Convex Relaxation Techniques for Control and System Identification

Bo Wahlberg
Department of Automatic Control
KTH Royal Institute of Technology
Stockholm, Sweden

Outline

1. Autonomous Transport Systems

2. Convex Relaxation for Sparseness

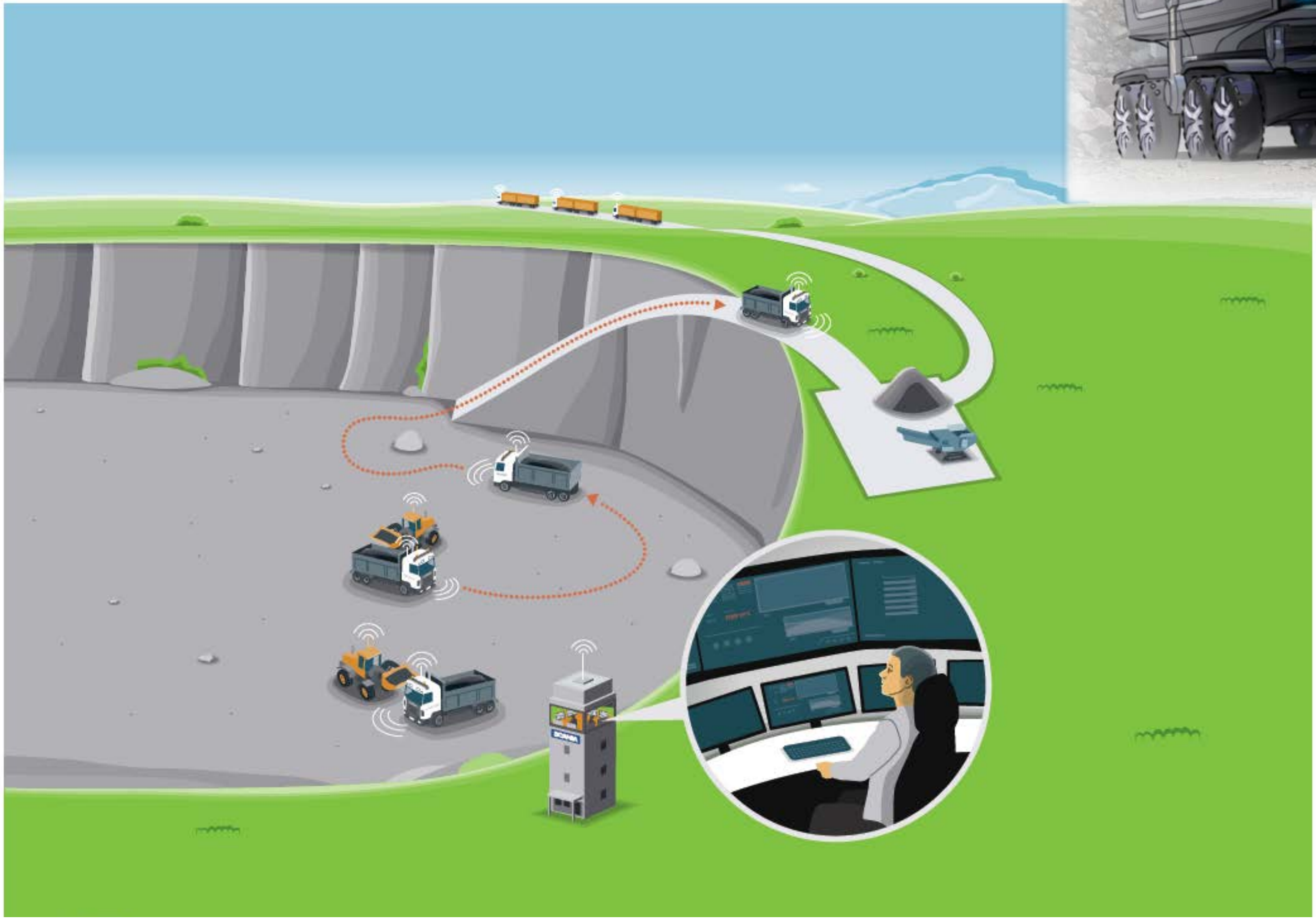
3. Model Predictive Control for Smooth and Accurate Driving

4. Optimal Experiment Design for System Identification

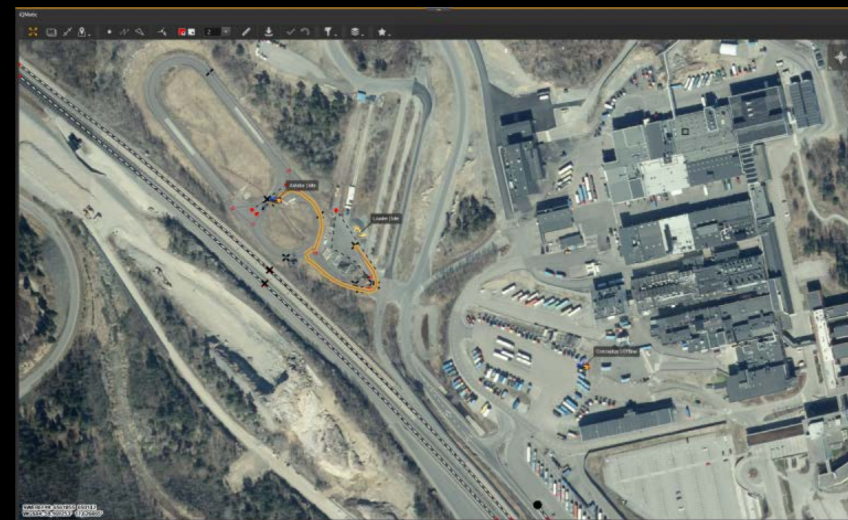
(References in my complementing CCC paper)

Part 1 Autonomous Transport Systems

- Efficiency
- Safety

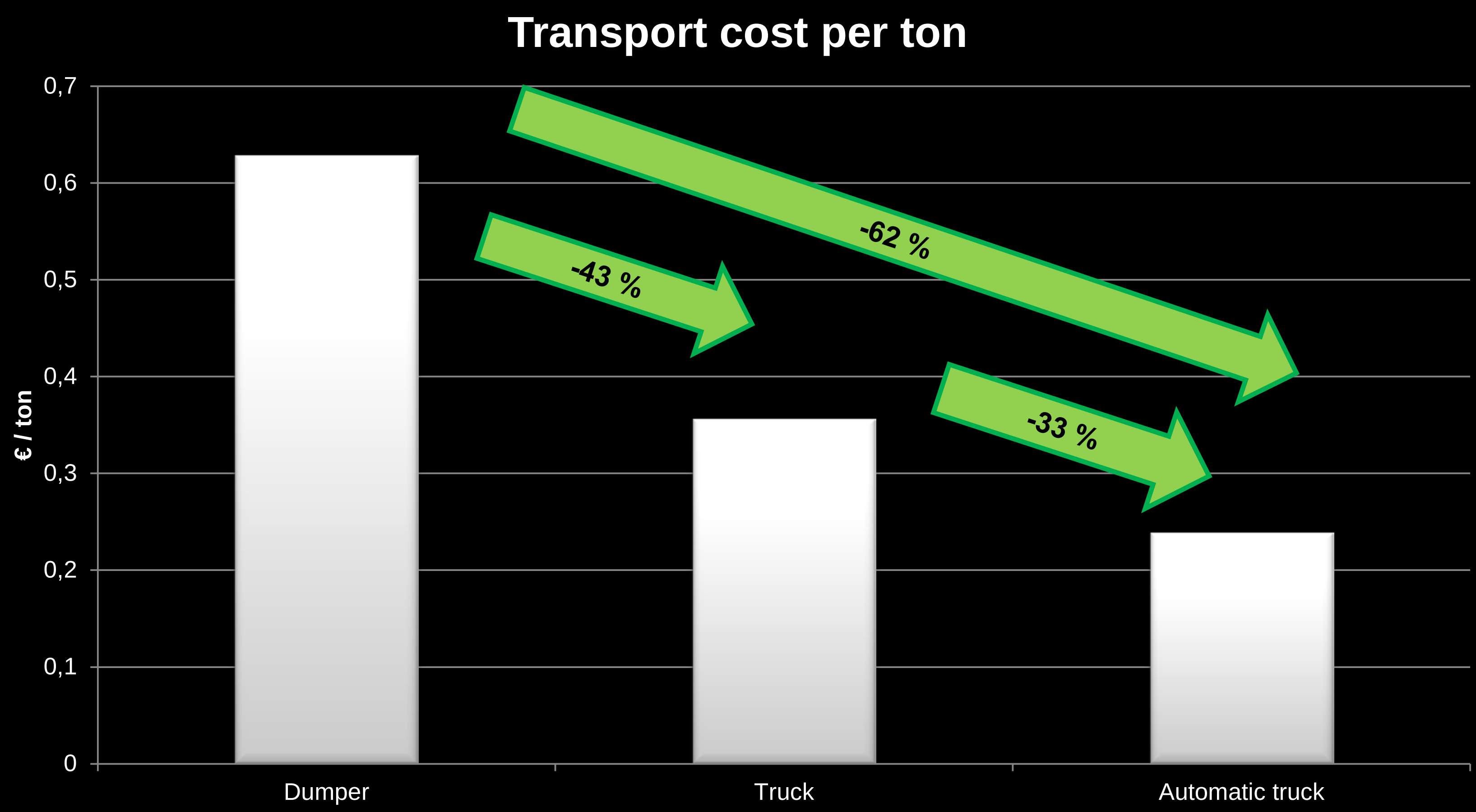


iQMatic - Project Summary



- Autonomous driving in closed-off areas
- Missions from a Command Centre
- Demonstrations (video)
- Tests at Customer site

