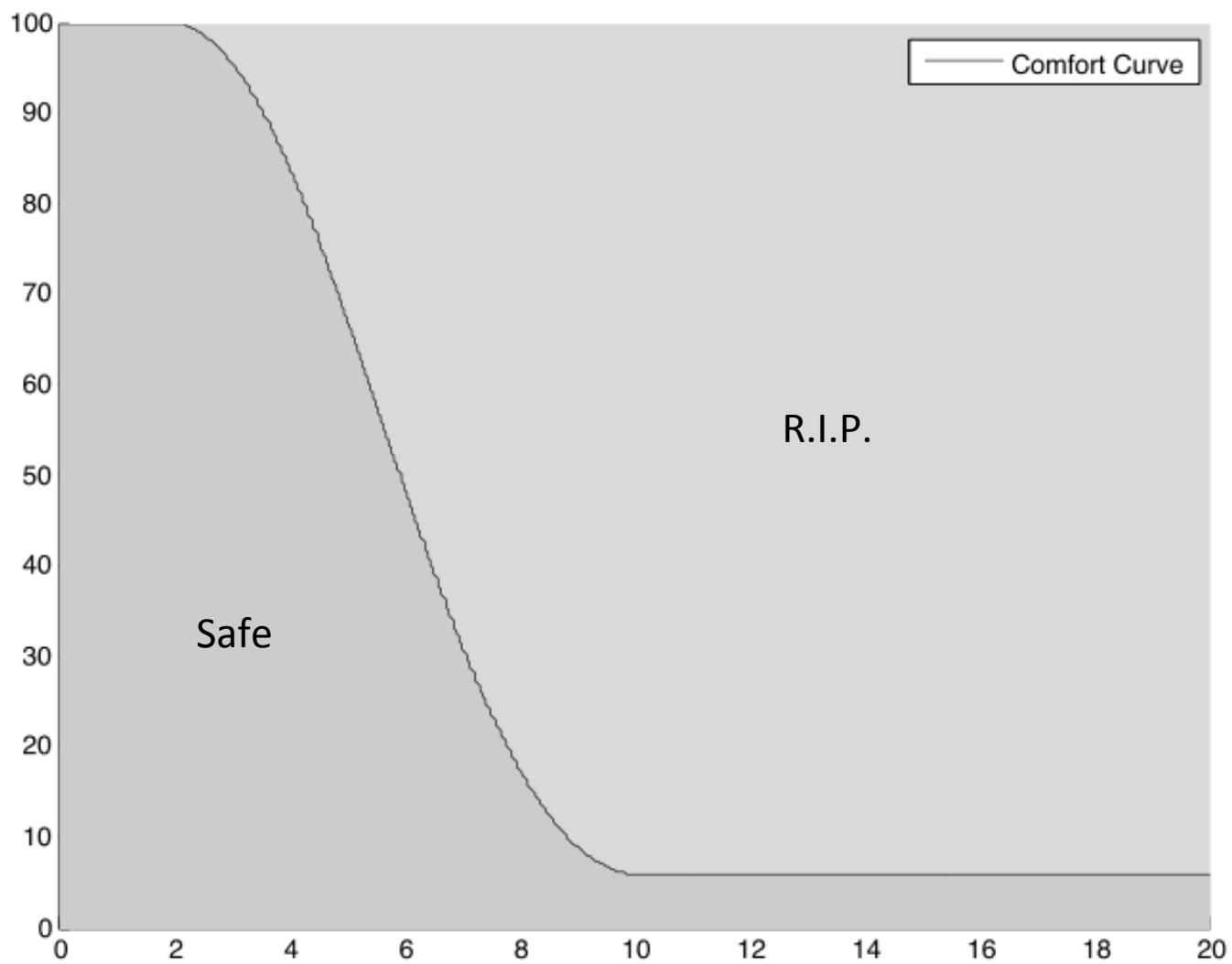


Scatter plot with comfort controller

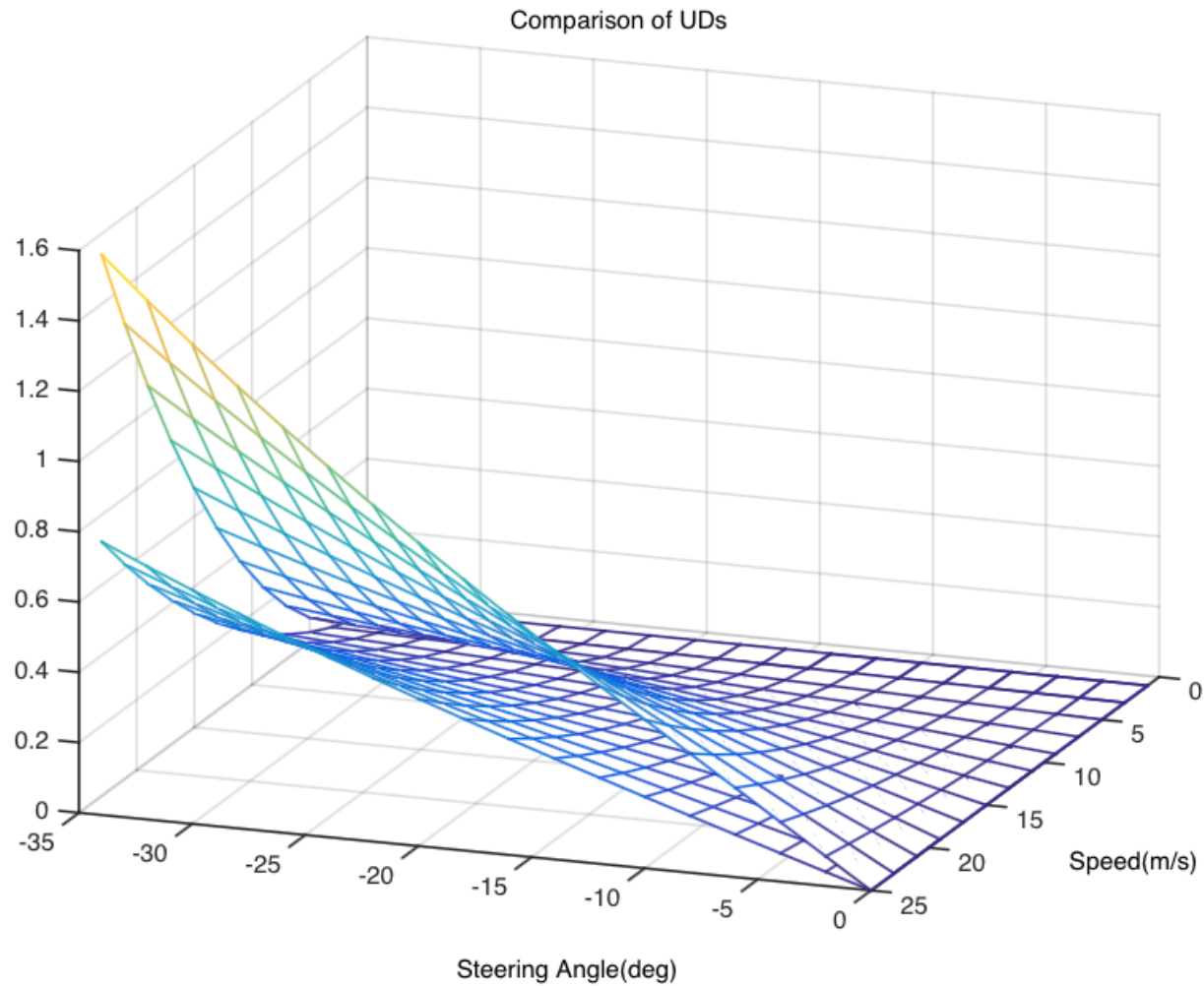


Simulation

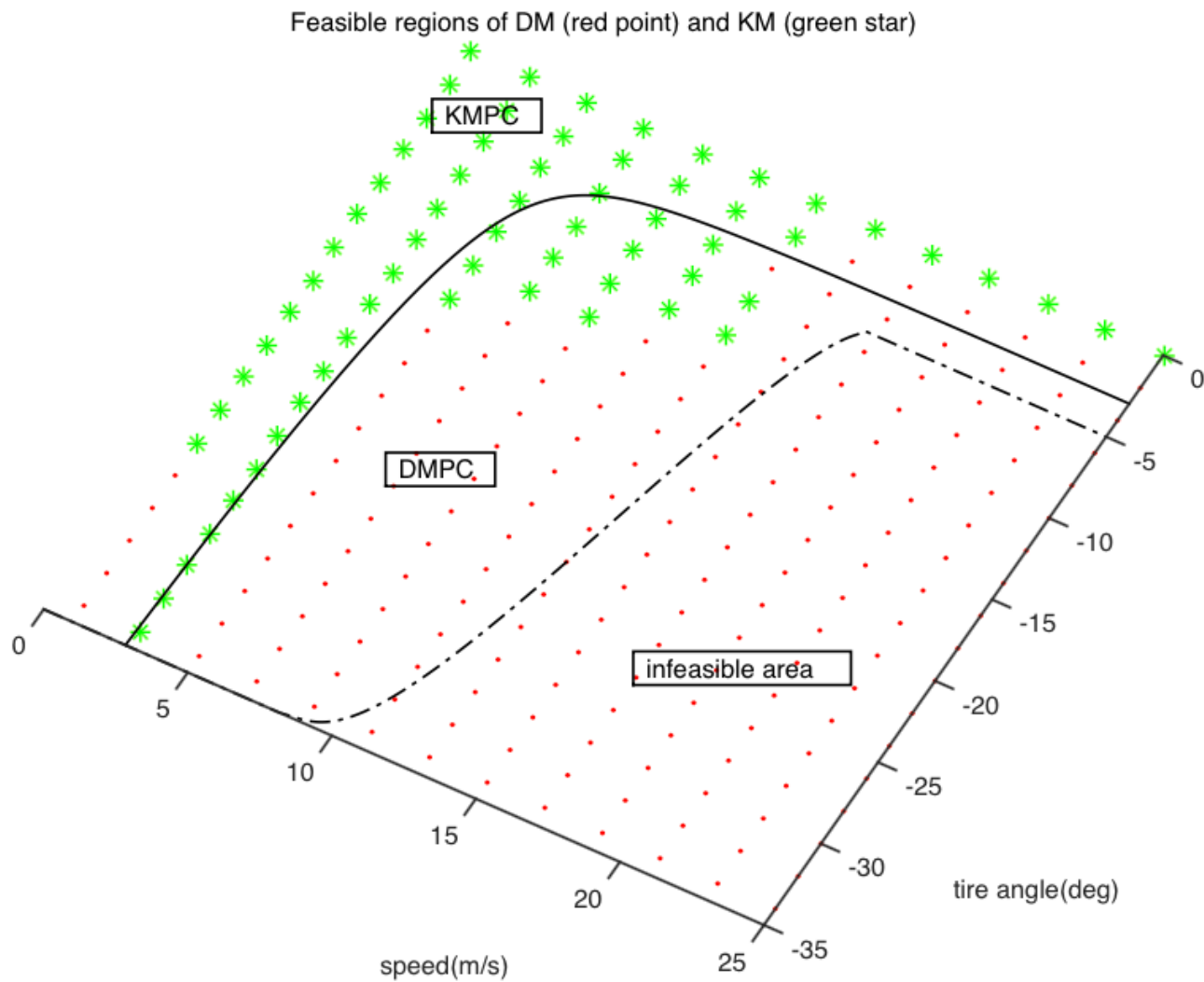




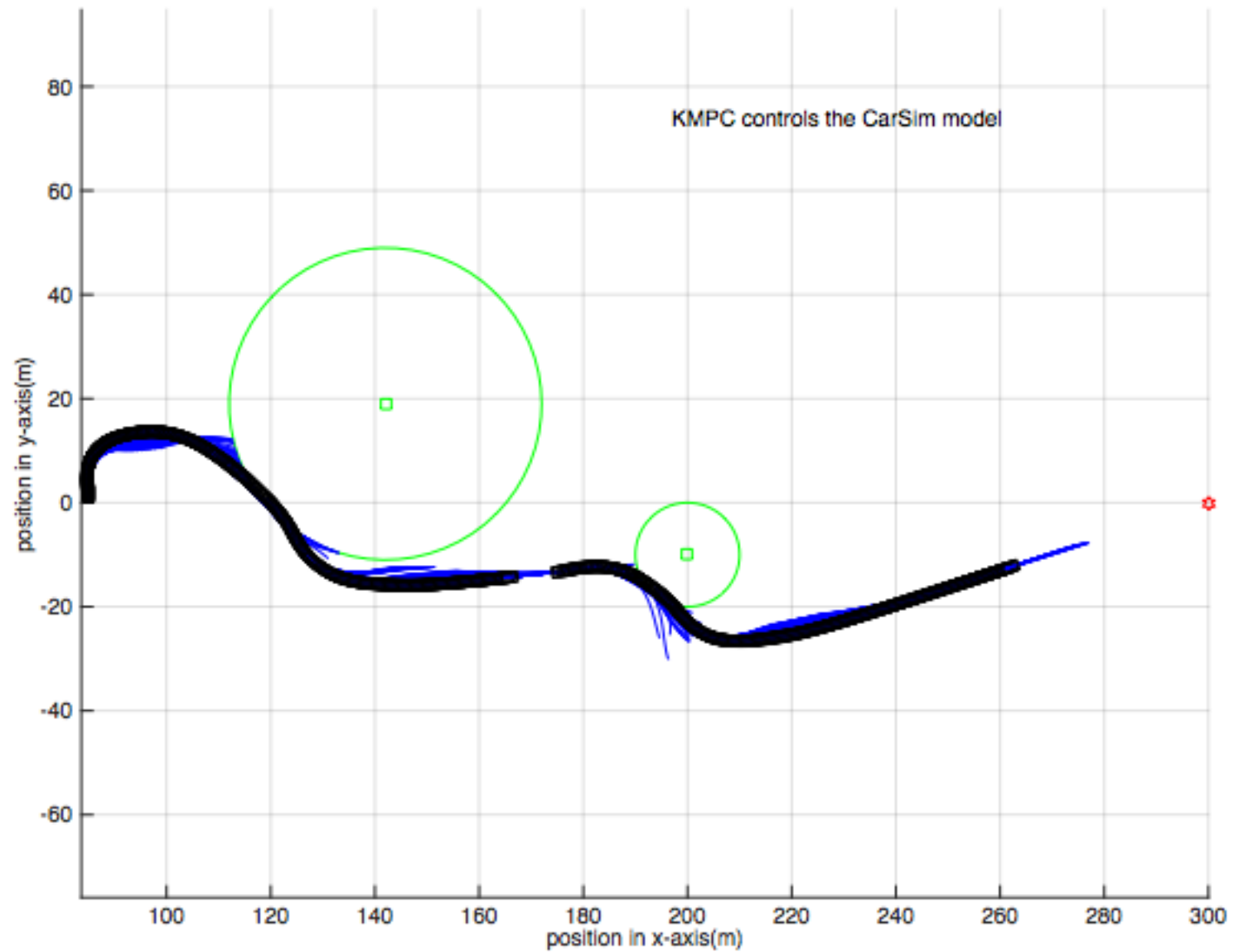
