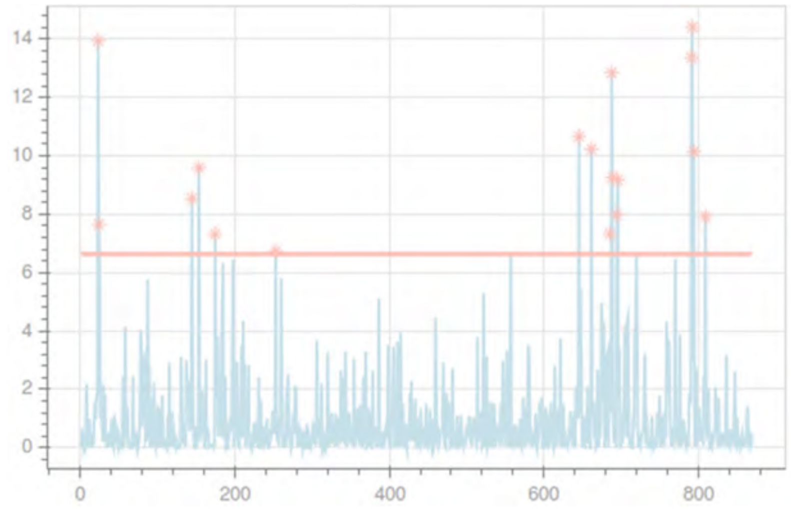
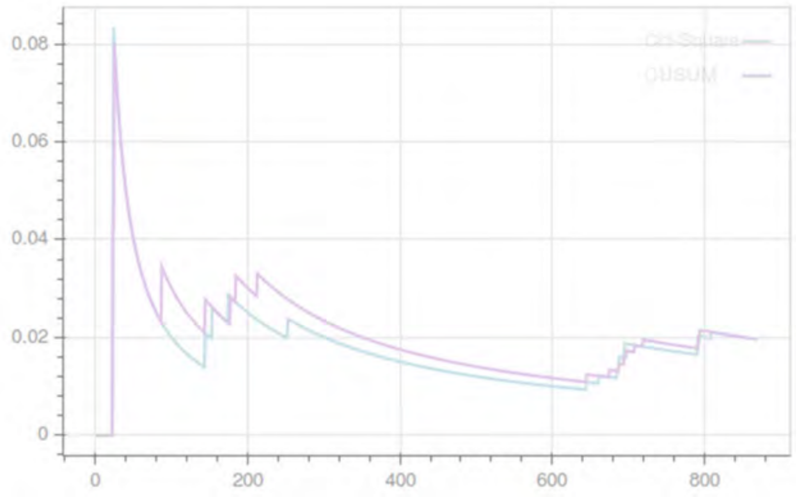
0.506
0.5055
0.505
0.5045
0.504
0.5035
0.503
0.5025
0
200
400
600
800
packets
true value
attacked value
estimate





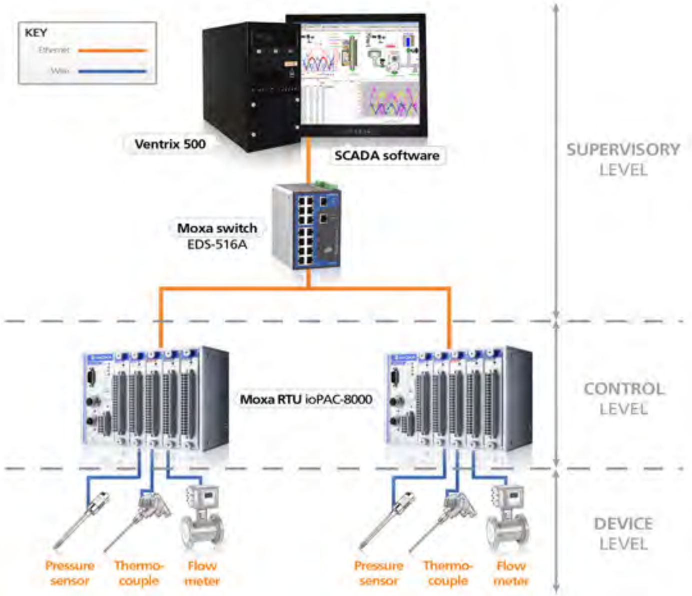
Attack Detection & Mitigation For Resilient Infrastructure & Automation

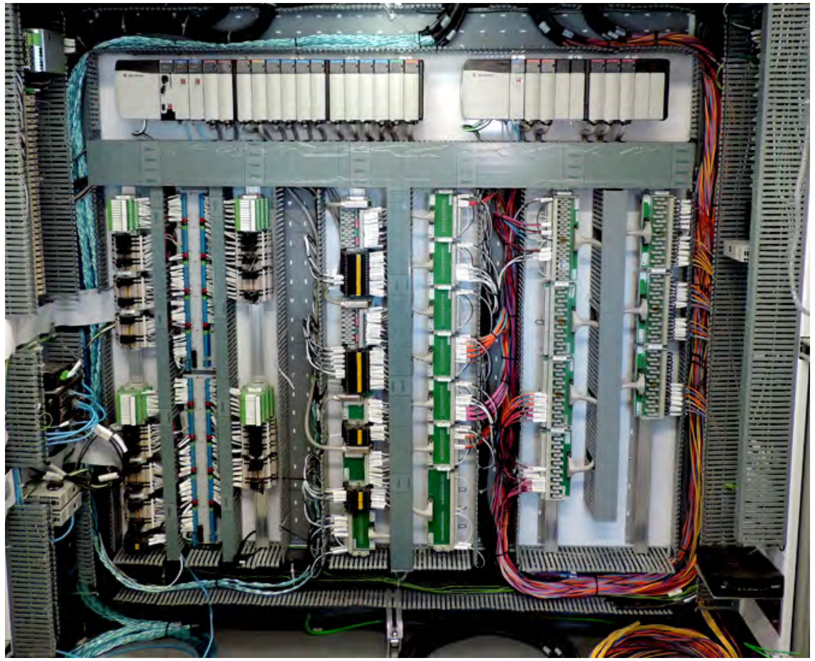
Justin Ruths

THE UNIVERSITY OF TEXAS AT DALLAS
UTD
EST. 1969

RCE Panel, Nov 10th 2022

1. digitization of control systems
2. non-uniform security guidelines
3. real-time requirements of control loops
4. computational capabilities of attackers