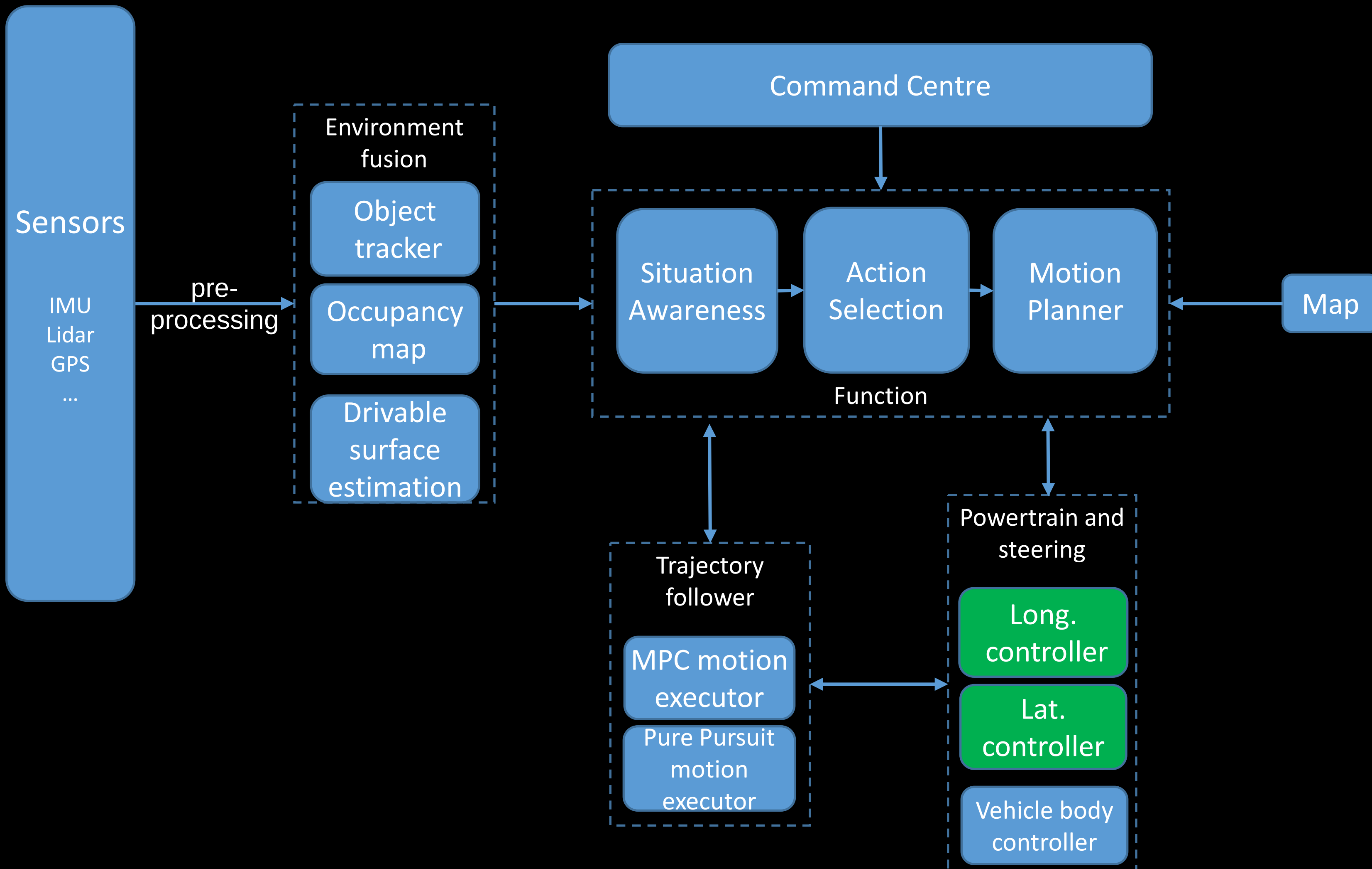
Mining – Business Case from Scania



Astator and its Sensors and Cruise Control

SCANIA

iQMatic – System Overview



To be continued

KTH VETENSKAP OCH KONST

SCANIA

Part 2: Convex Relaxation for Sparseness

Convex optimization methods that promote solutions that are

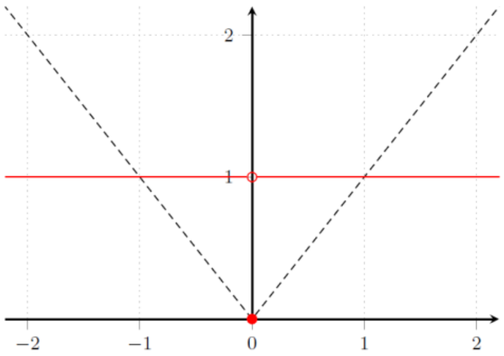
- Sparse
- Piecewise constant
- Piecewise linear

Applications in

1. Change Detection
2. Model Predictive Control

Tools to monitor when these algorithms work or not, and how to fix them!

The $\mathbf{l}_1$ Norm "Trick"



Measure of Sparseness:

The number of non-zero elements in a vector:

$\|\mathbf{x}\|_0 = \sum f_0(x_i)$

The l_1 norm of a vector $\|\mathbf{x}\|_1 = \sum |x_i|$

Convexification: Use $\|\mathbf{x}\|_1$ as substitute for $\|\mathbf{x}\|_0$

Piecewise constant: $\|\mathbf{D}_1\mathbf{x}\|_1 = \sum |x_{i+1} - x_i|$

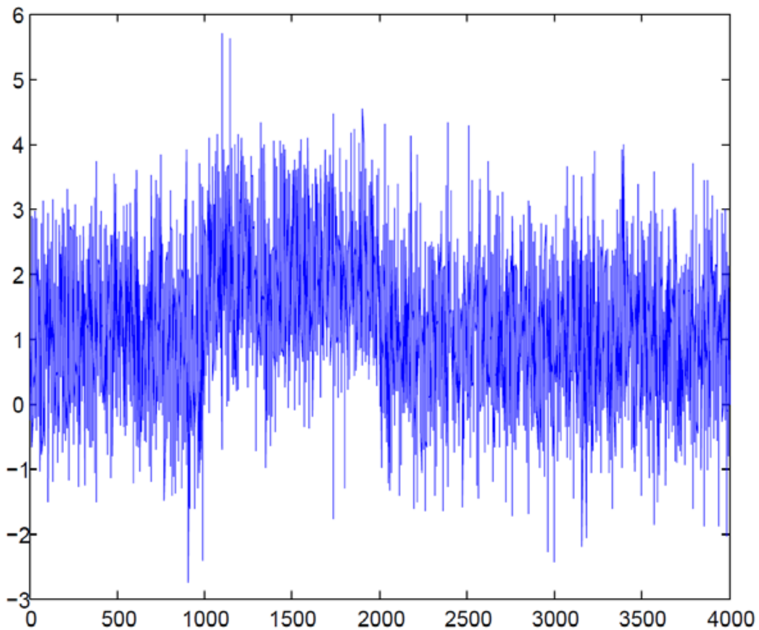
Piecewise linear: $\|\mathbf{D}_2\mathbf{x}\|_1 = \sum |x_{i+1} - 2x_i + x_{i-1}|$

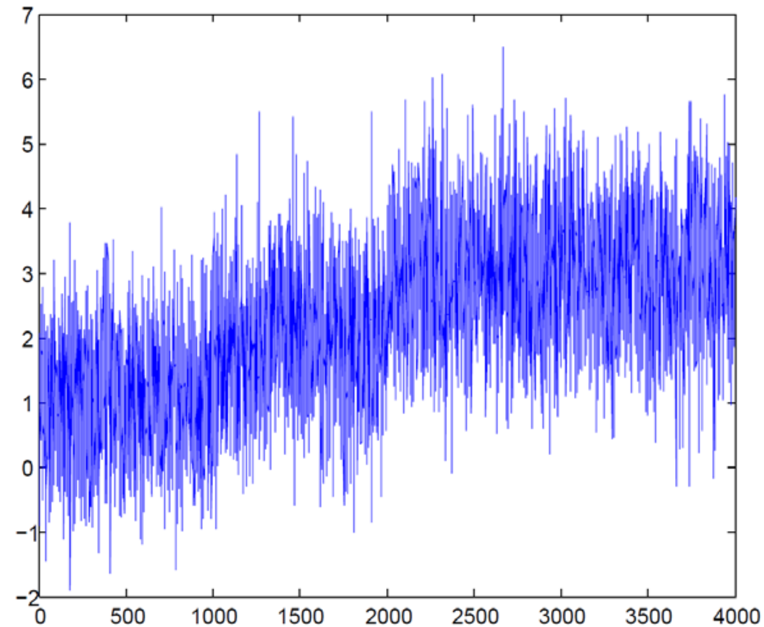
Mean Value Segmentation

Data: $\{y_t,\ t = 1, \ldots, N\}$

Model: $y_t \sim \mathrm{N}(m_t, 1)$, where m_t is piecewise constant

Problem: Estimate the means $m_t,\ t = 1, \ldots, N$.