Hybrid MPC Design



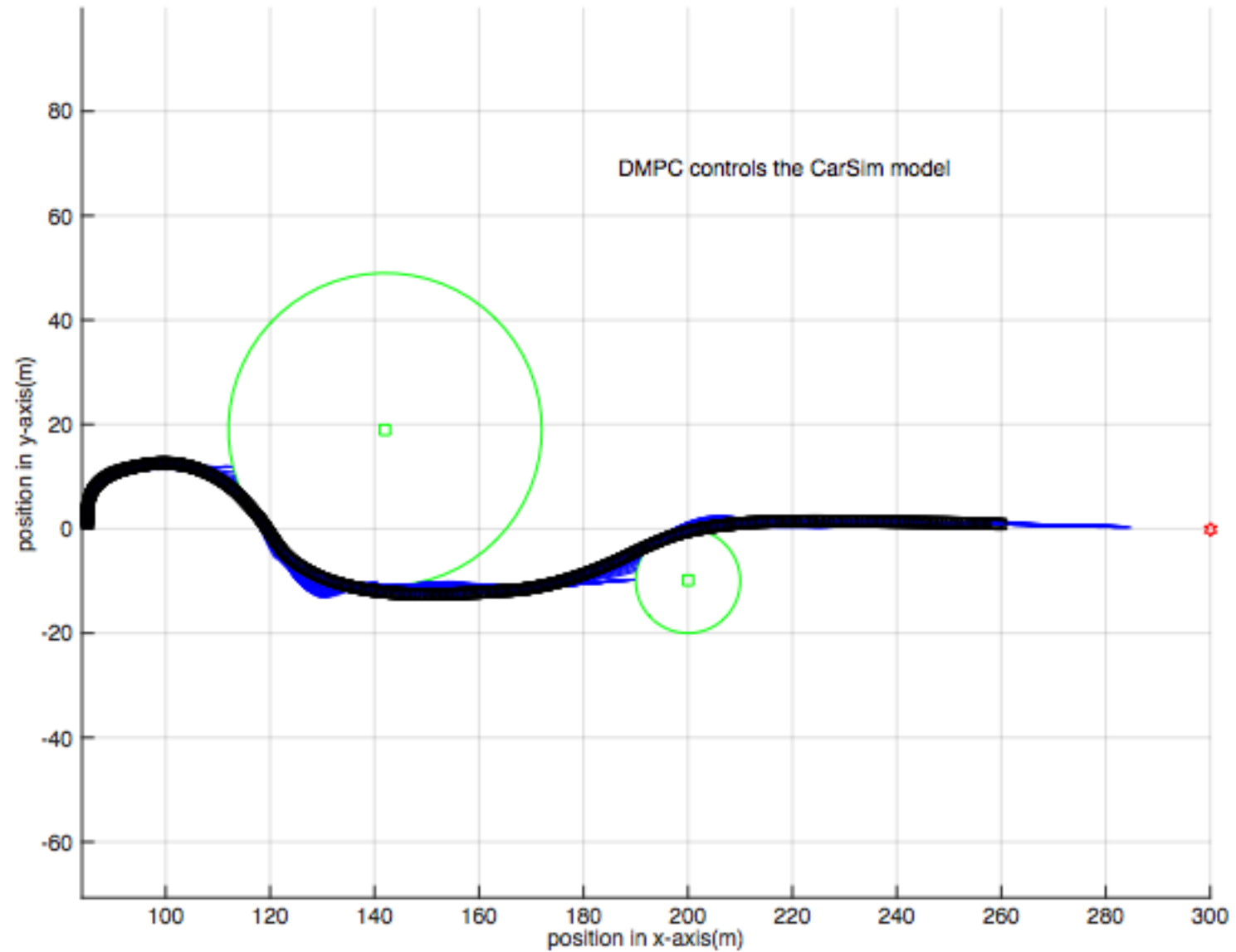
Hybrid MPC Design



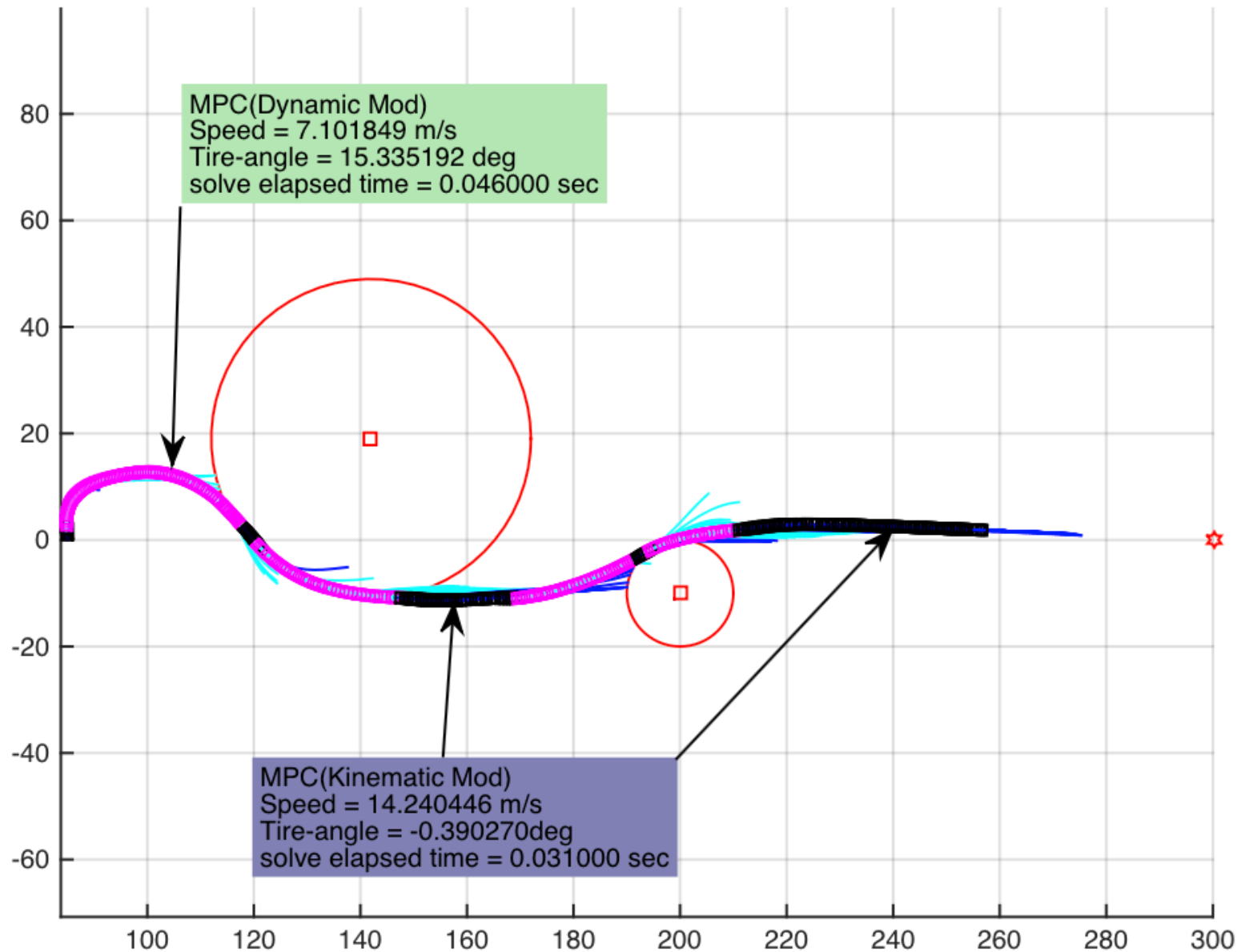
Simulation Results



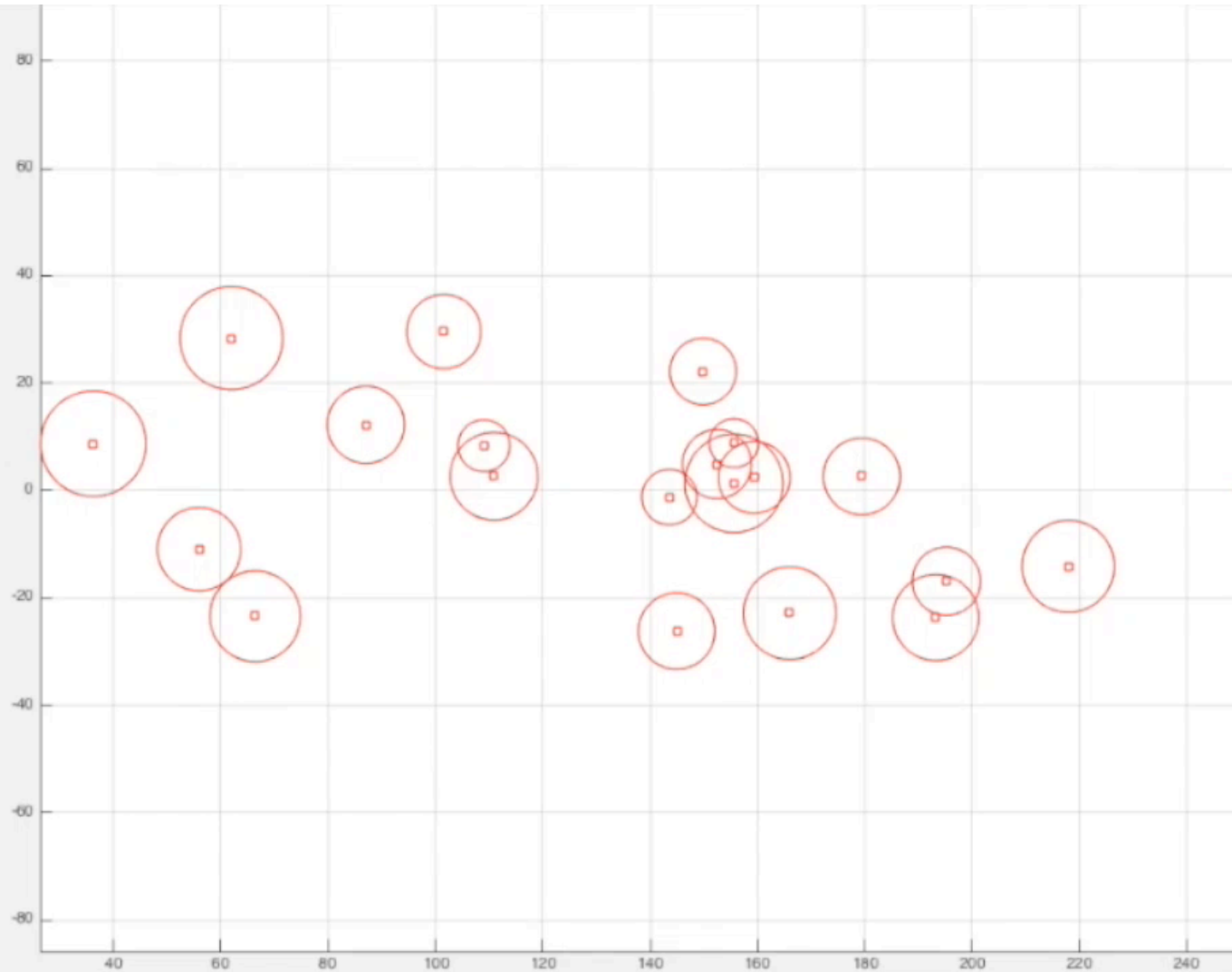
Simulation Result