new challenges: how do we…

Challeng
es

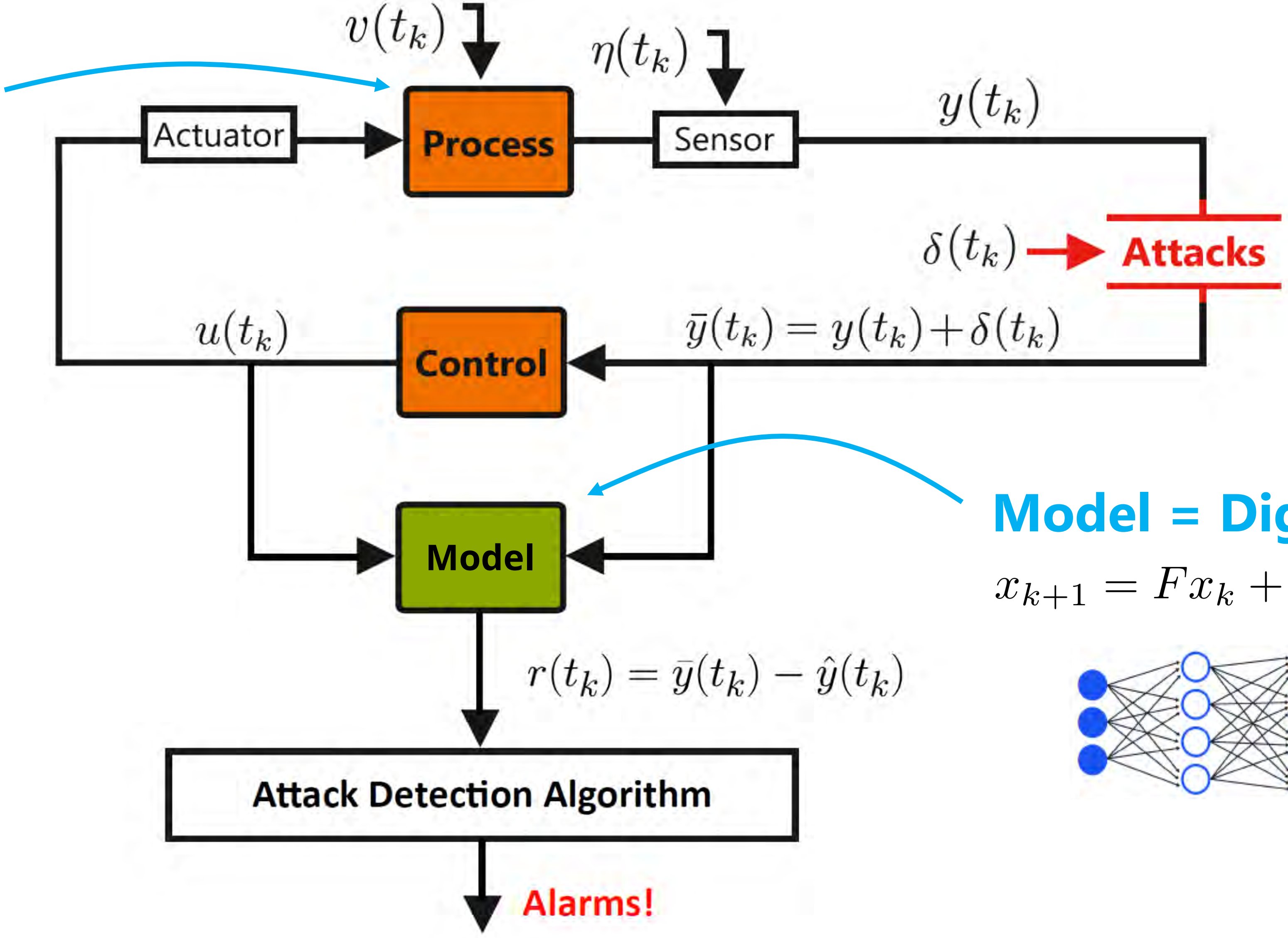
- detect attacks?
- react against attacks?
- design attack-resilient systems?

Attacks are worst case faults!

Technical
Question
s

- tune detector for desired performance?
- quantify the impact that an attacker can have?
- determine potential attacks?
- minimize the impact an attacker can have?