Fused Lasso

Method: ML+ TV penalty:

$$\boxed{\min_{m_t} \frac{1}{2}\sum_{t=1}^{N}[y_t - m_t]^2 + \lambda \sum_{t=2}^{N} |m_t - m_{t-1}|}$$

Promote **sparseness** of $[m_t - m_{t-1}]$ using the l_1 norm!

Large convex optimization problem that is well-suited for fast solvers such as ADMM

The Lasso method by Tibshirami, goes back to Laplace?

One design parameter: $\lambda \leq \lambda_{max}$ (analytic expression)

Properties of Optimal Solution

$$\min_{m_t, w_t} \frac{1}{2}\sum_{t=1}^{N}[y_t - m_t]^2 + \lambda \sum_{t=2}^{N} |w_t|$$

subject to: $w_t = m_t - m_{t-1}$

Dual variable: $z_t = \sum_{j=1}^{t-1}[m_j - y_j],\ t = 2, \ldots, N$

Optimality conditions:

$$|z_t| \leq \lambda,\ t = 2, \ldots, N, \quad z_1 = z_{N+1} = 0$$

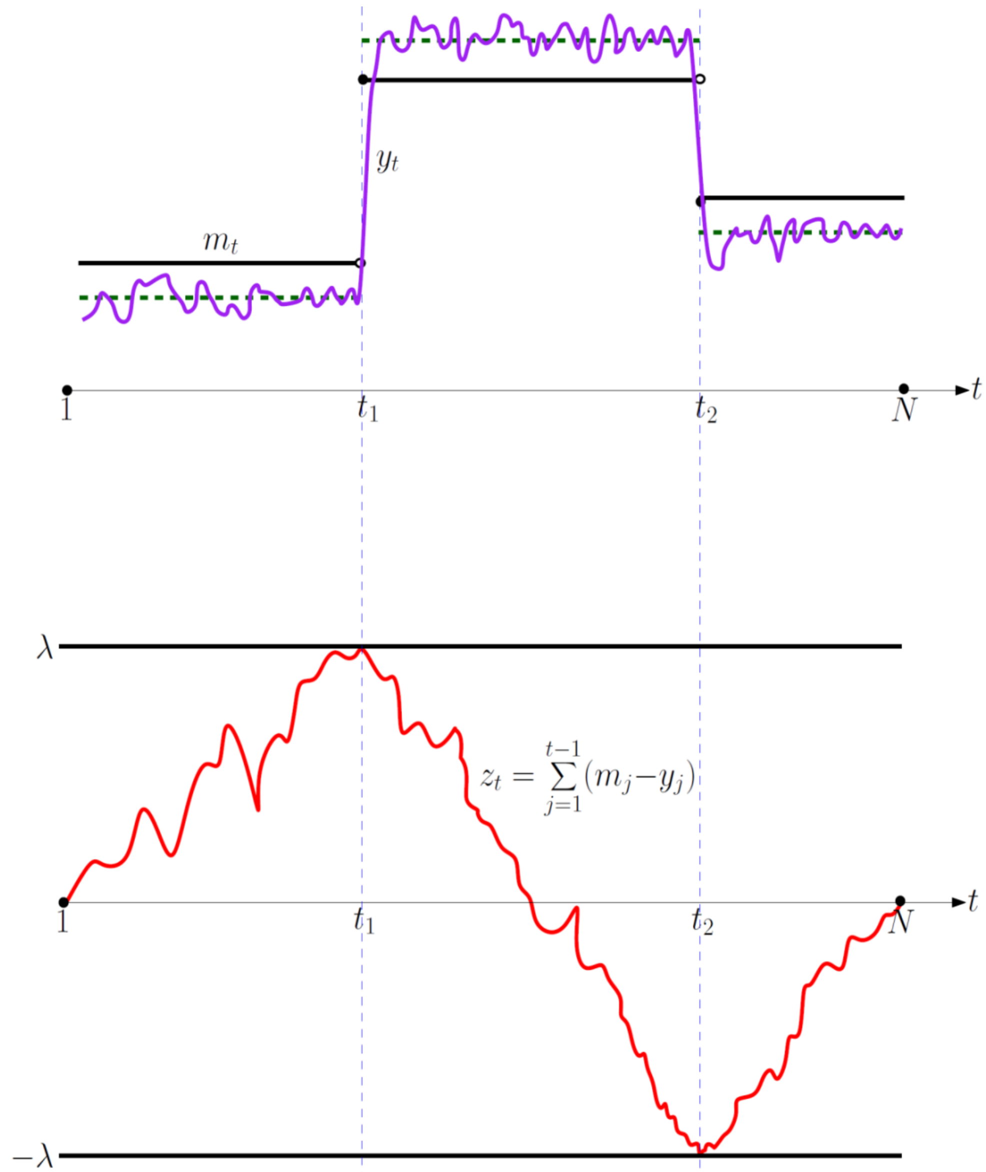
$$|z_t| < \lambda \text{ (constant)} \quad \Rightarrow\ m_t = m_{t-1}$$

$$|z_{t_k}| = \lambda \text{ (change)} \quad \Rightarrow\ \text{sign}[m_{t_k} - m_{t_k-1}] = \text{sign}[z_{t_k}]$$

where $t_0 = 1 < t_1 < \ldots < t_M \leq N$ are the optimal transition times

Reflective Brownian Bridge

The *optimal solution*
is characterized by
a random walk
with drifts, reflections,
& initial/end-constraints