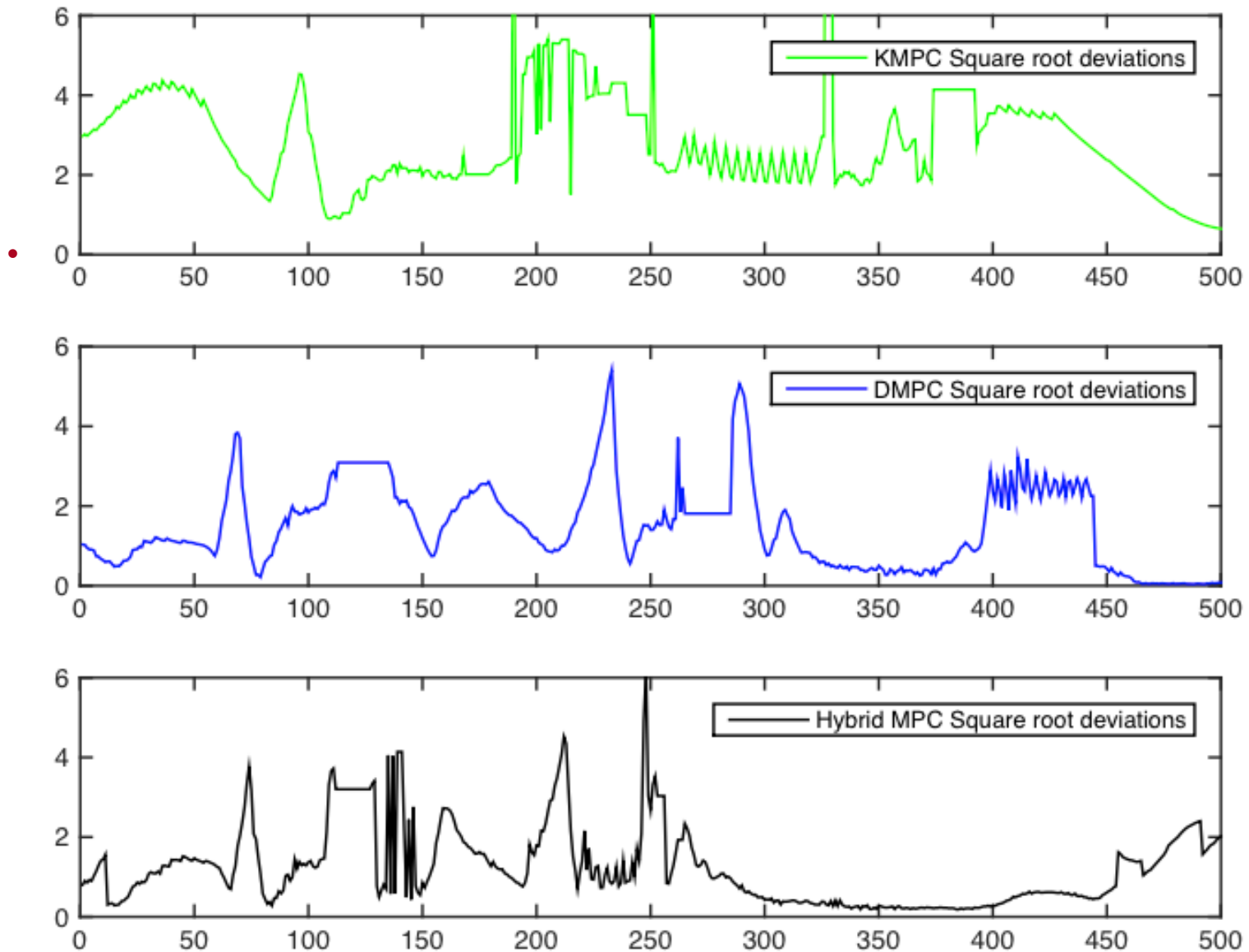
Simulation Result



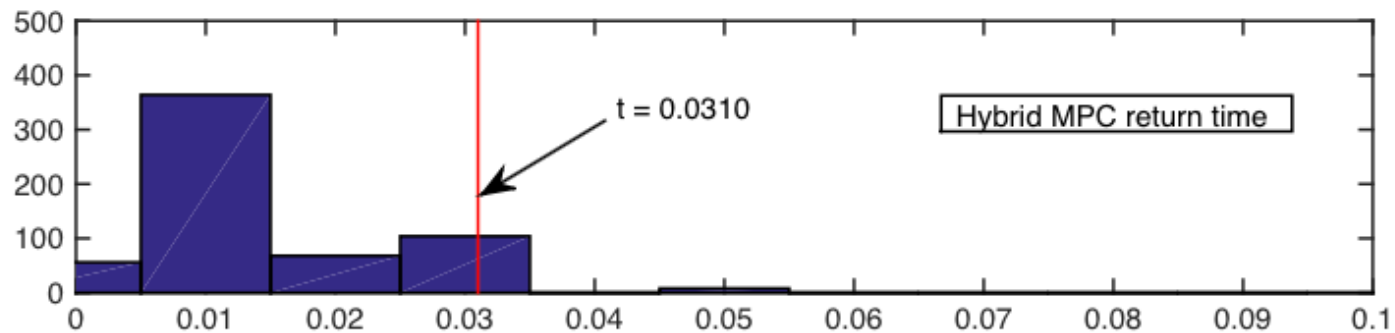
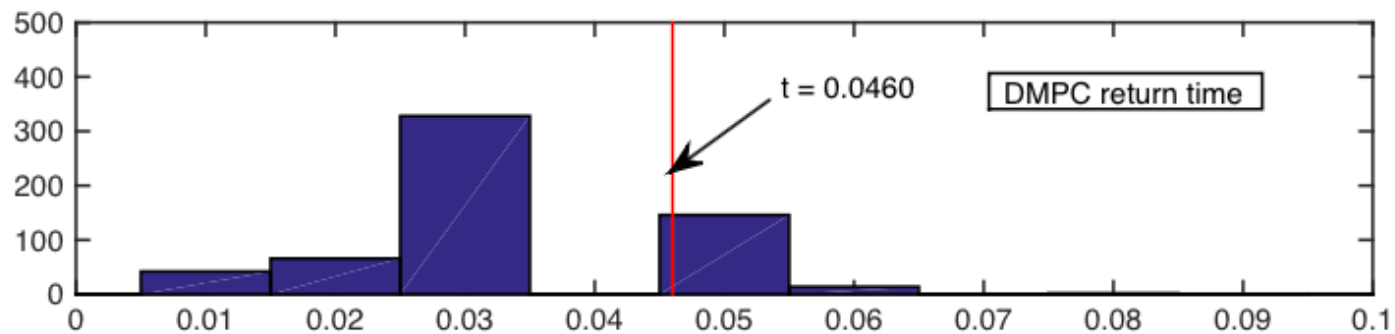
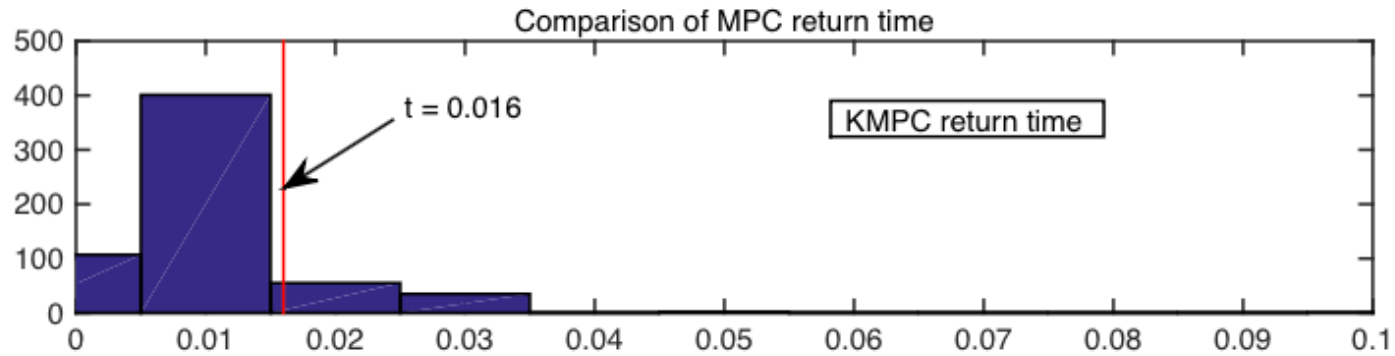
What about lots of obstacles?