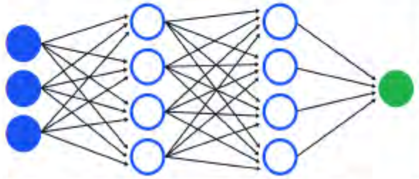
Actual System

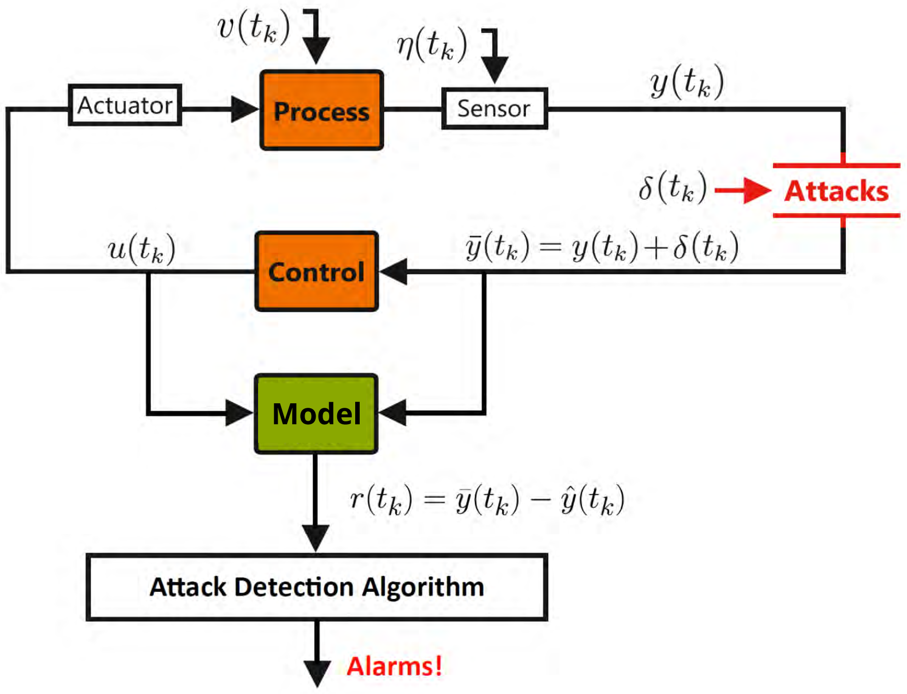


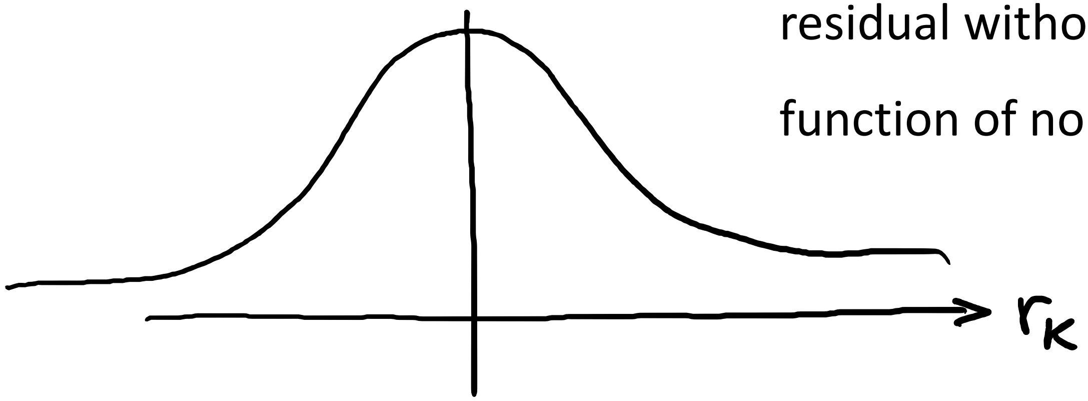
$v(t_k)$
$\eta(t_k)$
Actuator
Process
Sensor
$y(t_k)$
$\delta(t_k)$
Attacks
$u(t_k)$
Control
$\bar{y}(t_k) = y(t_k) + \delta(t_k)$
Model
$r(t_k) = \bar{y}(t_k) - \hat{y}(t_k)$
Attack Detection Algorithm
Alarms!

Model = Digital Twin

$$x_{k+1} = Fx_k + Gu_k + v_k$$




$v(t_k)$
$\eta(t_k)$
Actuator
Process
Sensor
$y(t_k)$
$\delta(t_k)$
Attacks
$u(t_k)$
Control
$\bar{y}(t_k) = y(t_k) + \delta(t_k)$
Model
$r(t_k) = \bar{y}(t_k) - \hat{y}(t_k)$
Attack Detection Algorithm
Alarms!



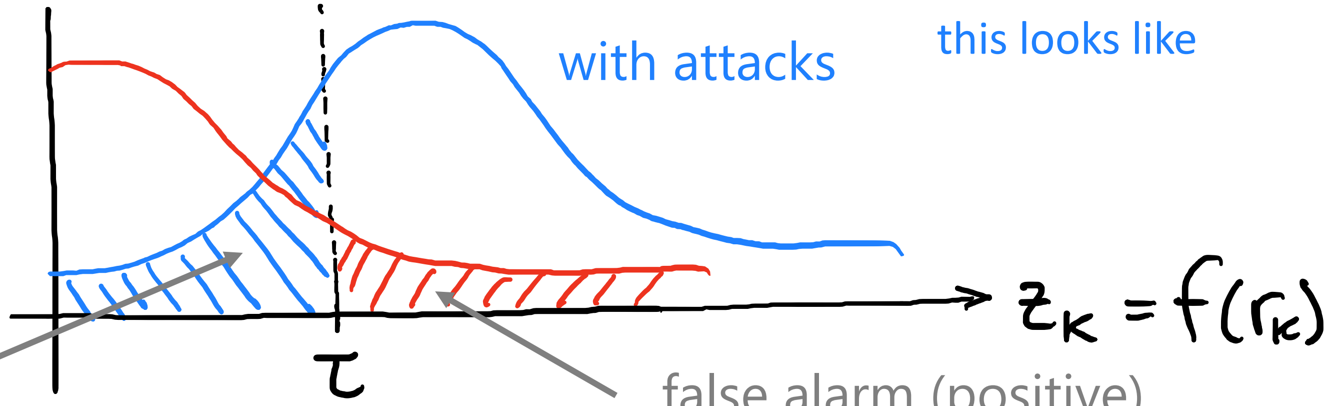
residual without attacks

function of noises only

design a distance measure

e.g., $z_k = |r_k|$

without attacks



with attacks

Attacks can be arbitrary, so we don’t know what this looks like

false negative
probability of not raising an alarm when there is an attack

false alarm (positive)
probability of raising an alarm with no attack

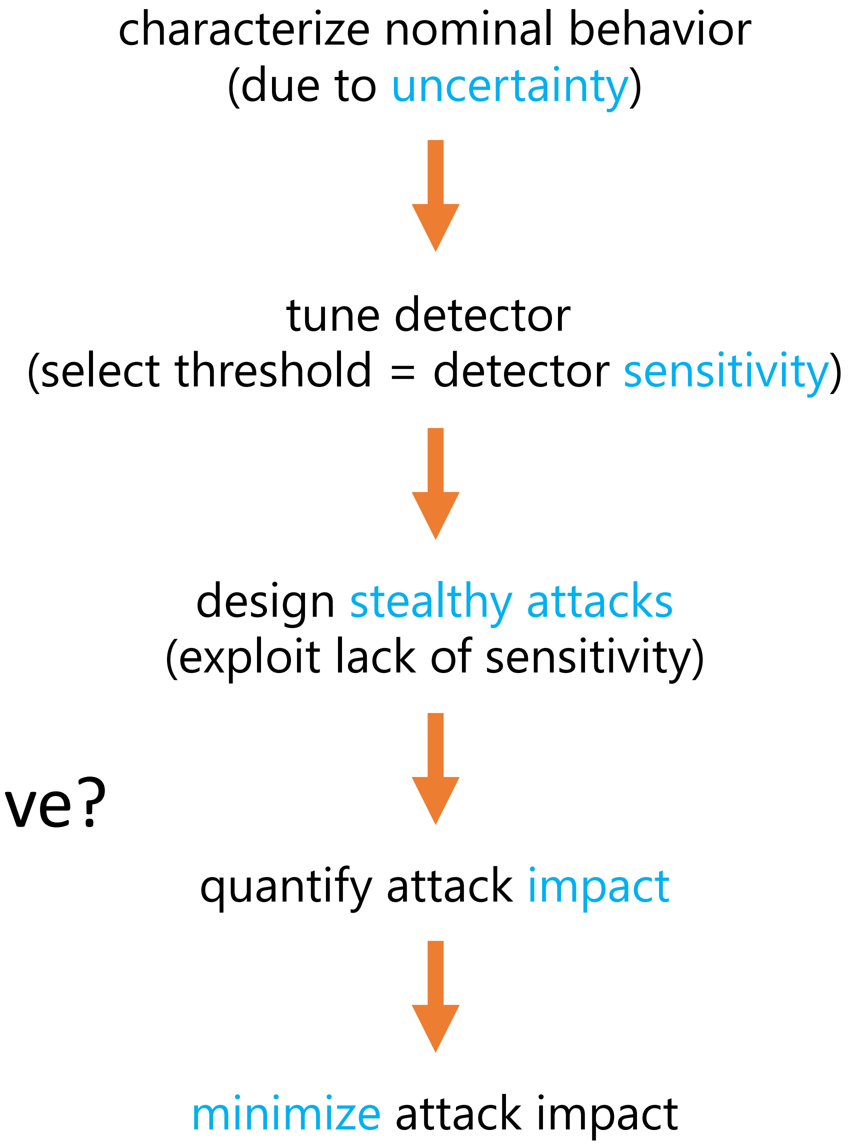
new challenges: how do we...