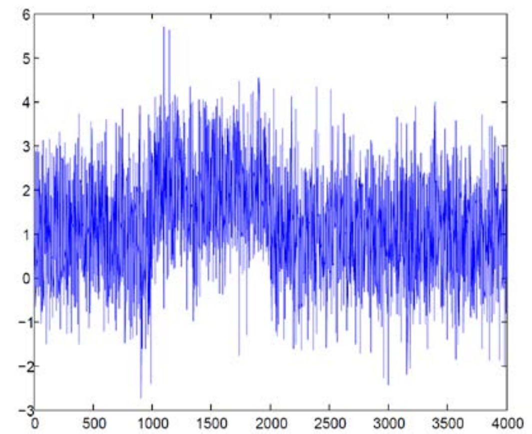
$$z_1 = 0$$

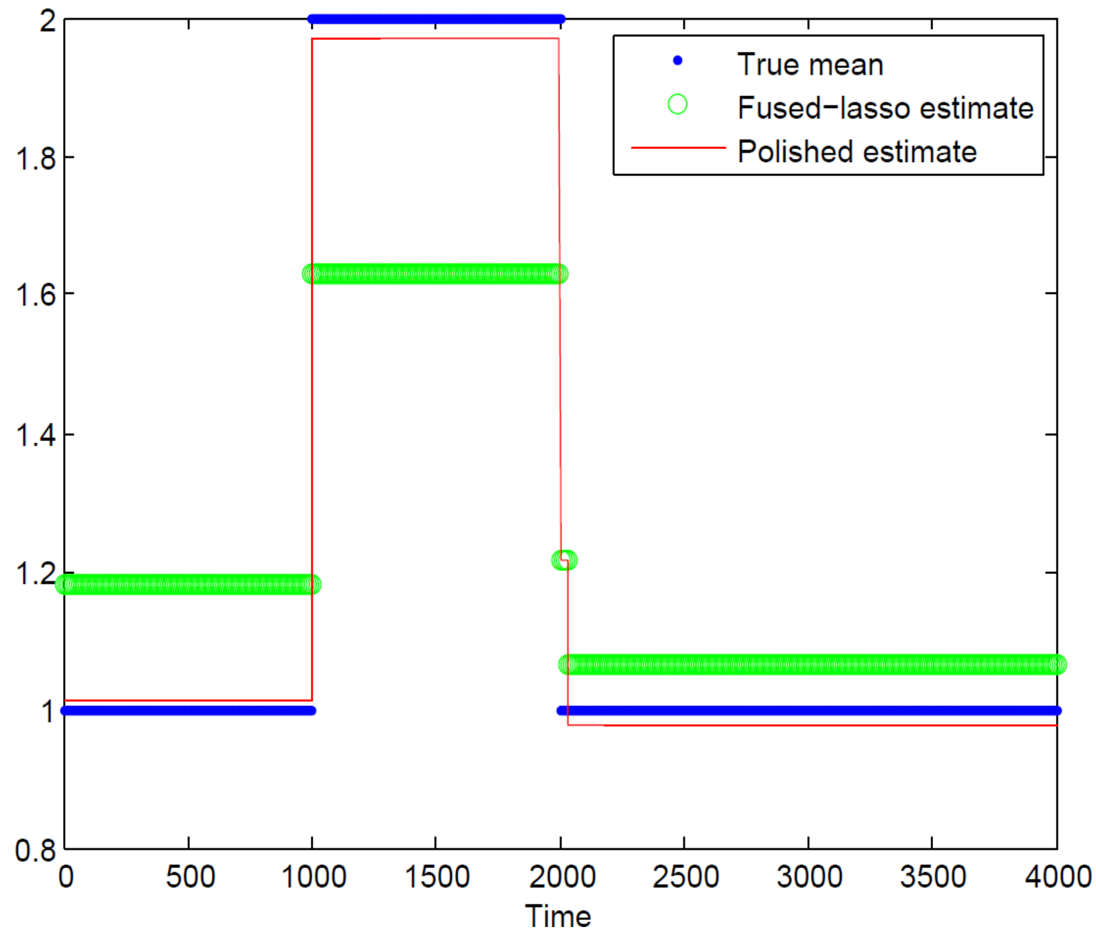
$$z_t = \sum_{j=1}^{t-1}[m_j - y_j],$$

$$z_{N+1} = 0$$

To be used for **monitoring**!

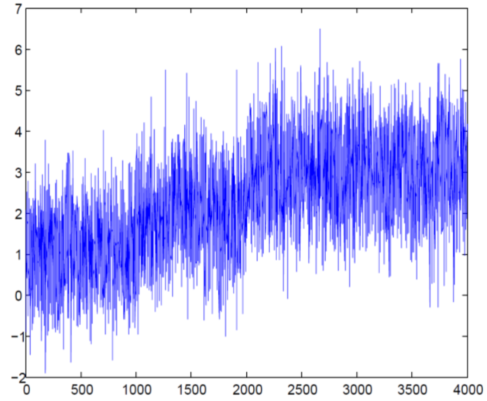
Example 1, Up/Down

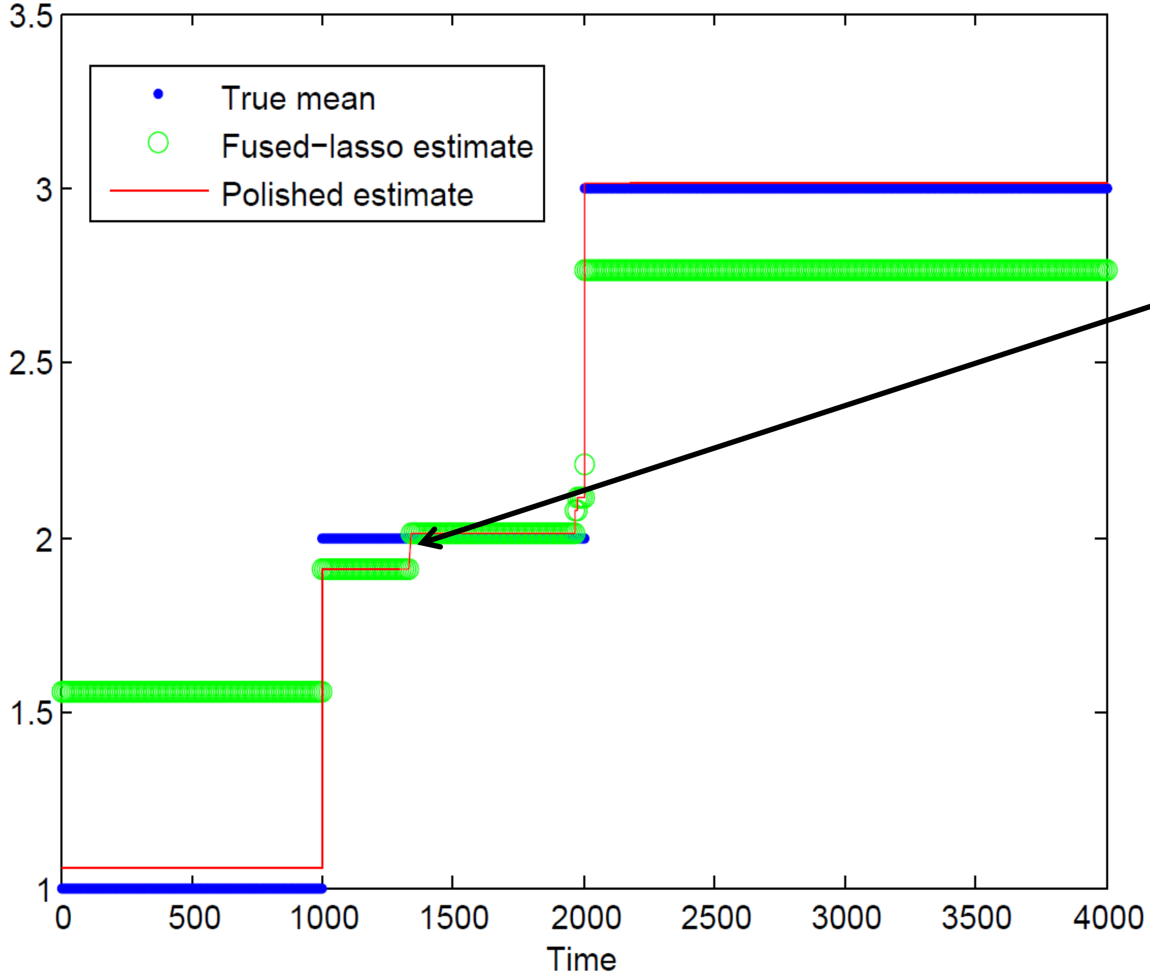




Works fine!