Simulation Result

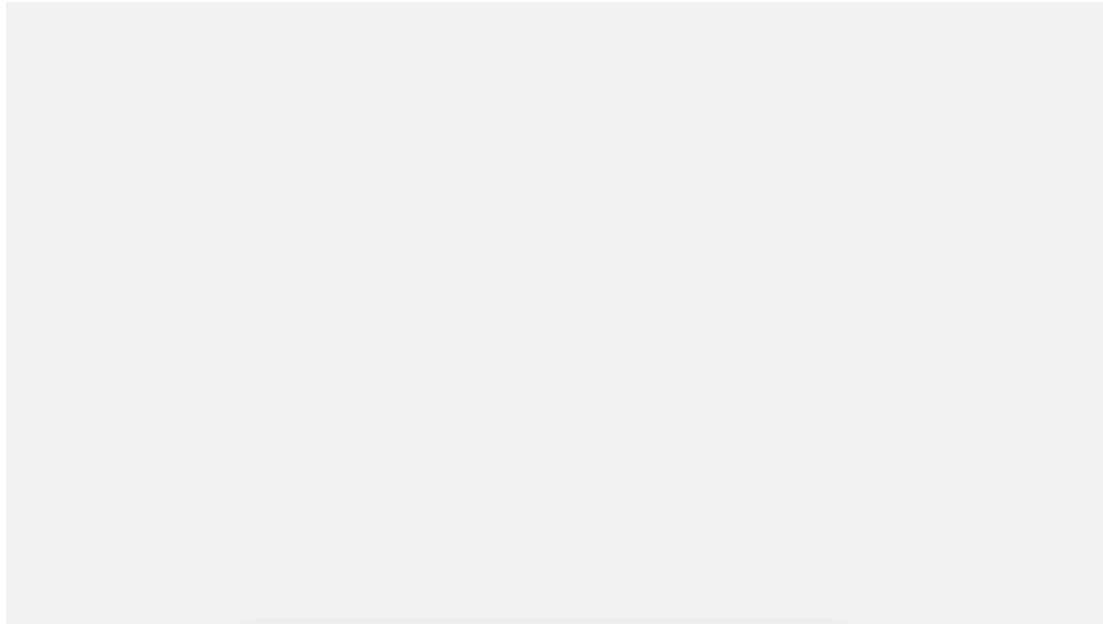


Simulation Result

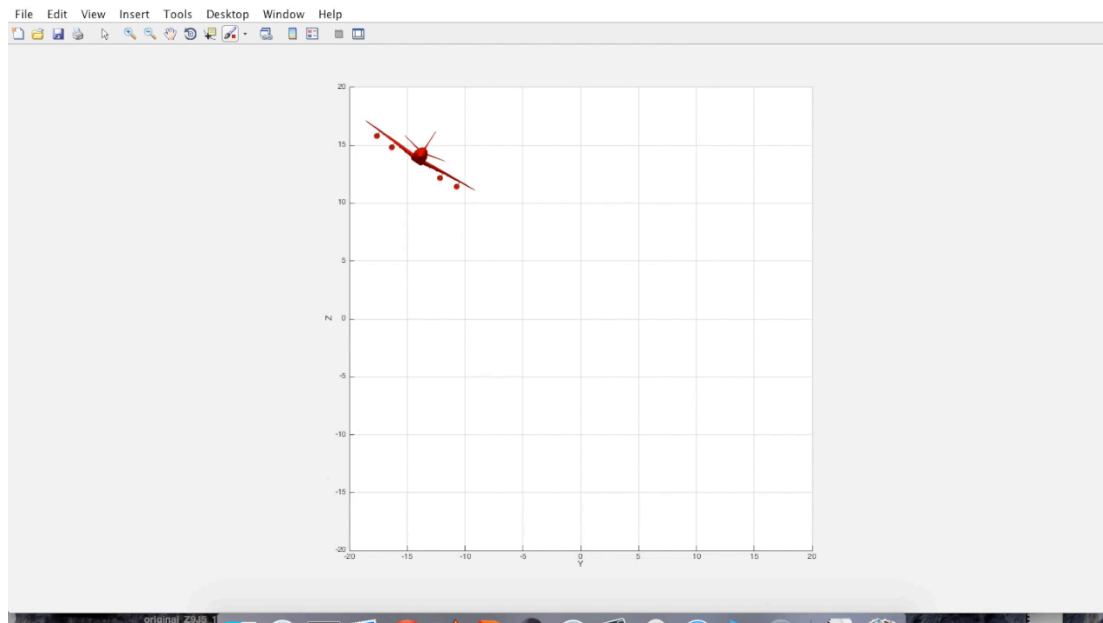


What if I don't have a car?

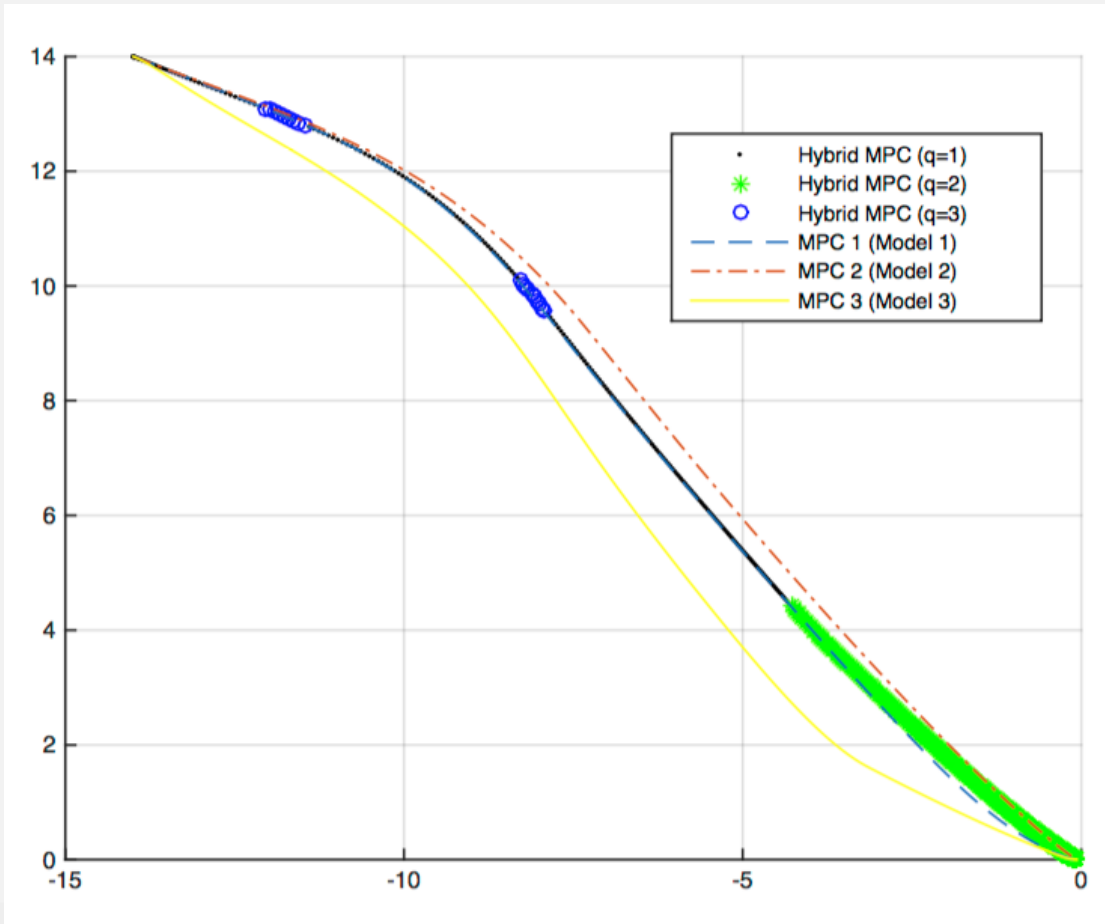
Unswitched



Switched



Application of switching MPC



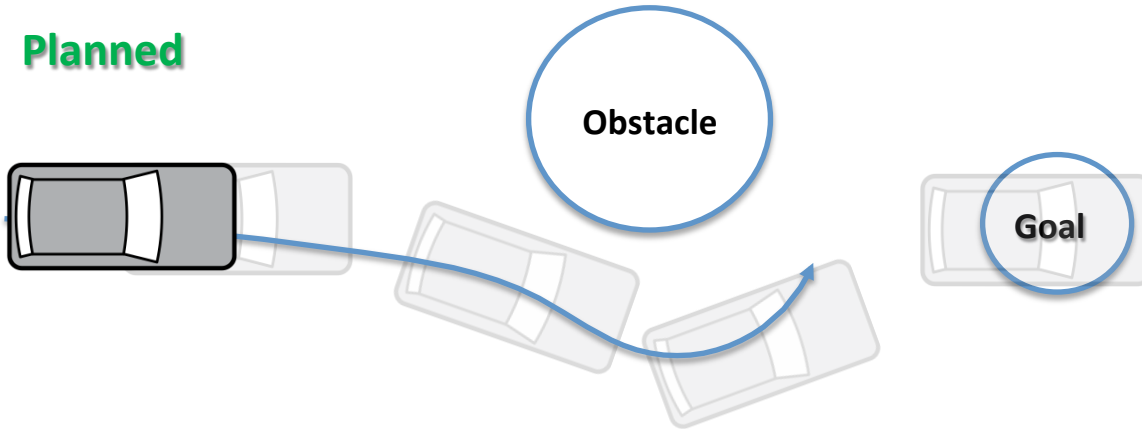
Application to Quadrotor Obstacle Avoidance

CACPS Current efforts: Design with time-awareness

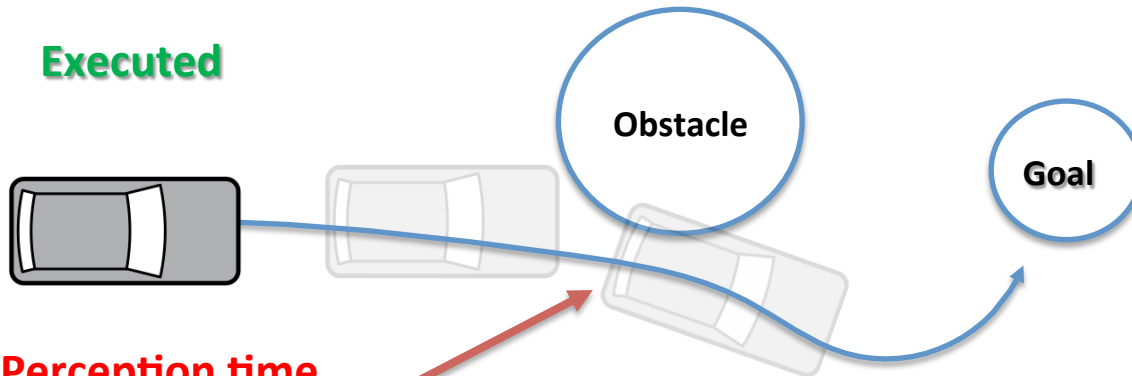
1. Overcoming time-awareness challenges

- Ensure timeliness for critical applications
- Reliability for verification and composition

Planned



Executed



Perception time,
Computation time lag

Approach

1. **Knowledge** of computational components time performance
2. **Design** to achieve computational constraints
3. **Self adaptation** to computational load, by trading accuracy with specification through an hybrid approach.