Challeng es

- detect attacks?
- react against attacks?
- design attack-resilient systems?

Technical Question s

- tune detector for desired performance?
- quantify the impact that an attacker can have?
- determine potential attacks?
- minimize the impact an attacker can have?



characterize nominal behavior
(due to uncertainty)

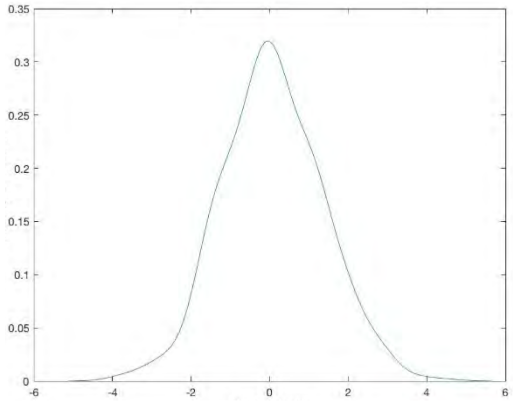
tune detector
(select threshold = detector sensitivity)

design stealthy attacks
(exploit lack of sensitivity)

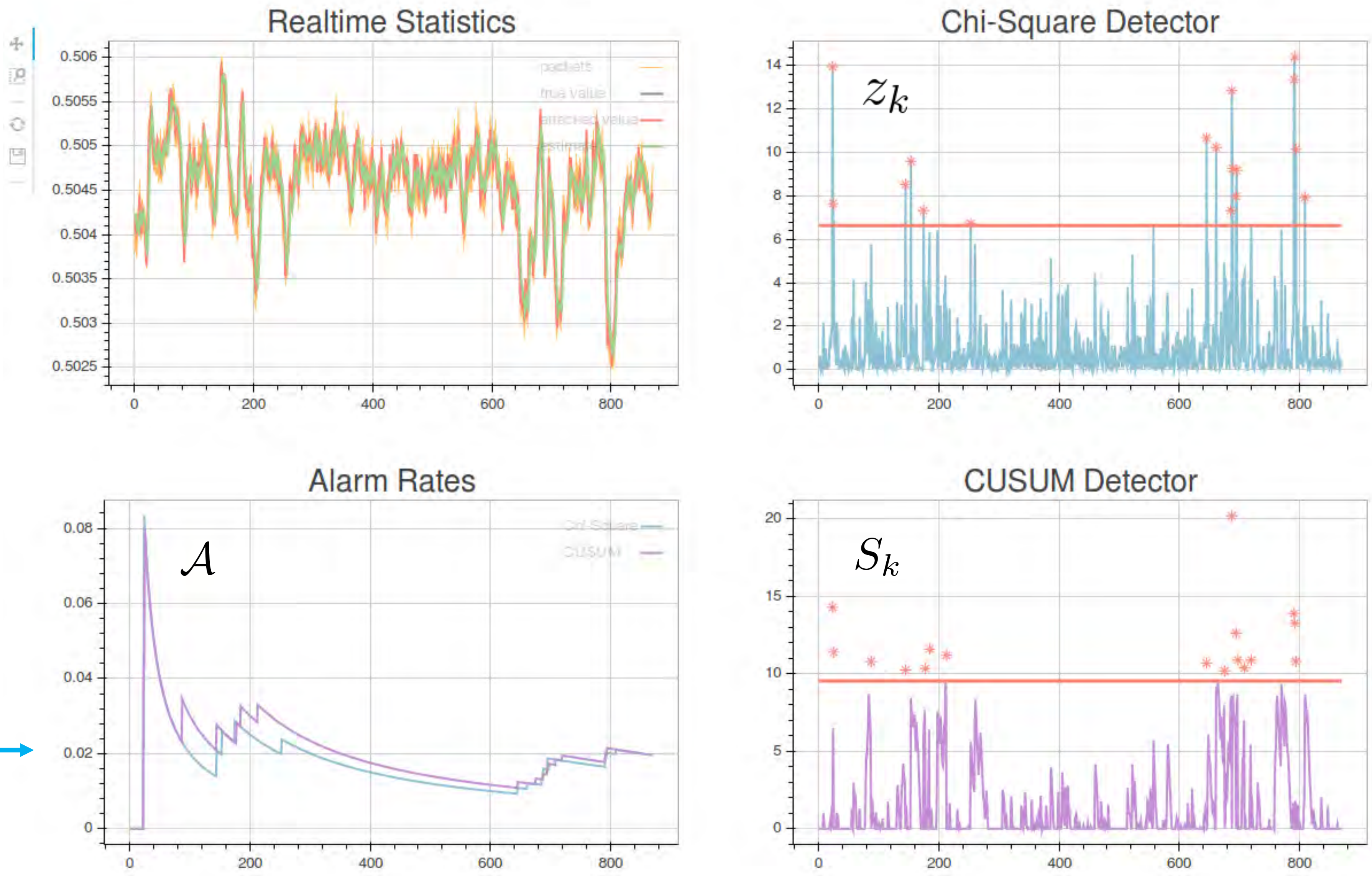
quantify attack impact

minimize attack impact

detector tuning

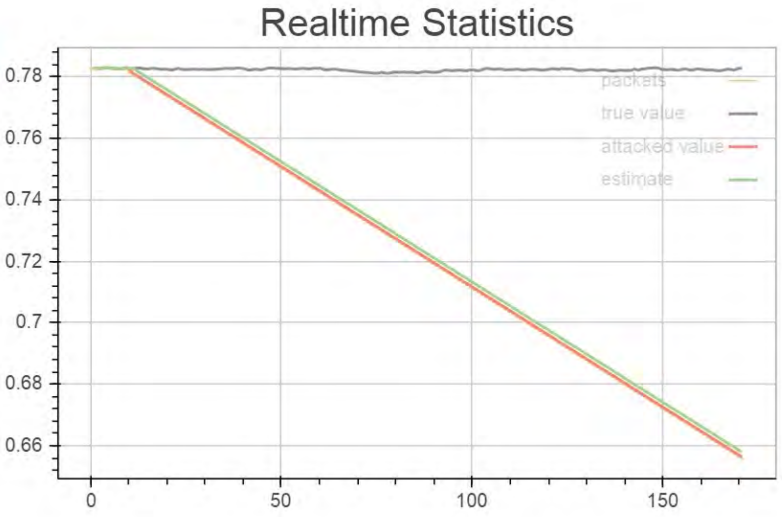


