OSE SEMINAR 2012

System identification in the presence of trends and outliers

AMIR SHIRDEL

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
AT ÅBO AKADEMI UNIVERSITY

ÅBO NOVEMBER 29 2012

OSE OPTIMIZATION AND SYSTEMS ENGINEERING

Åbo Akademi University

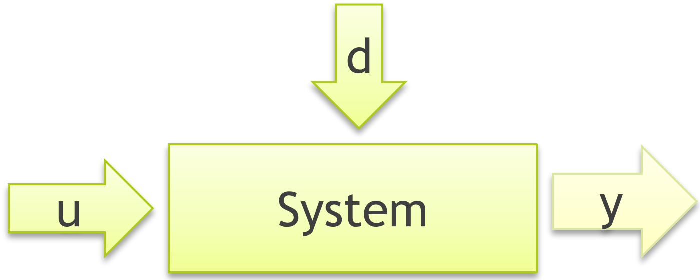
Background

- System identification is difficult when process measurements are corrupted by structured disturbances, such as trends, outliers, level shifts
- Standard approach: removal by data preprocessing
- However: difficult to separate between the effects of known system inputs and unknown disturbances (trends, etc.)

Present contribution

- Identification of system model and disturbances simultaneously
- Sparse representations of structured disturbances
- Sparse optimization used to system identification problem
- Simulated examples demonstrate that system model and structured disturbances can be identified simultaneously

System Identification



Model: $$y(k) = \varphi(k)^T \theta + d(k)$$

where $$\varphi(k)^T = [y(k\text{-}1),\ldots,y(k\text{-}r),u(k\text{-}1),\ldots,u(k\text{-}r)]$$

Identification problem: given measured input-output data $\{u(k),y(k)\}$, determine model parameter $\hat{\theta}$ by minimizing sum of squared errors

$$\sum (y(k) - \varphi(k)^T \hat{\theta})^2$$

Problem formulation

We consider a linear system:

$$y_0(k) = a_1 y_0(k-1) + \cdots + a_n y_0(k-n) + b_1 u(k-1) + \cdots + b_m u(k-m) + e(k)$$

where is *e(k)* random noise disturbance. It is assumed the measured output is given by

$$y(k) = y_0(k) + d(k)$$

where d(k) is a structured disturbance:

- outlier signal,
- level shifts,
- piecewise constant trends

Disturbance models

Sequence of outliers:

$$d_0(k) = \begin{cases} d_i, & k = k_i, i = 1, \ldots, M_0 \\ 0, & \text{otherwise} \end{cases}$$

Level shifts:

$$d_1(k) = d_i, \;\; k_i \leq k < k_{i+1}, i = 1, \ldots, M_1$$

Sequence of trends:

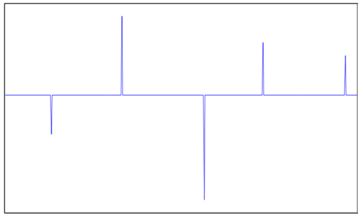
$$d_2(k) = d_2(k-1) + \beta_i, k_i \leq k < k_{i+1}, i = 1, \ldots, M_2$$

Sparse representation of disturbance

For the structured disturbances, the vectors $D_i d$ are sparse,

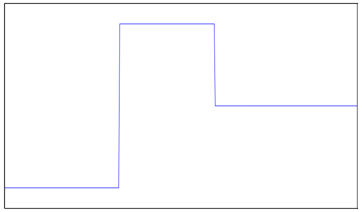
where $d = [d(1) \cdots d(N)]^{\mathrm{T}}$ and D_i depends on disturbance

Outliers: $D_0 = I$



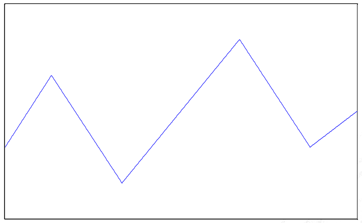
Level shifts:
$D_1 = diag([1 \quad -1])$

$$D_1 = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix}$$



Trends:
$D_2 = diag([1 \quad -2 \quad 1])$

$$D_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}$$



Sparse optimization approach

Identification by sparse optimization:

$$\min_{\hat{\theta},\hat{d}} \sum_k (y(k) - \hat{y}(k))^2$$

subject to

$$\left\| D_i d \right\|_0 \le M$$

where

$$\| \cdot \|_0 = \text{number of nonzero elements}$$

This is an intractable combinatorial optimization problem.