CACPS Current efforts: Design with time-awareness

2. Quantifying computational burden

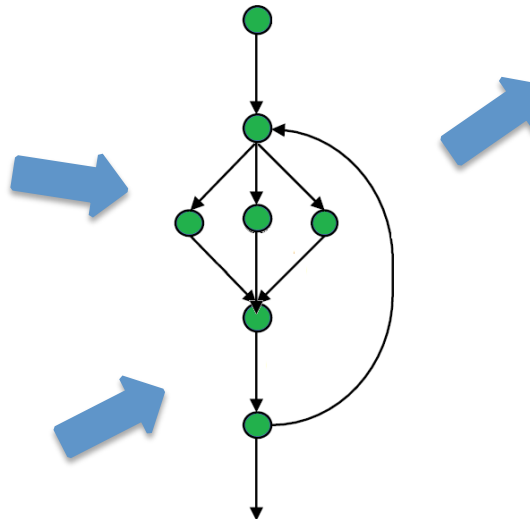
➤ Quantifying time performance

- Hardware: system's architecture
- Software
- Network induced

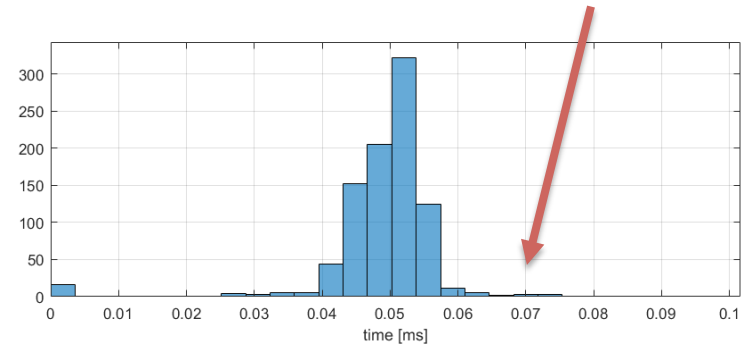
➤ Approach

```
1 function [bboxes, scores] = detectPeopleACF(I, varargin)
2 [params, detector] = parseInputs(I, varargin{:});
3 Iroi = vision.internal.detector.cropImageIfRequested(I, params.ROI, params.ROIType);
4 [bboxes, scores] = vision.internal.acf.detect(Iroi, detector, params);
5 bboxes = round(bboxes);
6 bboxes = clipBox(bboxes, size(Iroi));
7 bboxes(:,1:2) = vision.internal.detector.addOffsetForROI(bboxes(:,1:2),
8 if params.SelectStrongest
9 [bboxes, scores] = selectStrongestBox(bboxes, scores, ...
10 'RatioType', 'Min', 'OverlapThreshold', 0.65);
11 end
12 %-----
13 function m = getModel(params)
14 persistent model id name
15 if isempty(model)
16 % first time initialization
17 model = cell(1,2);
18 [data, id, name] = loadModel(params.Model);
19 model{id} = data;
20 elseif strcmp(params.Model, name)
21 % model changed
22 [name, id] = getModelNameAndID(params.Model);
23 if isempty(model{id})
24 % model not loaded yet. load it now.
25 model{id} = loadModel(params.Model);
```

Network



WCET / WCRT

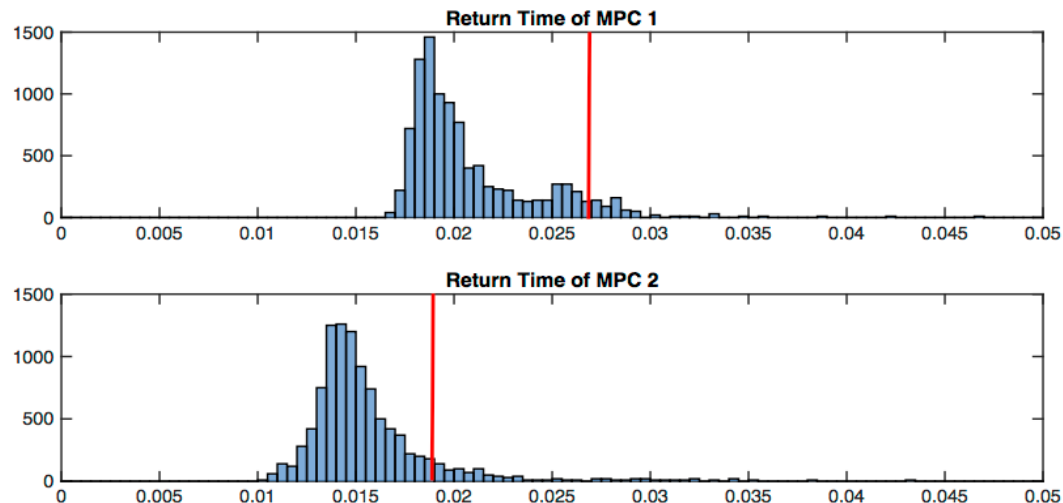


CACPS Current efforts: Design with time-awareness

2. Quantifying computational burden

➤ Quantifying time performance

- Characterizing return time and defining boundaries for algorithms



➤ Reactive and time-adaptive controller design

Incorporating time characterization in the design process

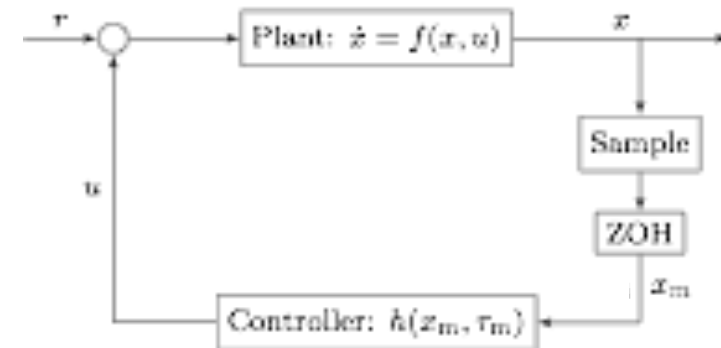
CACPS Current efforts: CA Predictive Control for CPS

3. Predictive control that is robust to computation time uncertainty

➤ **Computation time** T_{comp} is uncertain due to

- Hardware variability
- Delays and data losses
- Computational complexity

➤ **Approach**

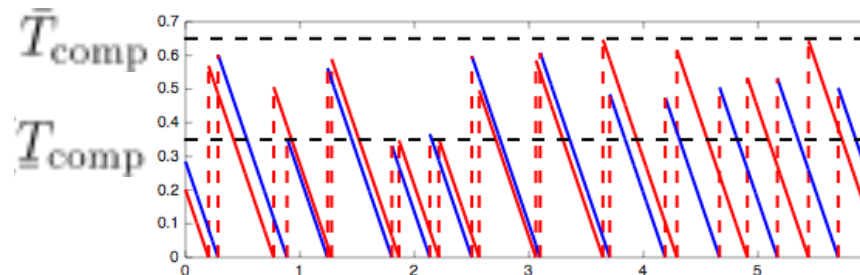


Hybrid system modeling

At sampling instants, T_{comp} is updated via the *difference inclusion*:

$$T_{comp}^+ \in [T_{comp}, \bar{T}_{comp}]$$

while in between sampling events, it decreases to zero linearly.



Design control algorithm to compensate for all possible values of

T_{comp}

Accomplishments during first year of performance and ongoing efforts

Summary of milestones reached (first period of performance):

- Formulation of model of computationally-aware CPS (CACPS)
- Derivation of uncertainty divergence (UD)
- Application to ground and aerial vehicle applications

Ongoing efforts:

- Overcoming time-awareness challenges
- Quantifying computational burden
- Predictive control that is robust to computation time uncertainty

CACPS Student Team:

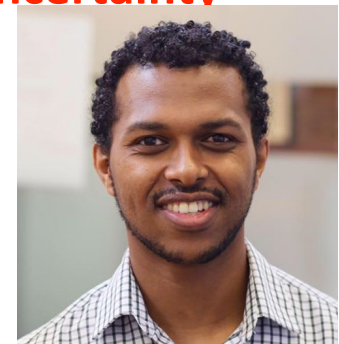
+ 2 Undergrad students
+ 2 High School inters



Berk Altin
(postdoc, UCSC)



Nathalie Risso
(grad, UofA)



Yegeta Zeleke
(grad, UCSC)

CPS: Synergy: Collaborative Research Computationally Aware Cyber-Physical Systems

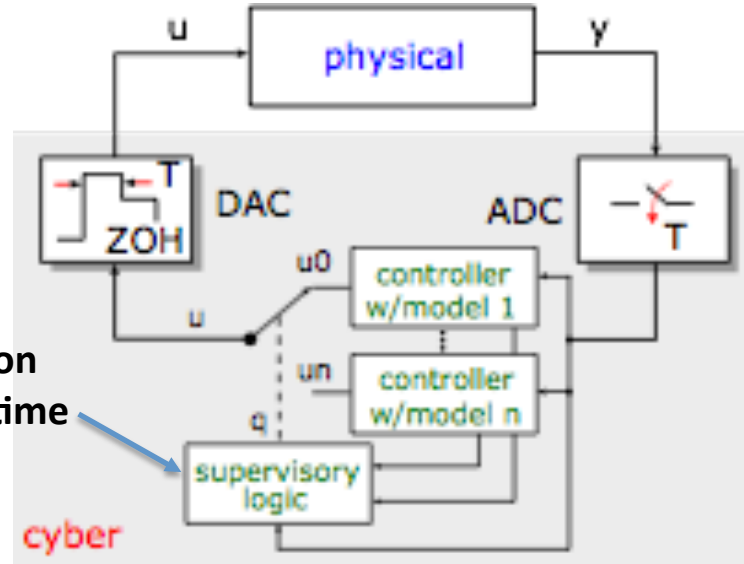
Ricardo Sanfelice (UC Santa Cruz NSF 1544396) , Jonathan Sprinkle (University of Arizona, NSF 1544395)

Challenge:

- Include computation time information at runtime as part of the controller's design for switching criteria
- Do not give up accuracy to improve computation time

Solution: logic based on computation time

- Metric for switching criteria
- Basis in accumulated error for predictive controller
- Hybrid control approach
- Trade off between time to compute and accuracy

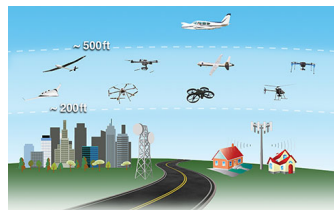


Scientific Impact:

- More specific controllers made available, based on available computation power
- Increased confidence in runtime controller, if switched

Broader Impact:

- Less conservative controllers mean closer flight patterns
- Benefits to ground and aerial vehicle control systems
- Applications to undergraduate team test beds
- ~20% reduction in average return time for dynamical models without compromising stability



UAVs in the National Air Space
(courtesy NASA)



Self-driving Cars

Thank you for your attention!

Back up slides

Sufficient conditions for asymptotically ultimately bounded

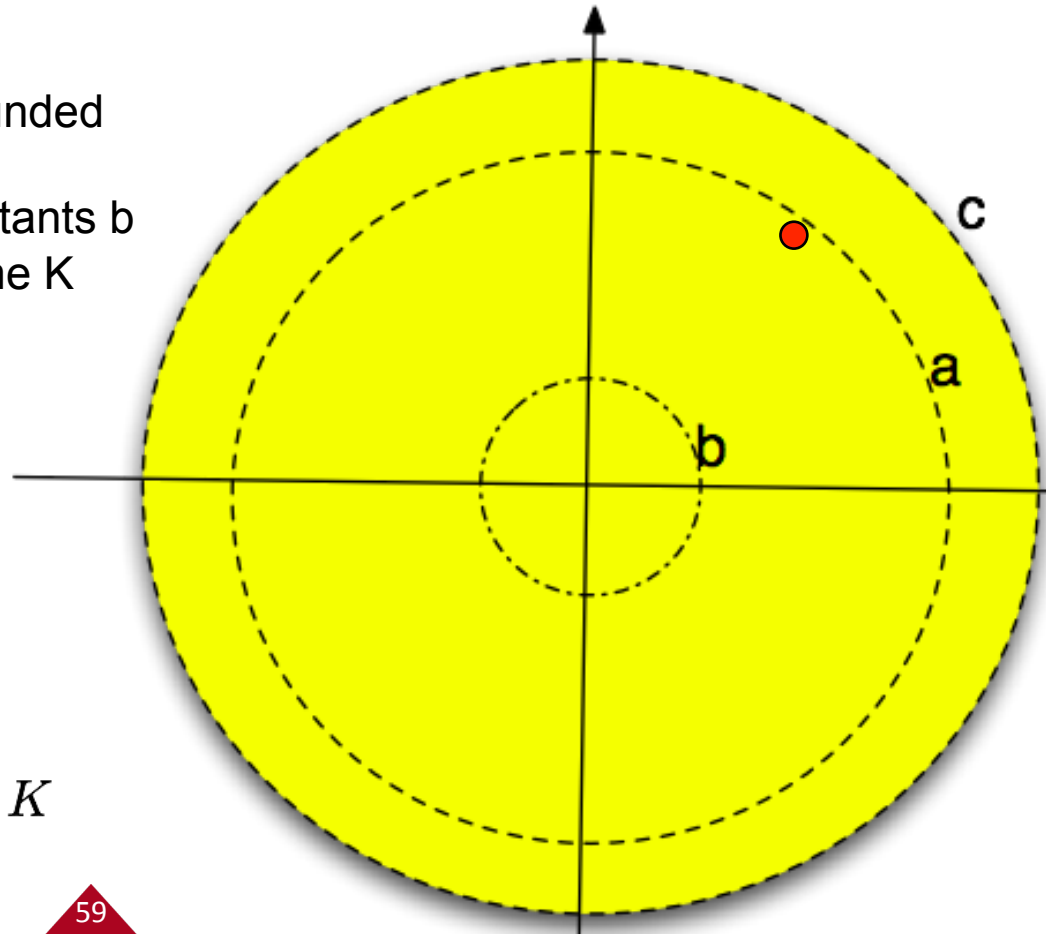
Why consider boundedness?