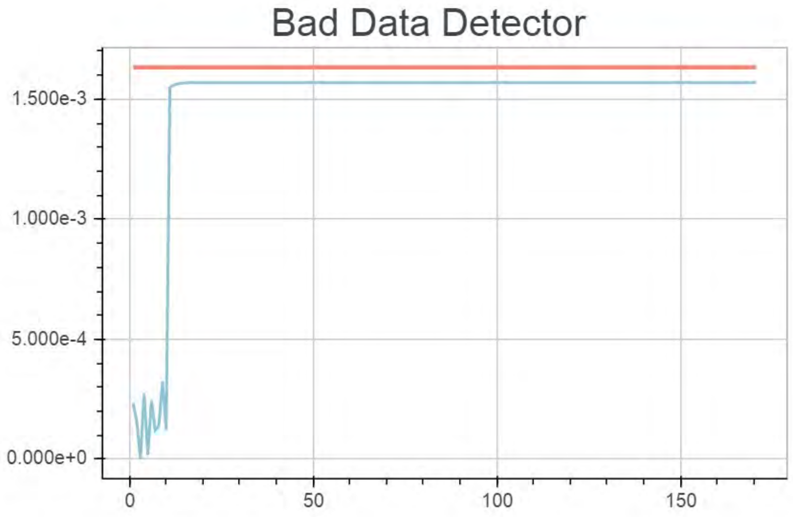
residual PDF

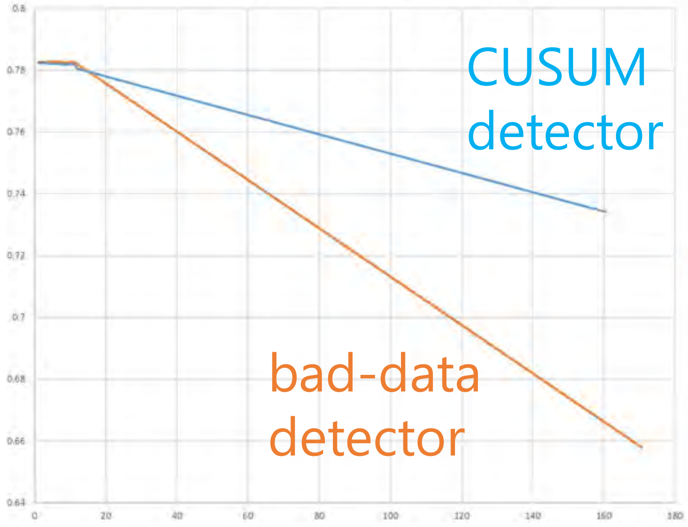


$$\mathcal{A}^* = 0.02$$

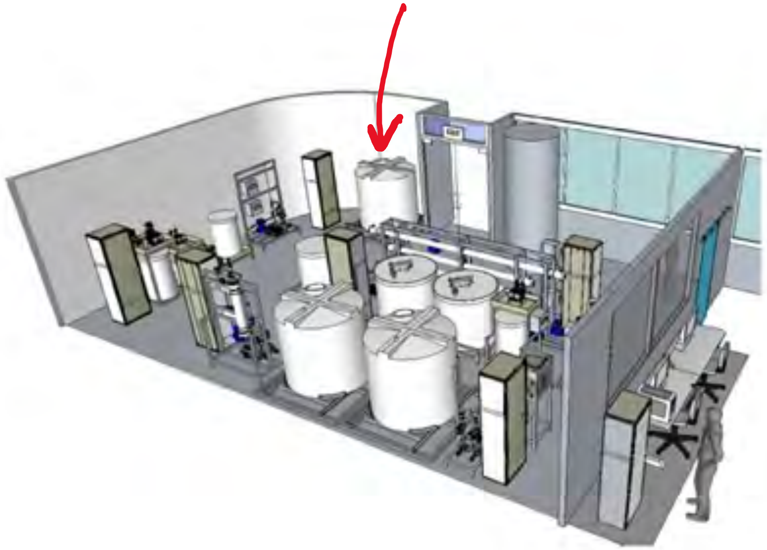
attack impact


Realtime Statistics

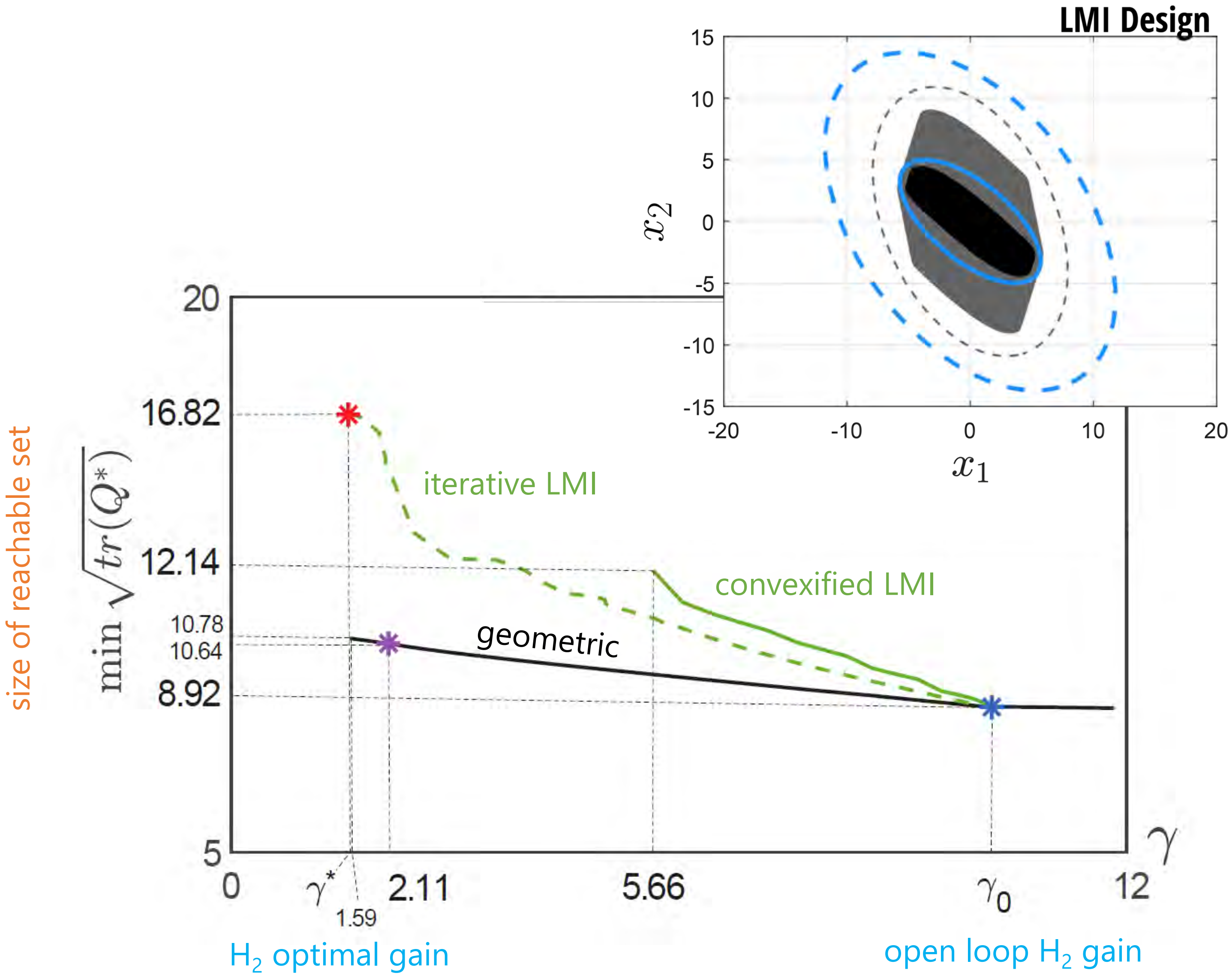

Bad Data Detector

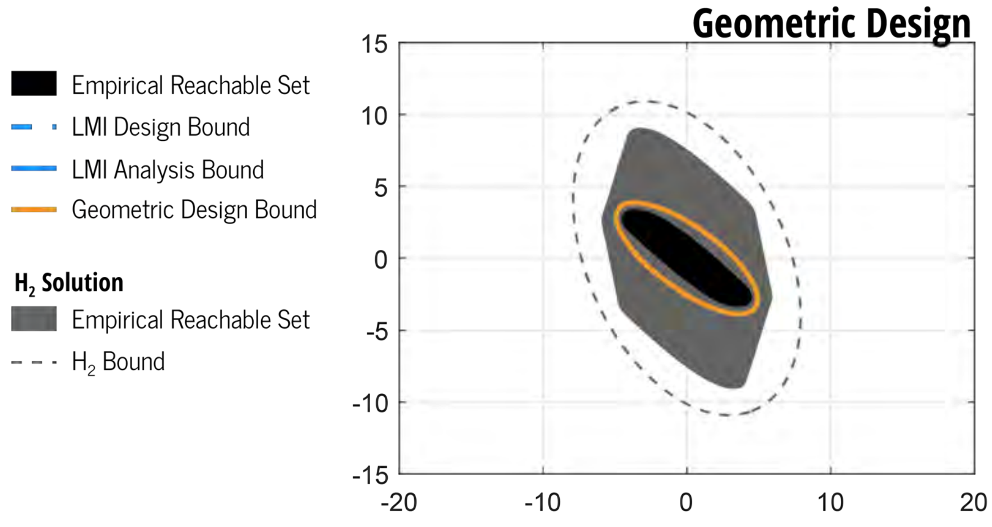

CUSUM detector