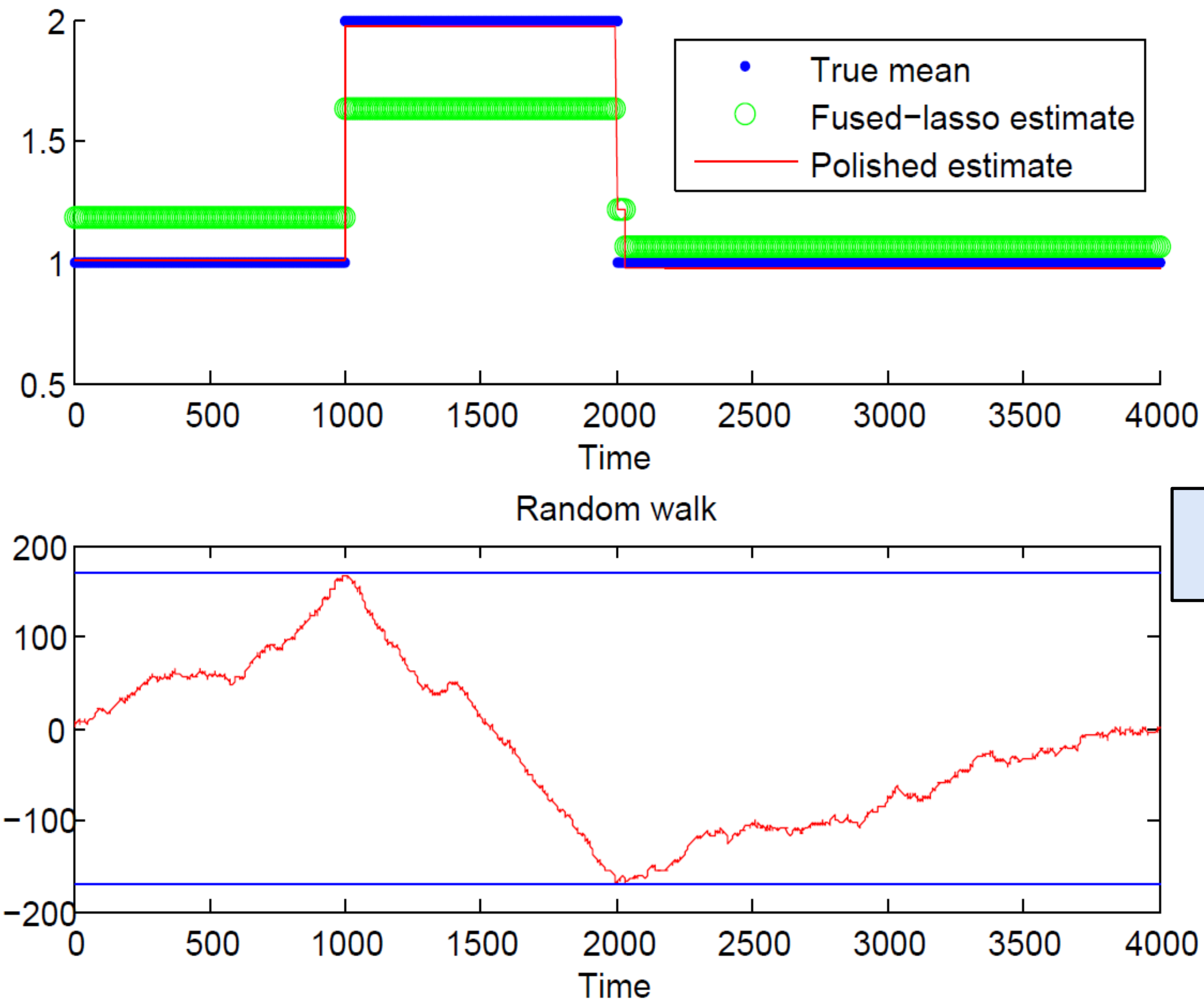
Example 2, Staircase





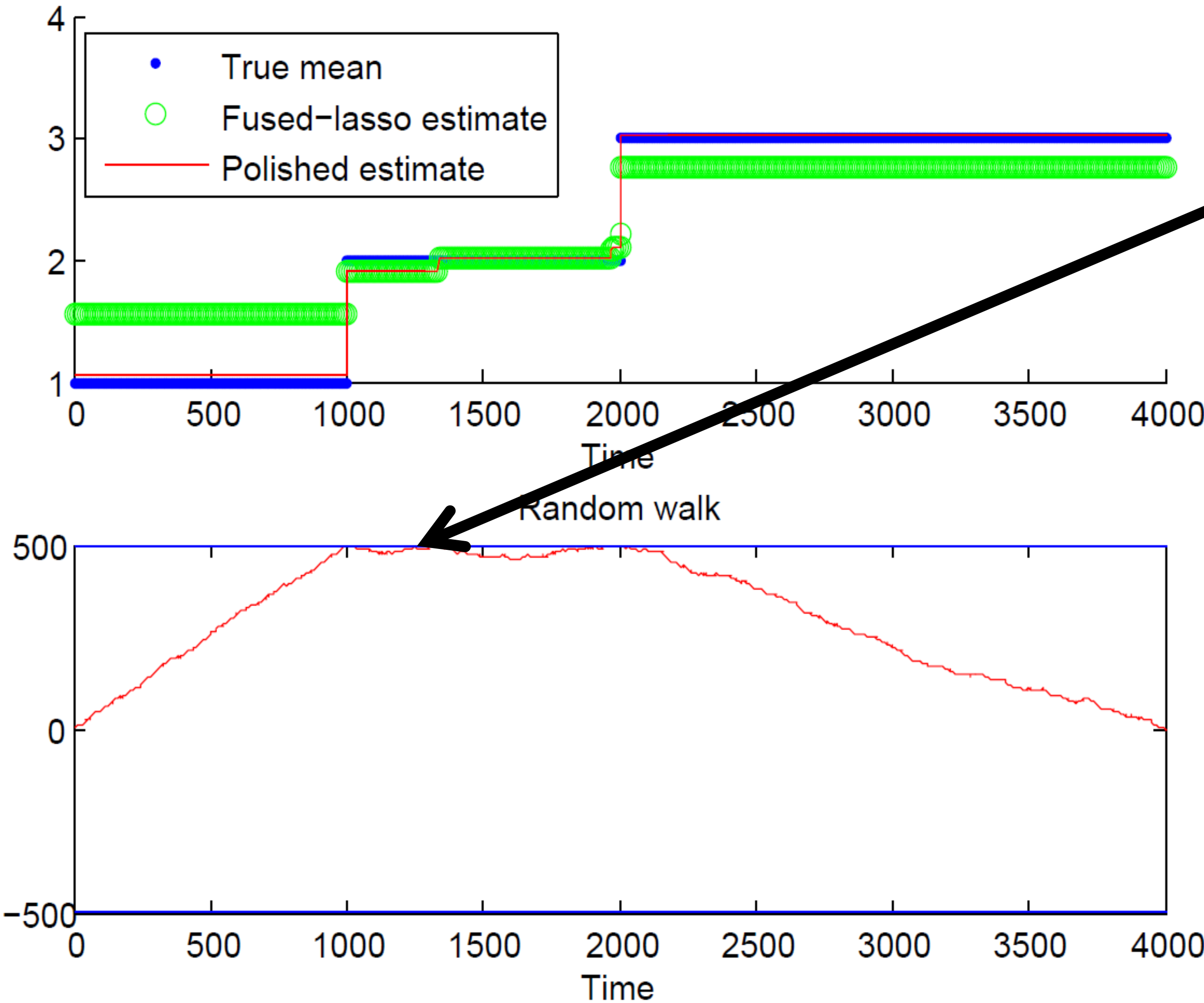
False change point!

Example 1: Monitoring Up/Down



As expected!

Example 2: Monitoring Staircase



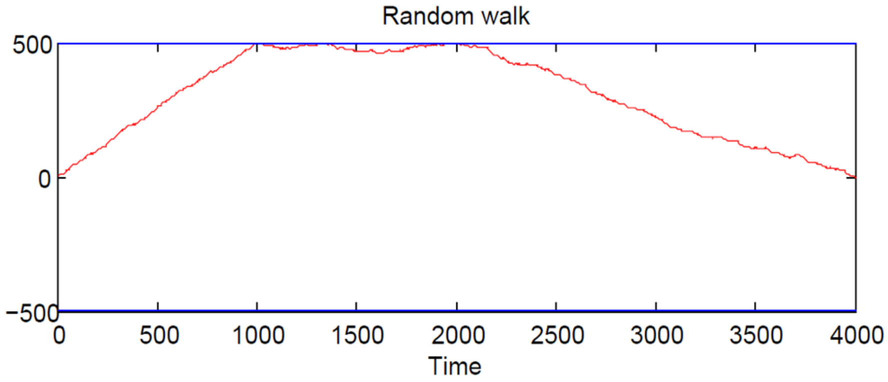
The drift is zero!

Why False Detections

The drift rate equals

$$\frac{1}{t_{k+1} - t_k}\lambda[\mathrm{sign}[m_{t_{k+1}} - m_{t_k}] - \mathrm{sign}[m_{t_k} - m_{t_{k-1}}]$$

For the staircase the drift is **zero!**



The optimal solution is then sensitive to noise, which causes false change points!

Can we fix this?

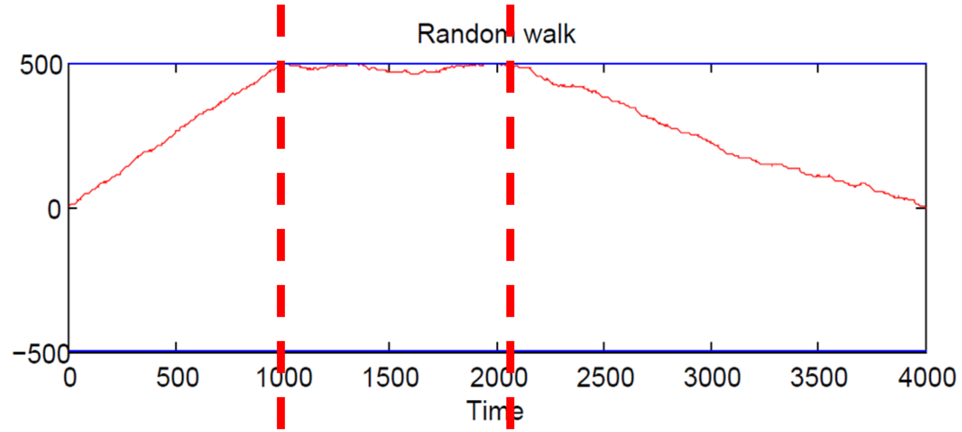
Yes - We can!

The **first** and the **last** change points are accurate!

Restart the algorithm with the interval from the first to the last change points.