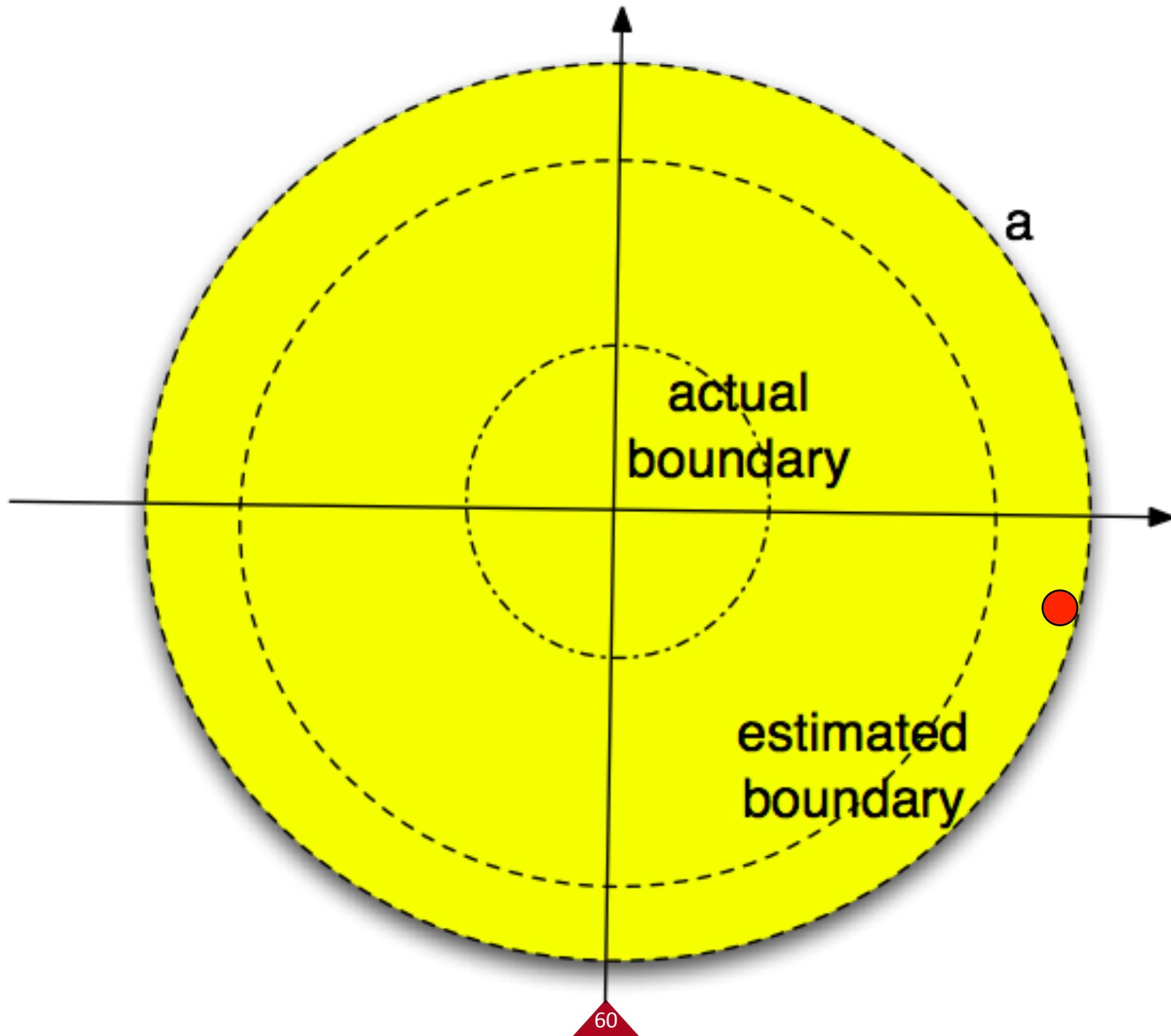
- (1) Minimizing UD does not automatically ensure stability.
- (2) An overall system consisting of stable subsystems is not necessarily stable.

These sufficient conditions are switching constraints.

A system is asymptotically ultimately bounded if the system evolves asymptotically to a bounded set, i.e., there are positive constants b and c , and for every $0 < a < c$, there is a time K such that

$$\|\xi_0 - \xi_r\| \leq a \Rightarrow \|\xi_k - \xi_r\| \leq b, \forall k > K$$

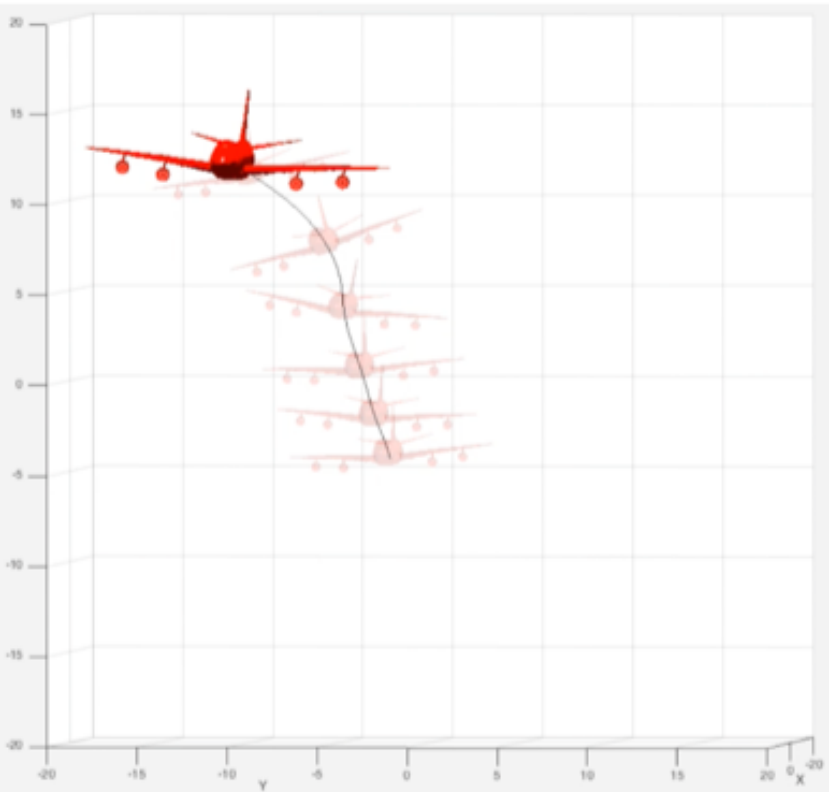




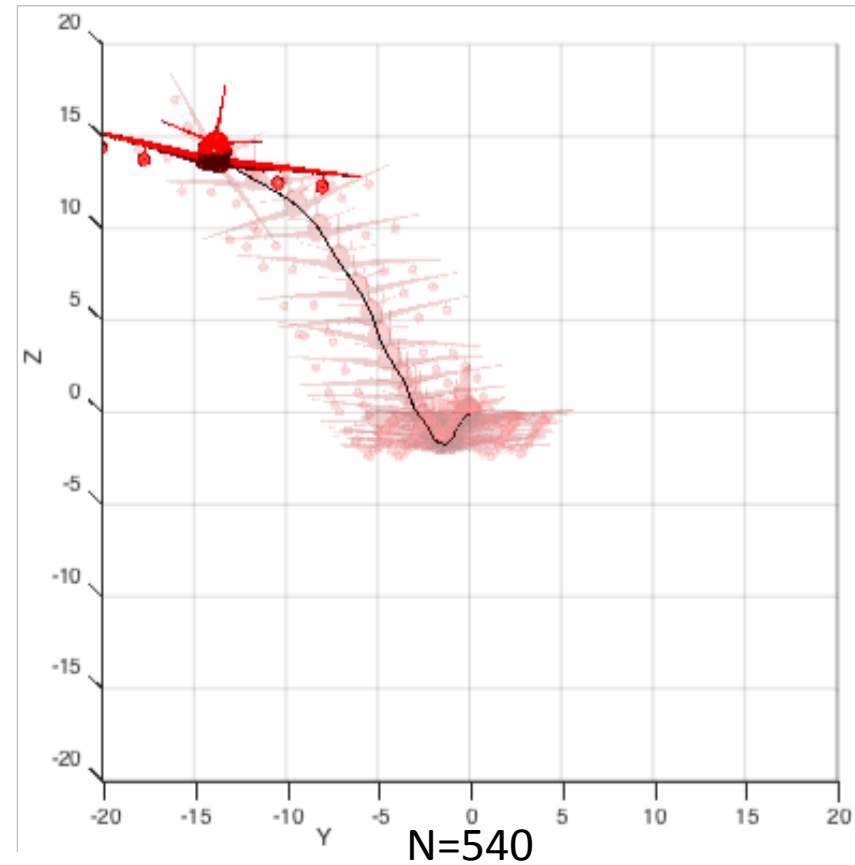
If the following assumptions are satisfied by a candidate CPS, then this work can be applied:

- The target state $\hat{\xi}_r$ for the system $\hat{f}(\hat{\xi}, \hat{u})$ and the predictive model $\hat{f}^q(\hat{\xi}, \hat{u})$, $q \in \mathbb{Q}$ is an equilibrium; namely, $\hat{\xi}_r \in \mathcal{X}$, and $\hat{f}(\hat{\xi}_r, \kappa^q(\hat{\xi}_r, 0)) = \hat{\xi}_r$ and $\hat{f}^q(\hat{\xi}_r, \kappa^q(\hat{\xi}_r, 0)) = \hat{\xi}_r$.
- The horizon is long enough; $\forall q, q' \in \mathbb{Q}$, $S_{\max}^q + S_{\max}^{q'} \leq N$. This prevents the depletion of control inputs while waiting for MPC return.
- In order to ensure that UD is bounded, $\frac{\partial \hat{f}^q}{\partial \hat{\xi}}$, $\frac{\partial \hat{f}^q}{\partial \hat{u}}$, $\nabla \kappa^q$, and $\hat{\Gamma}^q$ are required to be bounded $\forall \hat{\xi} \in \mathcal{X}$, $\forall \hat{u} \in \mathcal{U}$ and $\forall q \in \mathbb{Q}$.
- Implicit MPC control through $\kappa^q(\cdot, \cdot)$ will always drive the plant within the feasible state space \mathcal{X} , and $\kappa^q(\hat{\xi}, i) \in \mathcal{U}$, $\forall \hat{\xi} \in \mathcal{X}$ and $\forall i \in \{0, 1, \dots, N-1\}$.

Long horizon examples



N=36



N=540

This permits the use of a “dwell-time” approach to determination of stability.