bad-data detector



minimizing attack impact



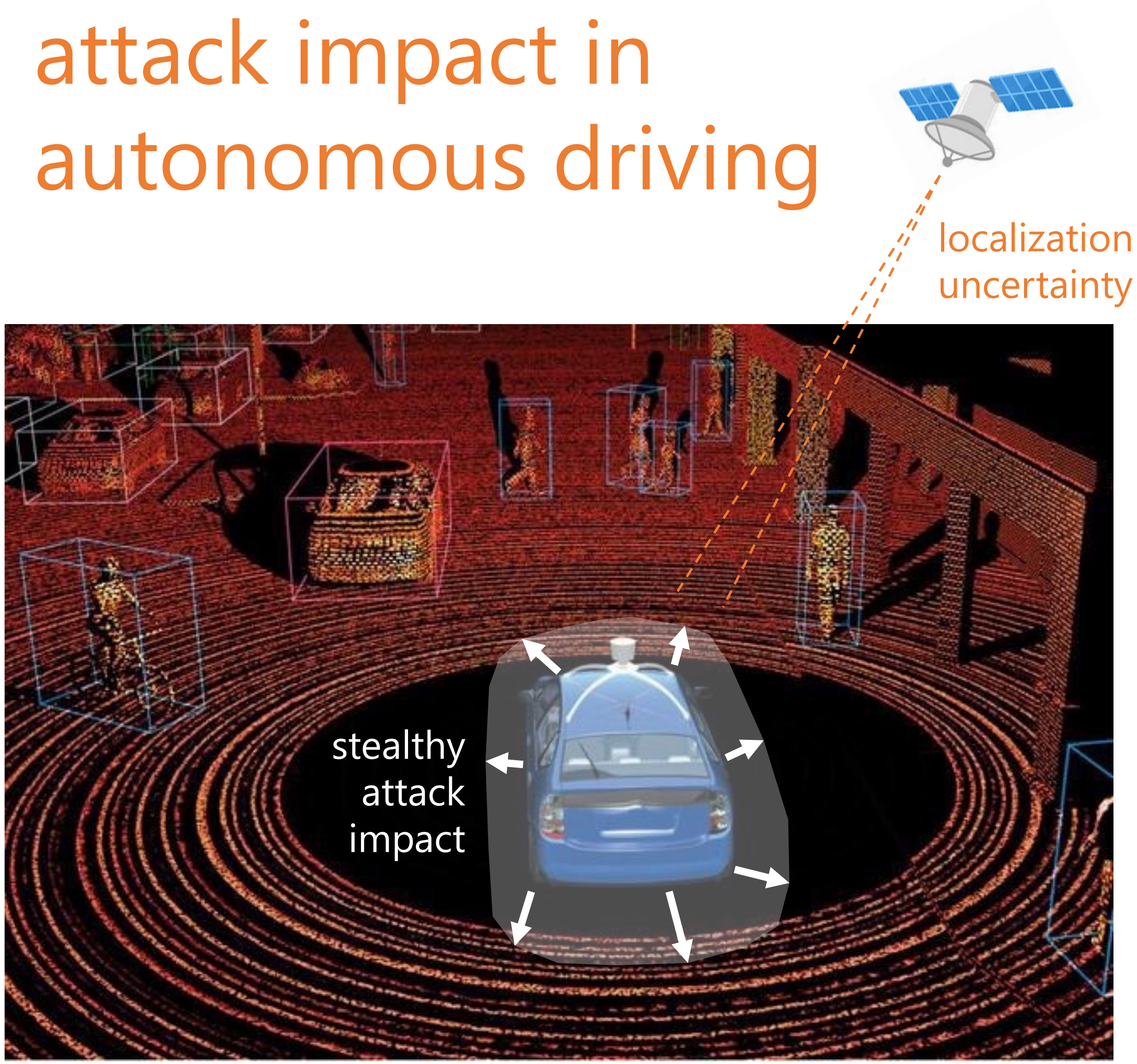


impose performance requirements to avoid trivial zero solution

output covariance constrained H_2:

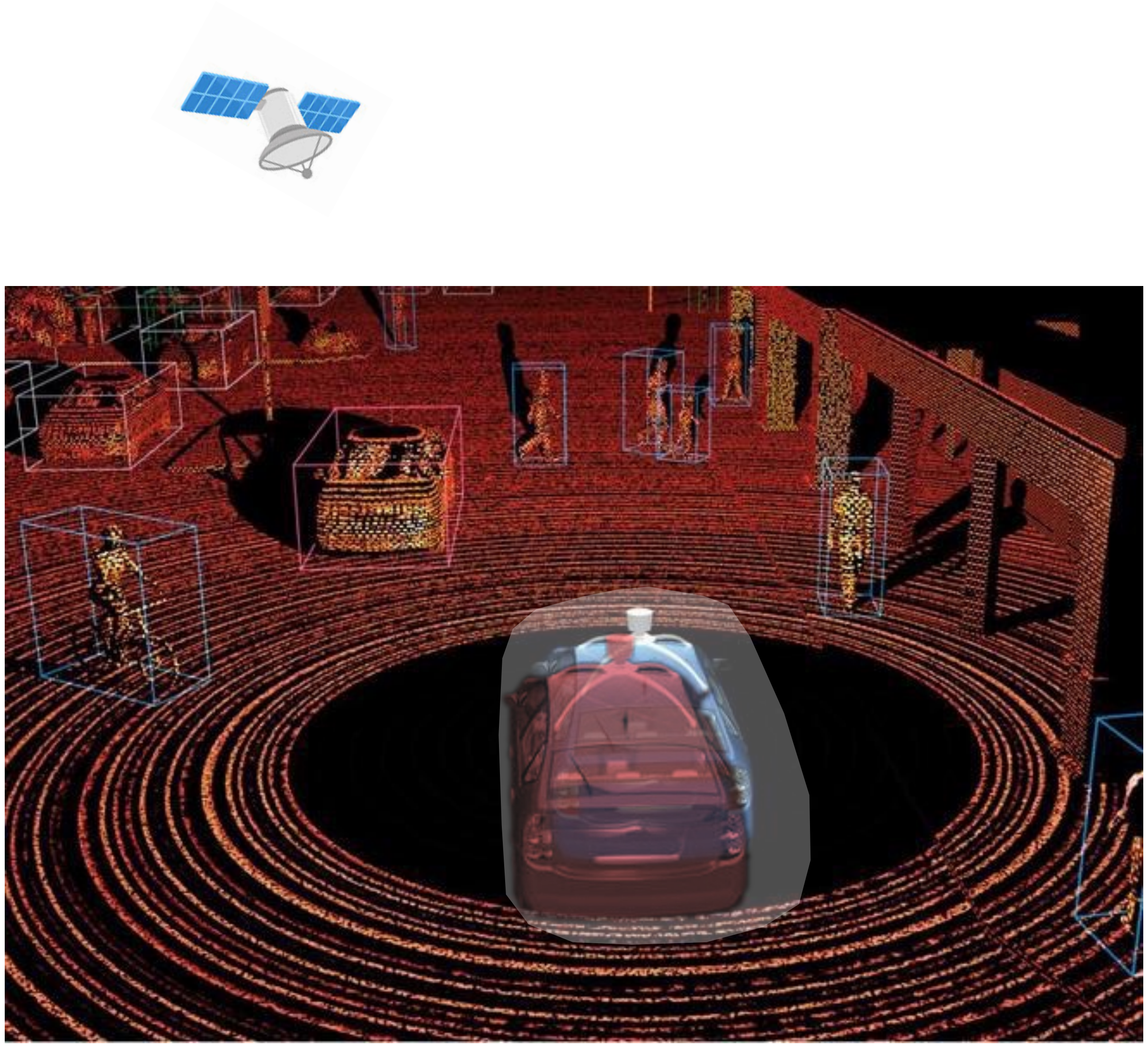
$$\sqrt{\frac{\mathbf{E}[y_k^T y_k]}{\mathbf{E}[\omega_k^T \omega_k]}} \leq \bar{\gamma}$$