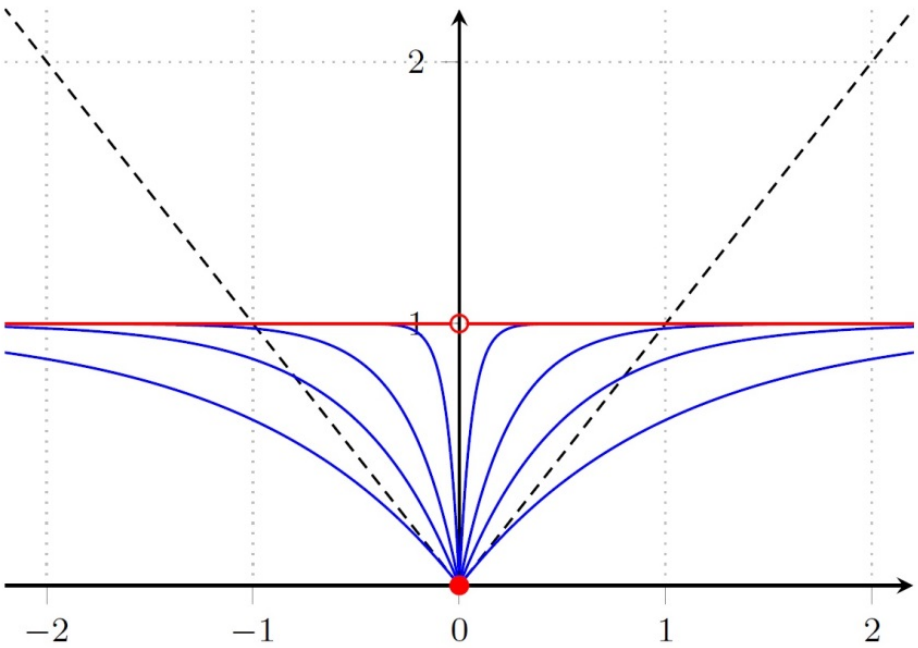
Iterate using the taut-string algorithm, which only solves for the first and last change points.

(Ottersten et.al. 2016)



Yes - We can!

Use a non-convex penalty function: $f(x) = 1 - e^{-x}$



Plot of $f(|x|/\sigma)$ for different σ

Yes - We can!

Preserve convexity:

$$\min_{m_t} \frac{1}{2} \sum_{t=1}^{N} [y_t - m_t]^2 + \lambda \sum_{t=2}^{N} f(|m_t - m_{t-1}|/\sigma)$$

Convex if $\lambda \leq C\sigma^2$

Iterative Adaptive Lasso Algorithm

(Malek-Mohammadi et.al. 2016)

Part 2 Summary

- The l_1-norm is very useful to promote sparseness.
- Monitor the solutions! Improve the algorithms!
- Many applications and extensions!

Ebadat et.al.: Regularized Deconvolution-Based Approaches for Estimating Room Occupancies, *2015 Googol T-ASE Best Applications Paper Award*

Back to Control

Control Applications

Model Predictive Control with l_1 TV input signal penalty

$$\min_{\mathbf{u}} \sum_{k=t}^{t+h} \left\{ (y(t) - r(t))^2 + \lambda |u(t+1) - u(t)| \right\}$$

subject to the state-space equations and other contraints.

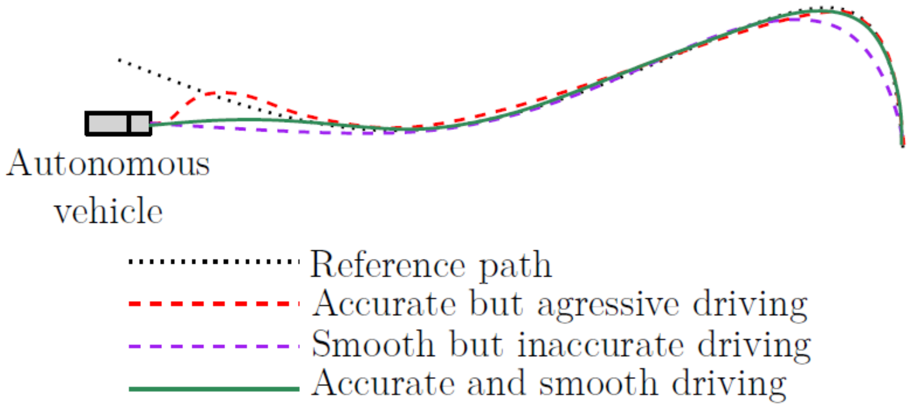
The control actuator should move as few times as possible while tracking the reference!

Gallieri and Maciejowski: **Lasso MPC**, ACC 2012, for over-actuated systems

Jovanovic: **Controller Architectures**,
Semi-plenary at 2016 European Control Conference

Part 3. MPC for Autonomous Driving

Optimize Driving
s.t. Dynamics
Constraints
Safety
Accuracy
Smoothness

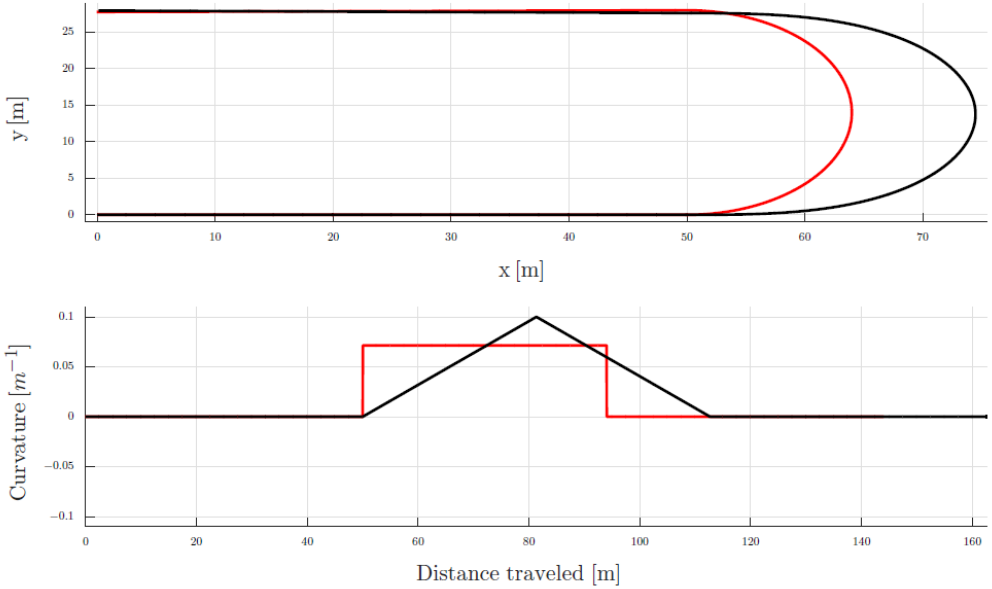


Smooth Driving - Clothoids

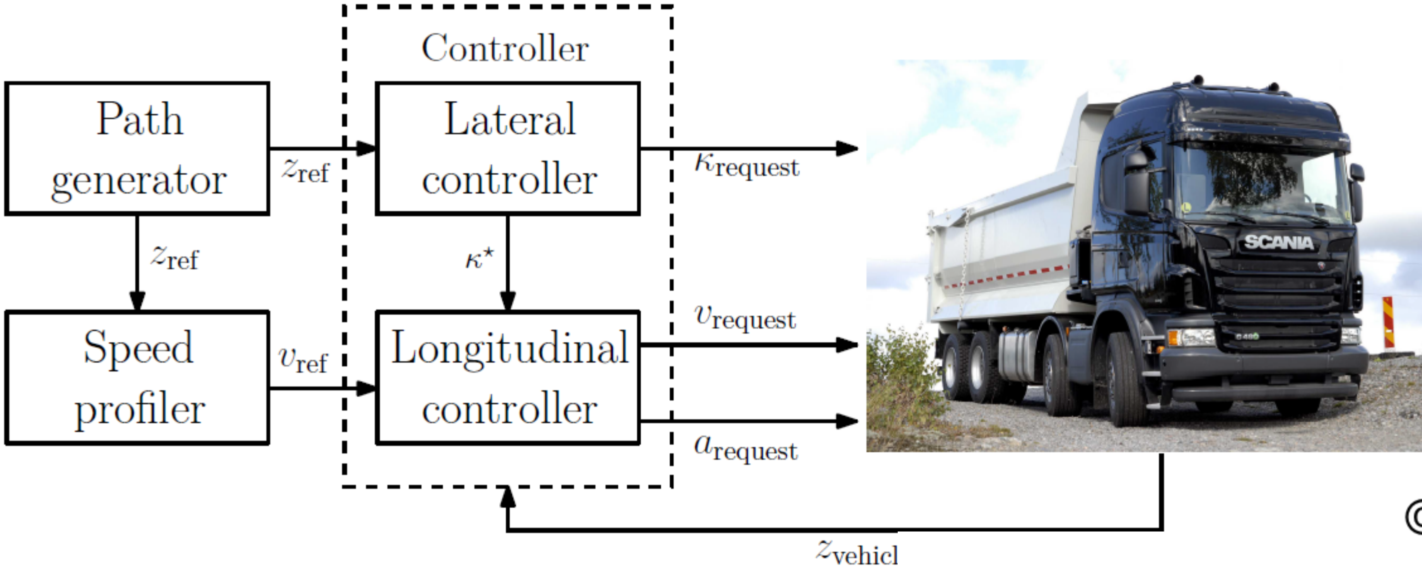
A *clothoid* is a curve whose curvature changes *linearly* with its arc length.