It doesn't always work.

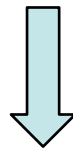
Reference Plant: The reference plant \hat{f} is formulated by

$$\begin{aligned}\ddot{x} &= -g \sin(\theta) + u_1 \cos(\theta)/m - u_2 \sin(\theta)/m - c\dot{x}/m \\ \ddot{y} &= g(\cos(\theta) - 1) + u_1 \sin(\theta)/m + u_2 \cos(\theta)/m - c\dot{y}/m \\ \ddot{\theta} &= ru_1/J\end{aligned}$$

remove the last term

Predictive Model 1 ($q = 1$): The first predictive model is derived from the reference plant by removing the last terms:

$$\begin{aligned}\ddot{x} &= -g \sin(\theta) + u_1 \cos(\theta)/m - u_2 \sin(\theta)/m \\ \ddot{y} &= g(\cos(\theta) - 1) + u_1 \sin(\theta)/m + u_2 \cos(\theta)/m \\ \ddot{\theta} &= ru_1/J\end{aligned}$$



linearized around 0

Predictive Model 2 ($q = 2$): The second predictive model is a linearized version of the reference model about the origin:

$$\begin{aligned}\ddot{x} &= u_1/m + (-g - u_2/m)\theta \\ \ddot{y} &= u_2/m + u_1/m\theta \\ \ddot{\theta} &= ru_1/J\end{aligned}$$

linearized around 0, but multiplied by rho

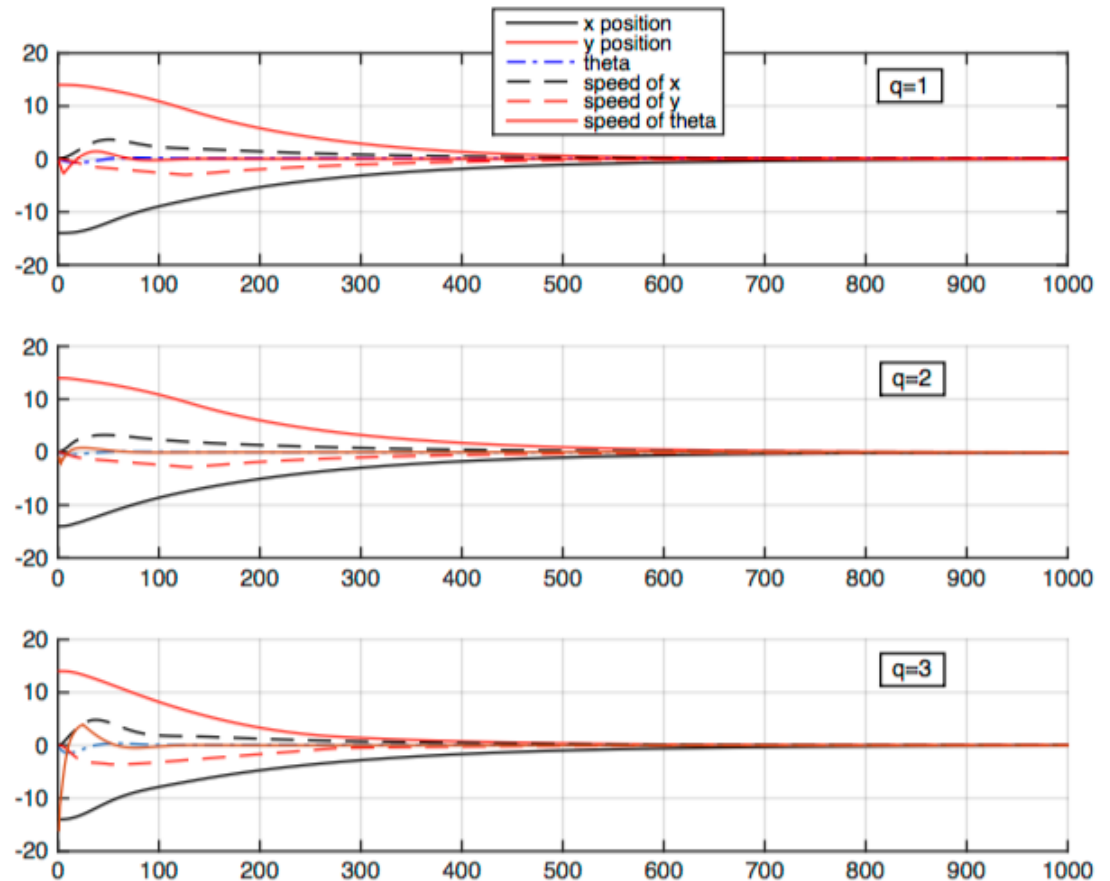
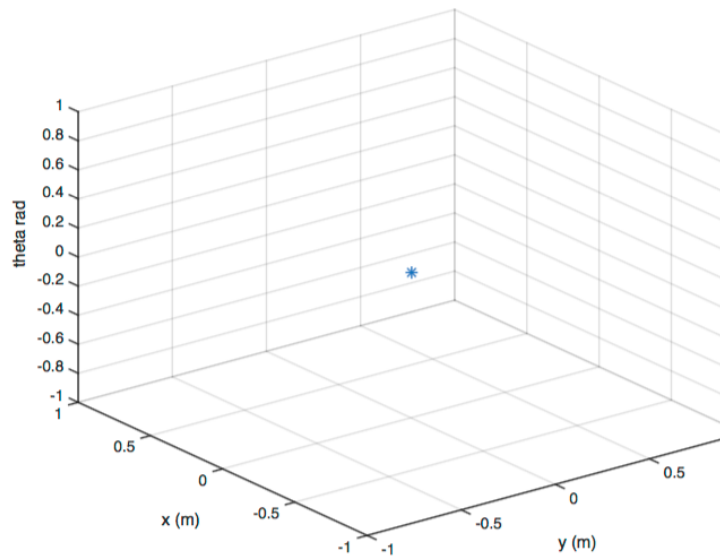
Predictive Model 3 ($q = 3$): The third predictive model is a linearized version of the reference plant about the origin, but the equations are then multiplied by ρ :

$$\begin{aligned}\ddot{x} &= (u_1/m + (-g - u_2/m) - c\dot{x}/m\theta) \rho \\ \ddot{y} &= (u_2/m + u_1/m\theta - c\dot{y}/m) \rho \\ \ddot{\theta} &= (ru_1/J) \rho\end{aligned}$$

where $\rho = 0.95$ is deliberately selected to adjust the model's accuracy.

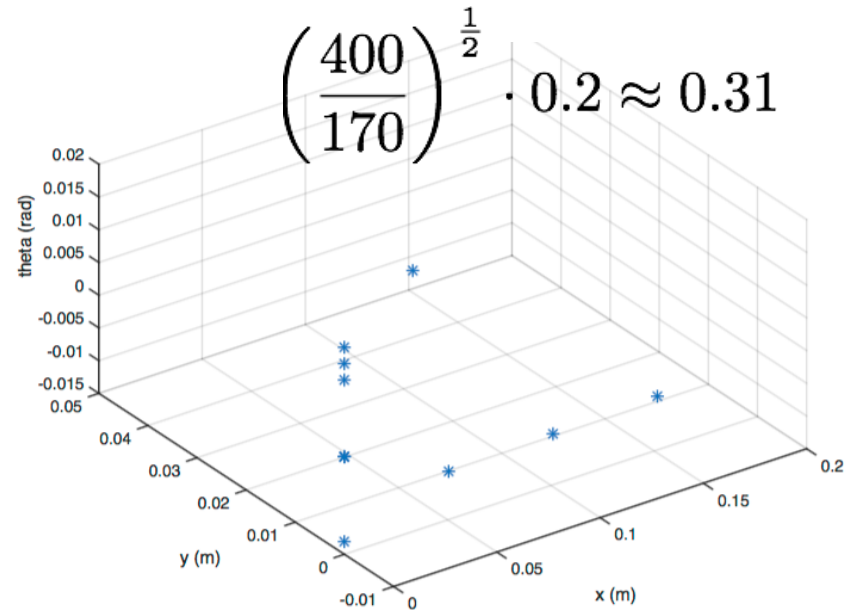
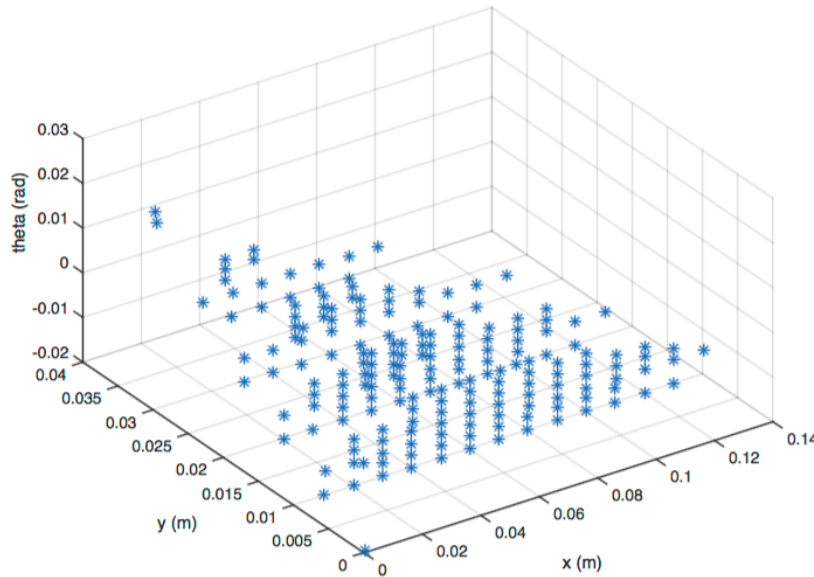
Each model converges

$$\begin{aligned}
 R^{\text{MPC}} &= \max \left(\max_q \left(\frac{b^q}{a^q} \right)^{\frac{1}{\sigma}}, \max_q \left(\frac{b^q}{a^q} \frac{b^{\bar{q}}}{a^{\bar{q}}} \gamma \right)^{\frac{1}{\sigma}} \right) \underline{\mathbb{R}}^{q+} \\
 &= \max_q \left(\left(\frac{350}{250} \right)^{\frac{1}{2}}, \left(\frac{350}{250} \frac{350}{250} 0.95 \right)^{\frac{1}{2}} \right) 0 \\
 &= 0
 \end{aligned}$$