attack impact in autonomous driving



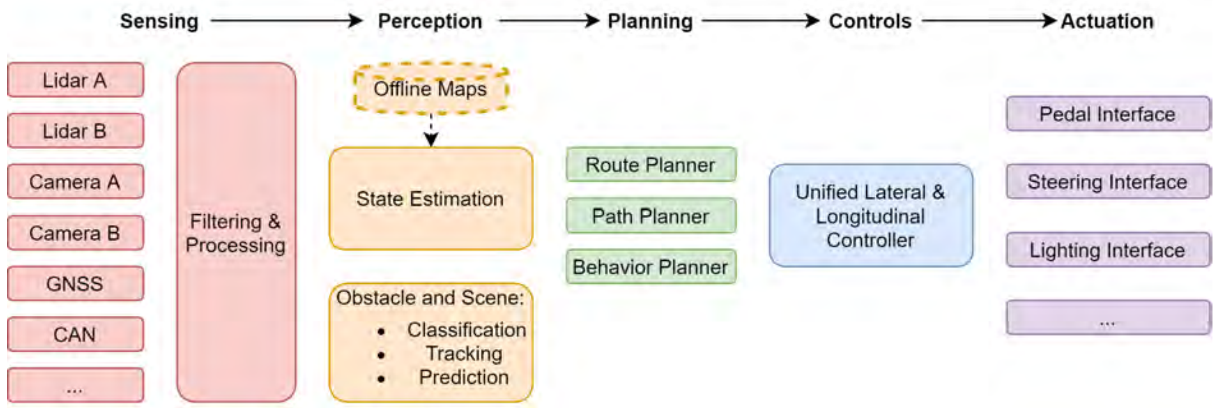
localization uncertainty

stealthy attack impact



Nova

at UT Dallas



Sensing → Perception → Planning → Controls → Actuation

Lidar A

Lidar B

Camera A

Camera B

GNSS

CAN

...

Filtering & Processing

Offline Maps

State Estimation

Obstacle and Scene:

- Classification
- Tracking
- Prediction

Route Planner

Path Planner

Behavior Planner

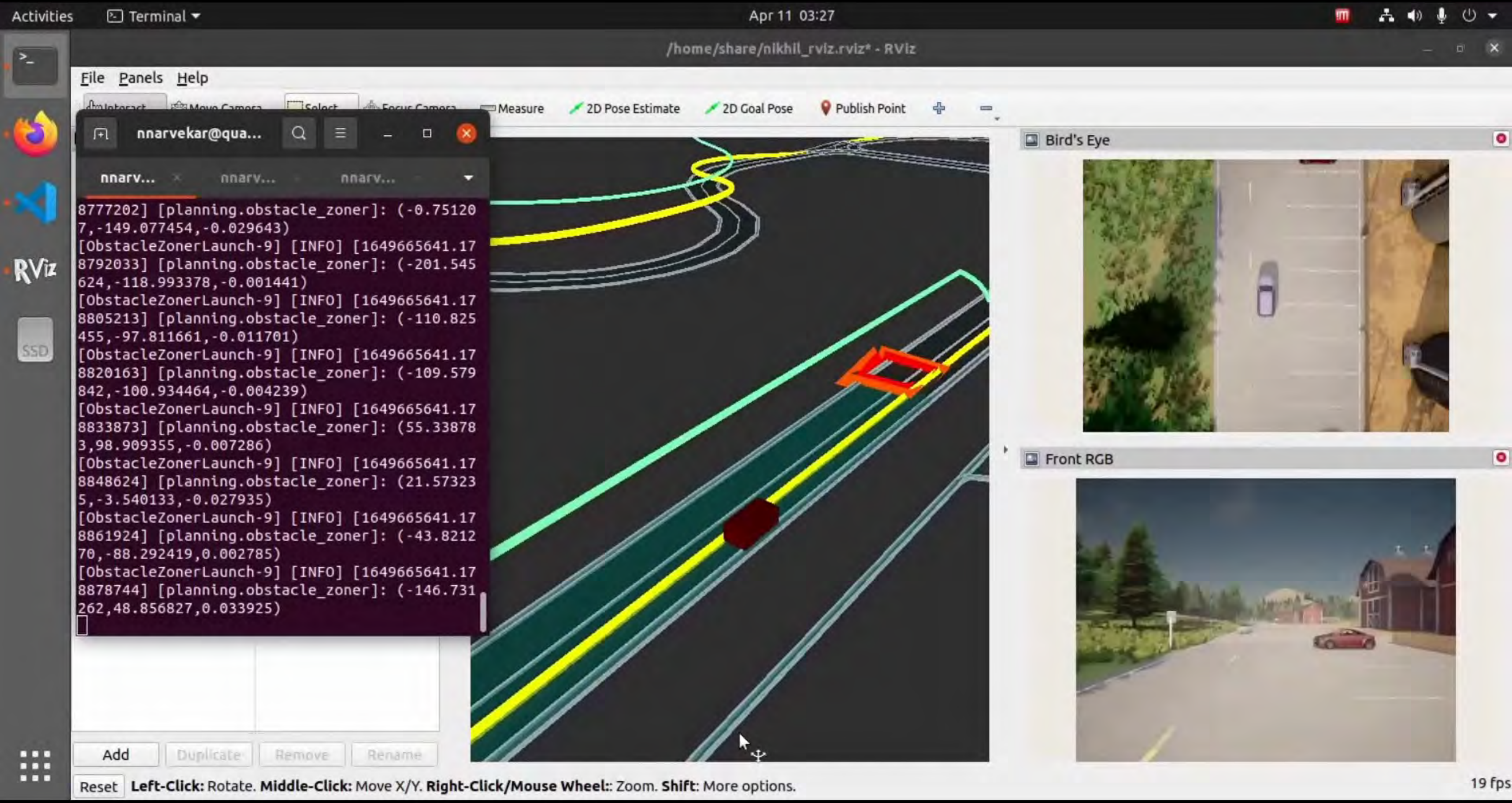
Unified Lateral & Longitudinal Controller

Pedal Interface

Steering Interface

Lighting Interface

...



```
8777202] [planning.obstacle_zoner]: (-0.75120
7,-149.077454,-0.029643)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8792033] [planning.obstacle_zoner]: (-201.545
624,-118.993378,-0.001441)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8805213] [planning.obstacle_zoner]: (-110.825
455,-97.811661,-0.011701)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8820163] [planning.obstacle_zoner]: (-109.579
842,-100.934464,-0.004239)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8833873] [planning.obstacle_zoner]: (55.33878
3,98.909355,-0.007286)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8848624] [planning.obstacle_zoner]: (21.57323
5,-3.540133,-0.027935)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8861924] [planning.obstacle_zoner]: (-43.8212
70,-88.292419,0.002785)
[ObstacleZonerLaunch-9] [INFO] [1649665641.17
8878744] [planning.obstacle_zoner]: (-146.731
262,48.856827,0.033925)
```

new challenges: how do we…

- detect attacks?
- react against attacks?
- design attack-resilient systems?

Nova
at UT Dallas