Roads are designed using clothoids to give smooth driving.

”The steering wheel” has a constant rate along a clothoid.



MPC for Autonomous Driving



© Scania.

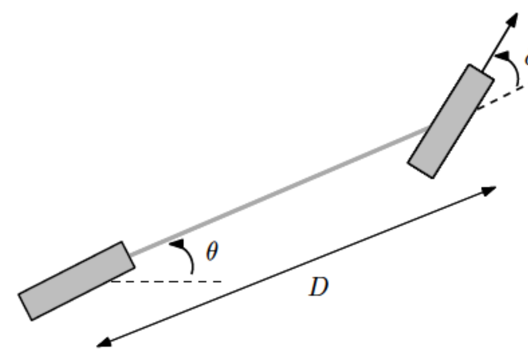
Lateral Control using Curvature κ

- *Minimize rate and rate change of the steering!*

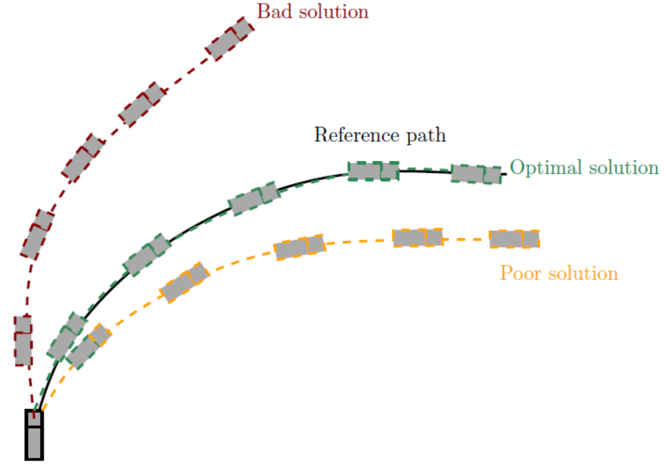
  Aggresive steering maneuvers lead to high values of lateral acceleration and jerk.

- Vehicle motion is predicted by a kinematic model.

- Constrain the vehicle to be in the vicinity of the reference path.



Spatial MPC



Use distance instead of time $ds = v(t)dt$

Current distance traveled: s_0

Prediction horizon $\sim$ distance ponts $s_1, s_2, \ldots, s_H$

Control curvature vector

$$\boldsymbol{\kappa} = [\kappa(s_1), \ldots \kappa(s_H)]$$

Curvature rate: $\mathrm{D}_1\boldsymbol{\kappa}$ (should be small)

Curvature rate change: $\mathrm{D}_2\boldsymbol{\kappa}$ (should small and sparse)

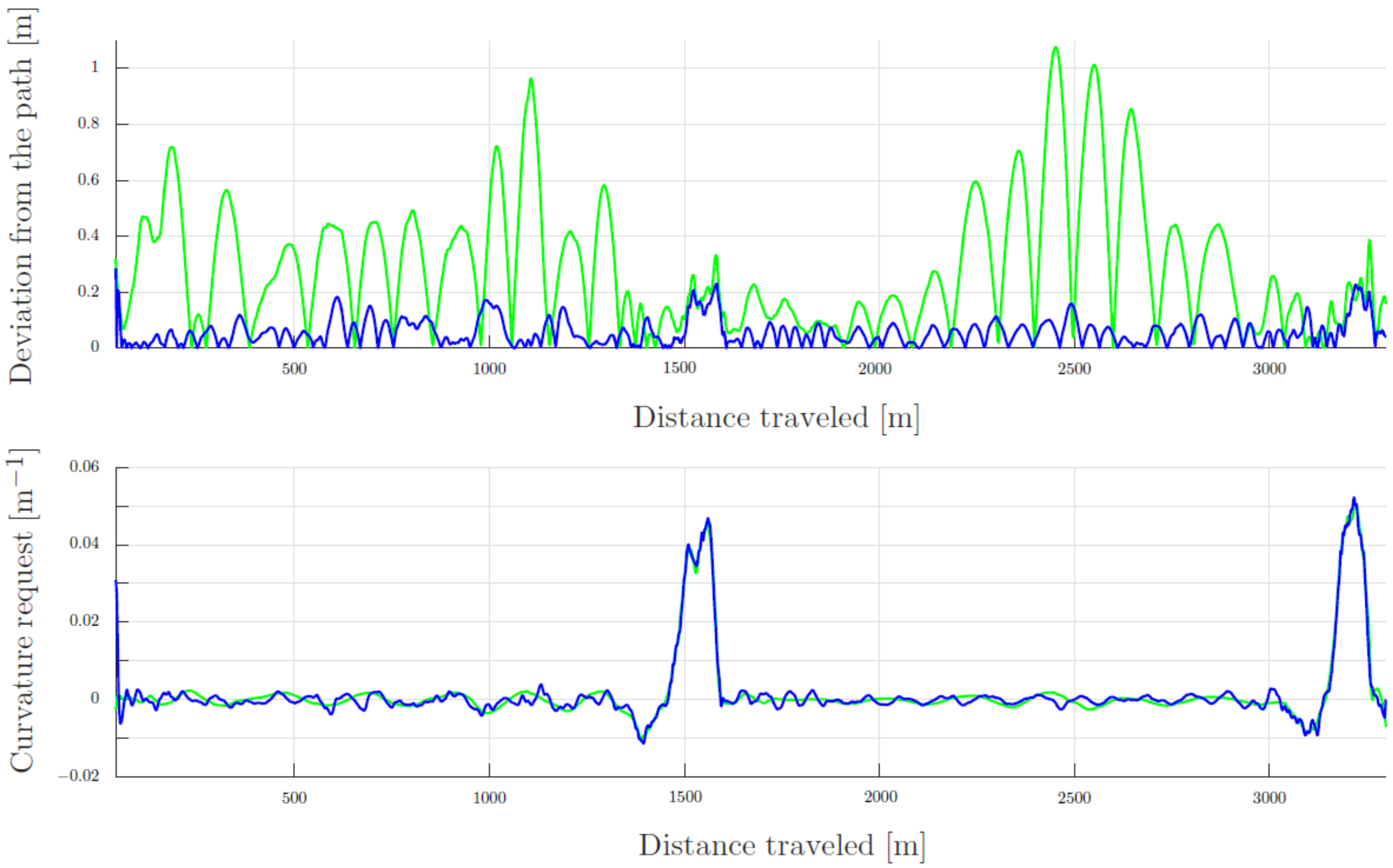
Spatial MPC for Smooth and Accurate Driving

$$
\begin{aligned}
\min_{\boldsymbol{\kappa}} \quad & ||\mathbf{D}_2\boldsymbol{\kappa}||_2^2 + \alpha||\mathbf{D}_1\boldsymbol{\kappa}||_2^2 + \lambda||\mathbf{D}_2\boldsymbol{\kappa}||_1 \\
\text{s. t.} \quad & |\mathbf{x}(\boldsymbol{\kappa}) - \mathbf{x}_{\text{ref}}| \leq \varepsilon_x \\
& |\mathbf{y}(\boldsymbol{\kappa}) - \mathbf{y}_{\text{ref}}| \leq \varepsilon_y \\
& |\mathbf{D}_1\boldsymbol{\kappa}| \leq \mathbf{1}c_{\text{max}} \\
& |\boldsymbol{\kappa}| \leq \mathbf{1}\kappa_{\text{max}} \\
& \kappa(s_0) = \kappa_{\text{vehicle}},
\end{aligned}
$$

(Lima et. al. 2016)

Fast Driving

Green: Human like Controller
Blue: MPC for Smooth and Accurate Driving



Part 3 Summary

- Passengers and vehicles (wear and tear) prefer smooth driving
- Need also to be accurate (stay on the road)
- **We have developed, implemented and tested a MPC for smooth and accurate autonomous driving of a heavy truck**

What about the Model?

Part 4. Optimal Experiment Design for System Identification

The **M** in MPC is often the most difficult part!

System Identification (Data-Driven Modelling) is a most active area of research!

- 8th IFAC Symposium on System Identification (Beijing 1988)
- 17th IFAC Symposium on System Identification (Beijing 2015)
- 18th IFAC Symposium on System Identification (**Stockholm** 2018)

Performing experiments can be both expensive and take extensive time.

System Identification for Control

SYNOPSIS

A technique for numerical identification of a discrete time system from input/output samples is described. The purpose of the identification is to design strategies for control of the system. The strategies are obtained using linear stochastic control theory.

The parameters of the system are estimated by Maximum Likelihood. An algorithm for solving the M. L. equations is given. The estimates are in general consistent, asymptotically normal and efficient for increasing sample lengths. These properties and also the parameter accuracy are determined by the information matrix. An estimate of this matrix is given.

The technique has been applied to simulated data and to plant data.

Åström and Bohlin: Numerical identification of linear dynamic systems from normal operating records, Proc. IFAC Symp on Self Adaptive Systems, UK, 1965 .