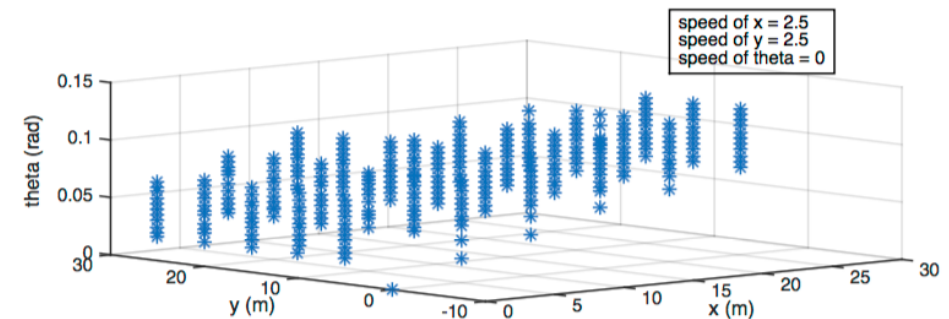
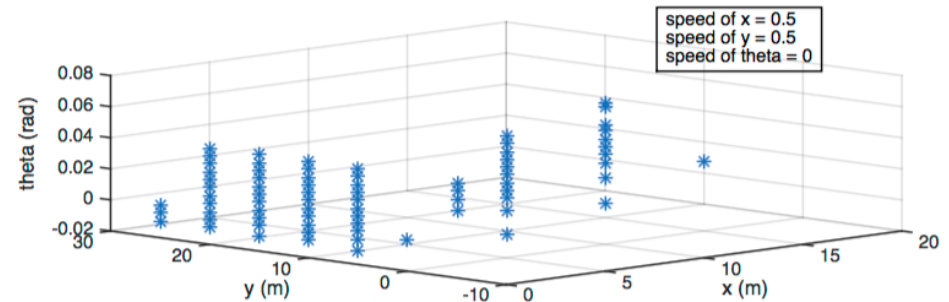
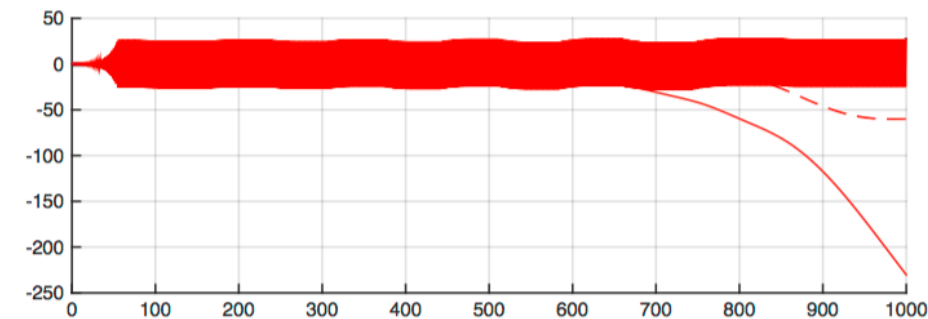
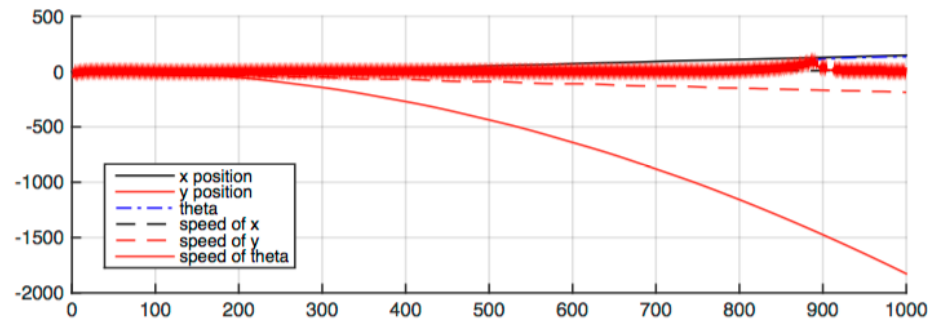
Numerical results of the superlevel sets of all three models suggest that each of these three MPCs can bound the aircraft to the origin, or at least a small ball containing the origin with a radius smaller than the discrete spatial steps.

Make model-3 less accurate



As d decreases: although the superset $L_d(\Delta V^{q=3}(\hat{\xi}))$ has fewer points, the points in these two figures cover a similar area, and the set $L_d(\Delta V^{q=3}(\hat{\xi}))$ will not shrink in an obvious way by decreasing d .

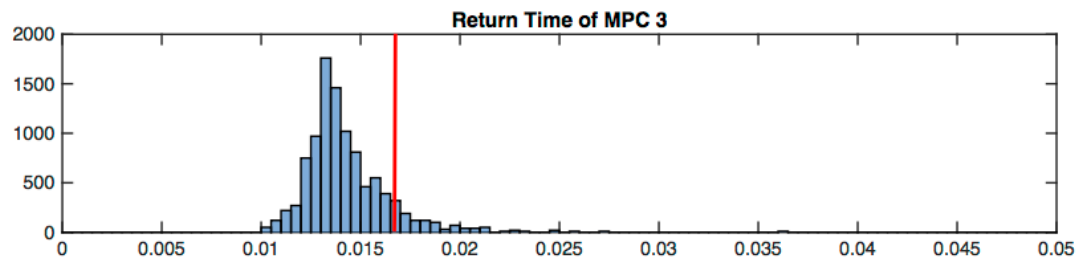
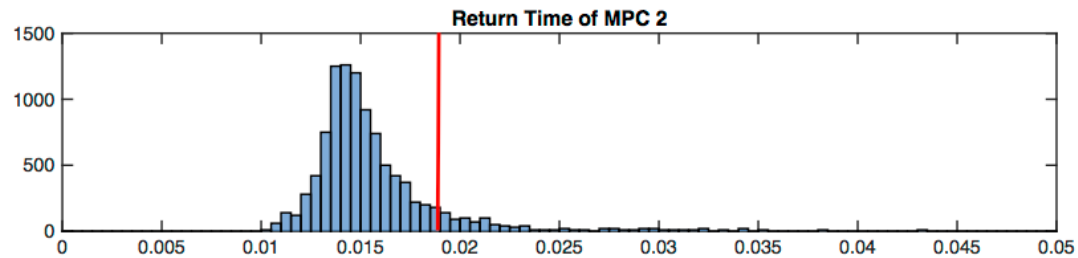
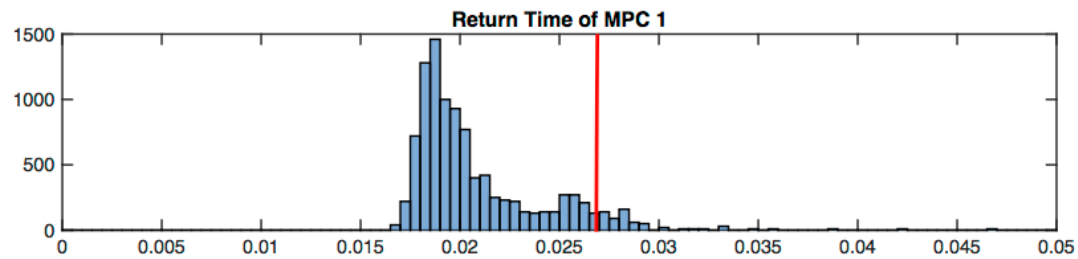
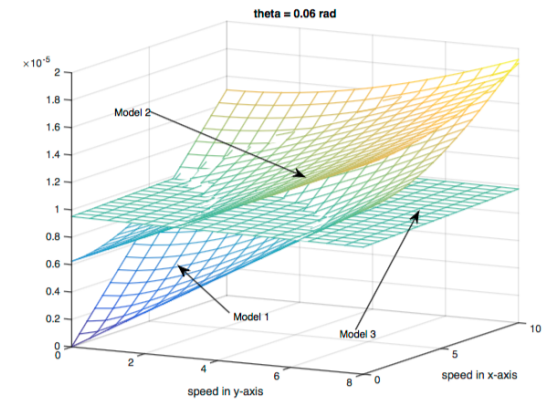
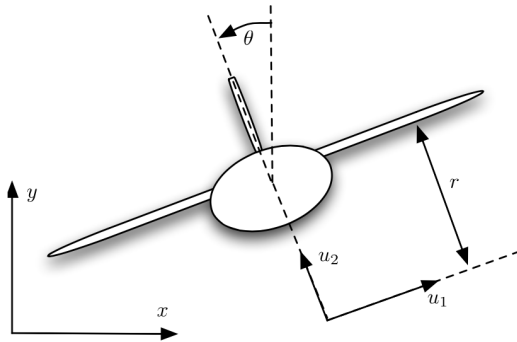
For $\rho=0.5$, and $d=1$



The boundary is not found even when the state reaches $x=30$ m and $y=30$ m.

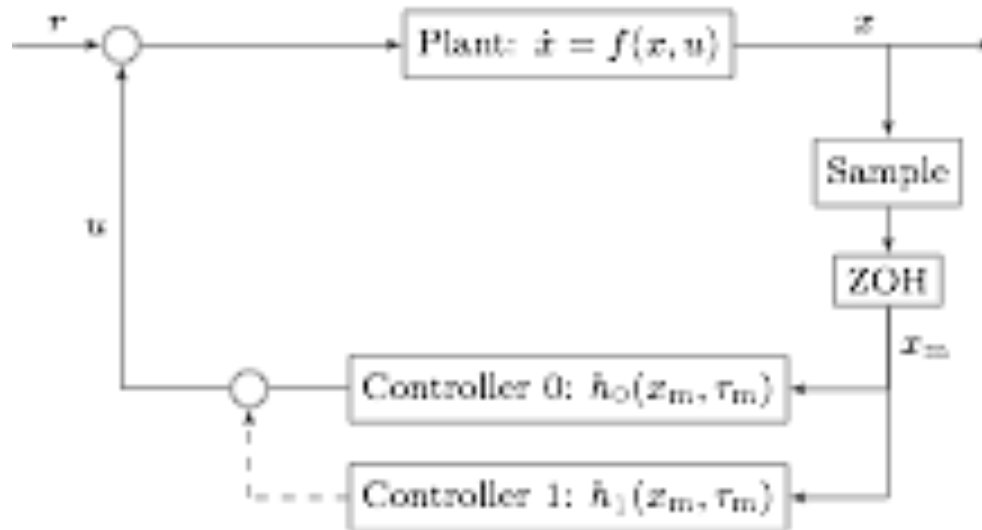
In this situation, the models do not satisfy the necessary conditions in order to use the uncontrollable divergence as the switching metric. It confirms that the approach requires some analysis in order to use the design methodology with confidence.

Return time



CACPS Current efforts: Real Time Predictive Control

Functions $f_q(x, u)$, $q = 0, 1, \dots, N$ of varying fidelity.

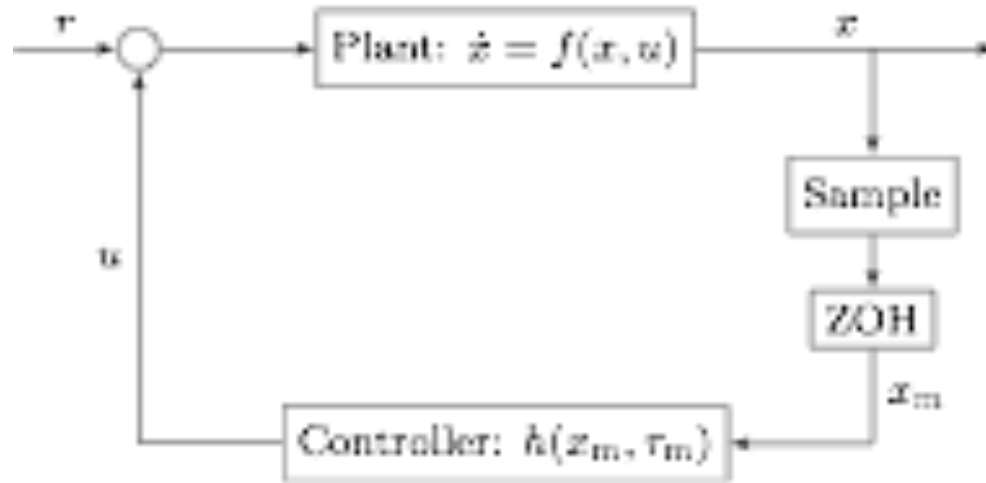


Switching MPC

Find a rule $S : x \rightarrow q$ to switch between prediction models f_q , with respect to a safety-accuracy trade-off.

Hybrid Model of RHC

Basic model with measurement state and timer : τ_m



$h(x_m, \tau_m)$ is the minimizer of

$$J(x_m, u) = \int_0^T \ell(x(s), u(s)) ds + V_f(x(T))$$

with initial condition x_m

Current Work: RHC with varying Horizons

Closed loop state:

$$z \triangleq \underbrace{(x, x_m, \tau_m, T, \tau_c)}$$

(plant state, sample, sampling timer, prediction horizon, controller timer)

Difference *inclusion* model with:

$$\dot{z} = f_{cl}(z), \quad \tau_m \in [0, T_m],$$

$$z^+ \in G_{cl}(z), \quad \tau_m \in \{0\},$$

The *set valued* map G_{cl} allows

- intermittent sampling (due to delays, losses etc.), and
- varying horizons for flexible control to compensate .

Stable provided V is a control Lyapunov function!