Torsten Bohlin 1931- 2016
Professor at KTH 1972-1996

EU-Project Autoprofit 2011-2014

Courtesy of Sasol, South Africa

“It is estimated that 75% of the cost related to a control project in industry is dedicated to the identification of a model.”

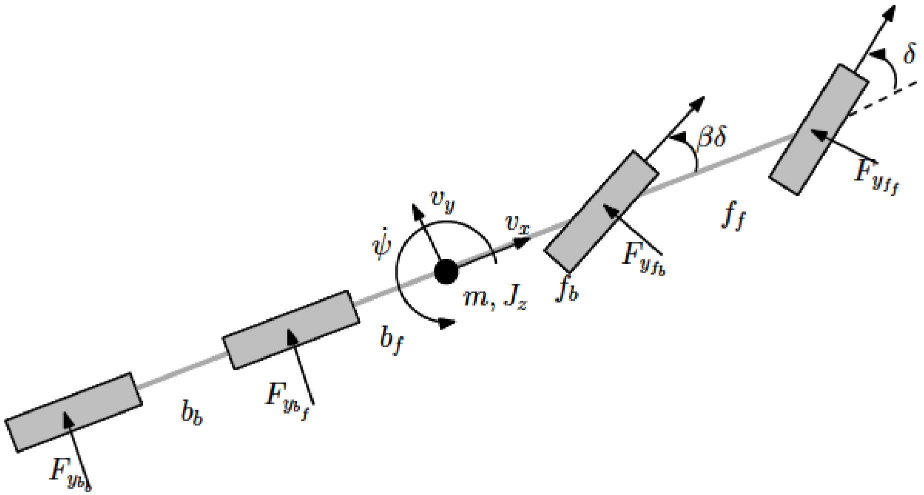
Optimal Input Design

Optimize Excitation

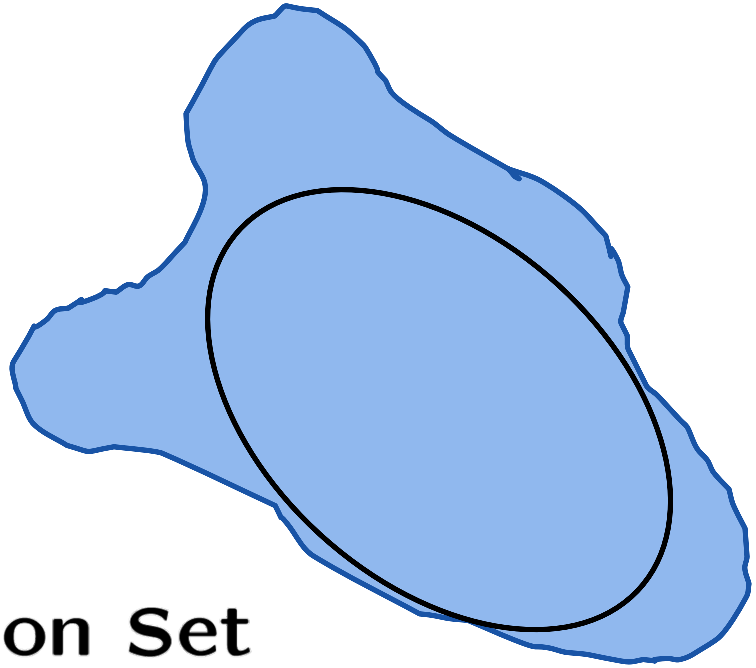
s.t. Identified Model Quality

Experimental Costs

Constraints



Model Based Control



All models in the **Application Set** satisfy the control specifications

Relates to parameter robustness

Ellipsoidal approximation:

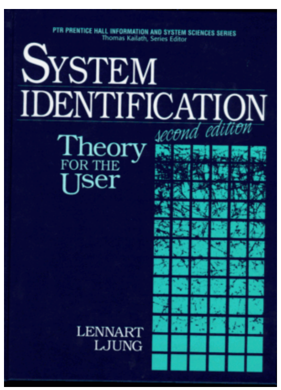
$$[\theta - \theta^o]^T V''_{app}(\theta^o)[\theta - \theta^o] \leq 2/\gamma$$

System Identification Set

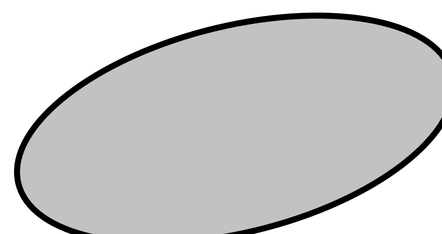
Asymptotically in data size N:

$$\hat{\theta} \in \left\{ [\theta - \theta^o]^T V''_{SI}(\theta^o)[\theta - \theta^o] \leq 2/\kappa \right\}, \ w.p.\alpha$$

See Ljung: System Identification (1999)
and corresponding Matlab SI Toolbox



Confidence ellipsoid for the estimated parameter vector $\hat{\theta}$



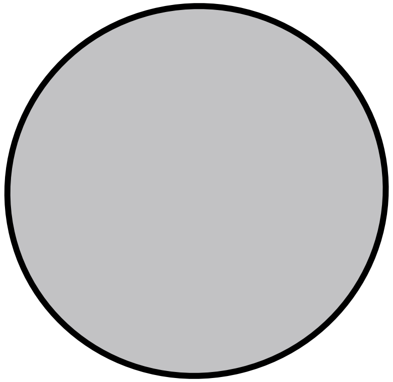
Classical Input Design Methods

Minimize the "size" of the covariance matrix of the parameter estimates:

$$\text{AsCov}(\hat{\theta}) = \frac{\lambda}{N}[0.5V''_{SI}]^{-1}$$

a function of the input signal.

Quality measures: *trace, max eigenvalue, determinant*, ..



Merge SI with the Application

If the **System Identification Set**

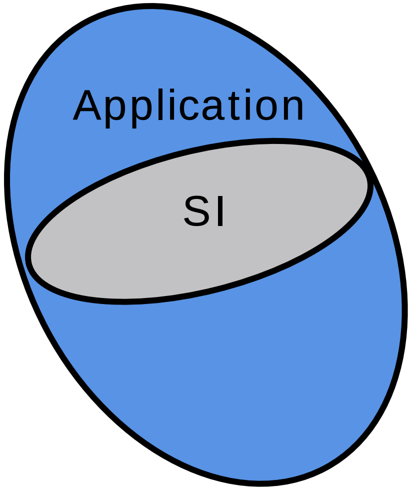
$$[\theta - \theta^o]^T V''_{SI}(\theta^o)[\theta - \theta^o] \leq 2/\kappa$$

is **inside** the **Application Set**

$$[\theta - \theta^o]^T V''_{app}(\theta^o)[\theta - \theta^o] \leq 2/\gamma$$

then $\hat{\theta}$ **satisfies** the application **specifications** w.p. α.

True if $\kappa V''_{SI}(\theta^o) \geq \gamma V''_{app}(\theta^o)$



(Matrix Inequality)

Application Oriented Input Design
Least Costly Identification Experiment

$$\underset{u}{\text{minimize}}\ \mathrm{E}\{y^2 + Cu^2\}$$

$$\text{subject to } \kappa V''_{SI}(\theta^o) \geq \gamma V''_{app}(\theta^o).$$

Minimize the output and input power in the SI experiment subject to the application constraint.

Convex optimization control problem in the input covariance function/power spectrum.

Time Domain Optimal Input Design

- Asymptotic SI properties only depends on the input spectrum/ second order statistics

- Signals are often time domain constrained.

Many possible time realization, e.g. filtered white noise

Time domain (non-convex) algorithms

Interesting ongoing research!

Model Based Optimal Input Signal Design

MOOSE2 is a model based optimal input design toolbox developed for MATLAB.

Complemented by an upcoming IEEE Control Systems Magazine article.

How much do you gain using optimal input design?

Application/controller dependent!

Conclusions

- Convex optimization for control and estimation based on the l_1 norm
- Convex optimal control for identification input design
- (Convex optimization for control and estimation based on the nuclear norm for rank constraint)

Future Challenges

- Learning to Control Dynamical Systems - how do we learn driving?
- Non-Linear Stochastic Systems - How do we assess and cope with uncertainty in autonomous driving?
- The interplay between Sensing, Planning and Control.
- Smart Transportation Systems by means of Connected Vehicles.

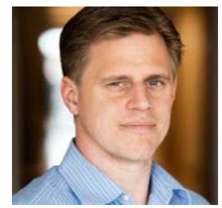
0.5 second

Thanks - This is a Team Work

Colleagues and post-doc

Cristian R. Rojas
Håkan Hjalmarsson
Jonas Mårtensson
Mohammadreza Malek Mohammadi






Former and current students

Mariette Annergren
Niclas Blomberg
Afrooz Ebadat
Per Hägg
Robert Mattila
Pedro Russo De Almeida Lima
Johan Ottersten