Instead we use $l_1 - \text{relaxation}$ and solve the convex problem

$$\min_{\hat{\theta},\hat{d}} \sum_k (y(k) - \hat{y}(k))^2 + \lambda \left\| D_i \hat{d} \right\|_1$$

Problem solution

Combining linear system and disturbances, gives

$$\begin{aligned} y(k) &= a_1\Big(y(k-1)-d(k-1)\Big)+\cdots+a_n\Big(y(k-n)-d(k-n)\Big)+d(k) \\ &\quad +b_1u(k-1)+\cdots+b_mu(k-m)+e(k) \end{aligned}$$

or

$$y(k) = \theta_a\varphi_y(k)+\theta_b\varphi_u(k)-\theta_a\varphi_d(k)+d(k)+e(k)$$

where

$$\begin{aligned} \theta_a &= [a_1 \cdots a_n]^{\mathrm{T}} \\ \theta_b &= [b_1 \cdots b_m]^{\mathrm{T}} \end{aligned}$$

$$\begin{aligned} \varphi_y(k) &= [y(k-1) \cdots y(k-n)]^{\mathrm{T}} \\ \varphi_u(k) &= [u(k-1) \cdots u(u-m)]^{\mathrm{T}} \\ \varphi_d(k) &= [d(k-1) \cdots d(k-n)]^{\mathrm{T}} \end{aligned}$$

Iterative reweighting

Cost function:

$$J(\theta_a,\theta_b,d)=\sum_{k=1}^{N}\Big(y(k)-\theta_a\varphi_y(k)-\theta_b\varphi_u(k)+\theta_a\varphi_d(k)-d(k)\Big)^2+\lambda\,\|D_i d\|_1$$

Contains bilinear term $\theta_a\varphi_d(k)$. Minimization using Bilinear Matrix Inequalities

Solution of sparse optimization problem by iterative reweighting:

Step 1. Minimize the weighted cost to give the estimates $\hat{\theta}_a$, $\hat{\theta}_b$ and $\hat{d}\,(k)$,

$$J_1(\hat{\theta}_a,\hat{\theta}_b,d)=\sum_{k=1}^{N}\Big(y(k)-\hat{\theta}_a\varphi_y(k)-\hat{\theta}_b\varphi_u(k)+\hat{\theta}_a\varphi_d(k)-\hat{d}(k)\Big)^2+\lambda\,\|W D_i\hat{d}\|_1$$

Step 2. Calculate new weights *W* and go to *step 1*. Continue until get the good convergance.

$$W=diag\Big(\frac{1}{\varepsilon+|D_i d(k)|}\Big)$$

Examples

We apply the proposed identification detrending method to the ARX model

$$y_0(k) = a_1 y_0(k-1) + a_2 y_0(k-2) + b_1 u(k-1) + b_2 u(k-2) + e(k)$$

$$y(k) = y_0(k) + d(k)$$

with parameter vector

$$\theta = [a_1 \quad a_2 \quad b_1 \quad b_2]^T$$

given by

$$\theta \quad = \quad [1.50 \quad -0.7 \quad 1.00 \quad 0.5]^T,$$

u(k) and e(k) are normally distributed signals with variances 1 and 0.1, and d(k) is unknown structured disturbance ,

Example1

Structured disturbance consisting of outliers (spikes) and piecewise trends:

$$d(k) = d_0(k) + d_2(k)$$

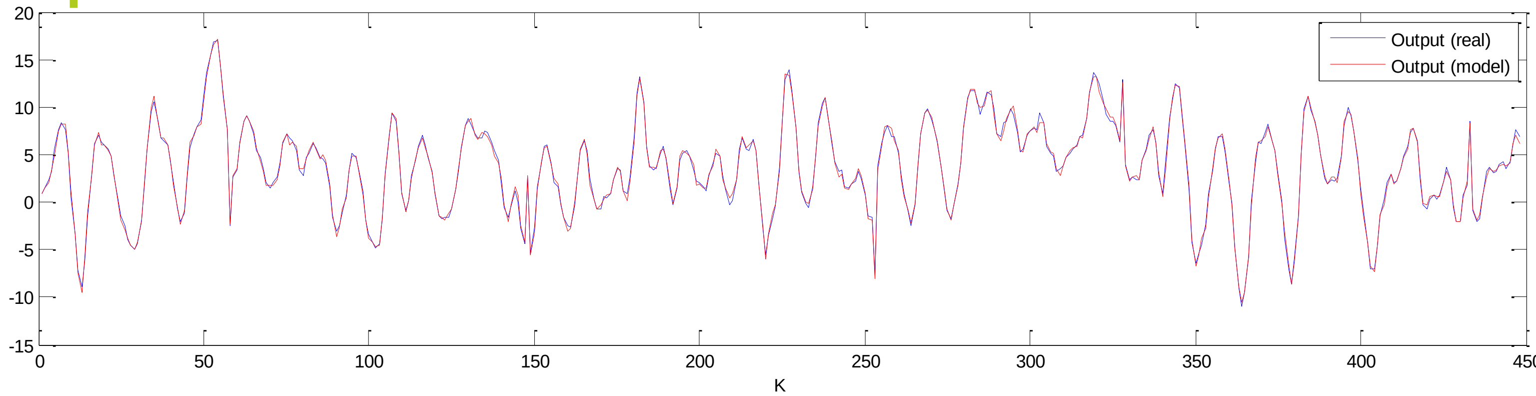
Outliers:

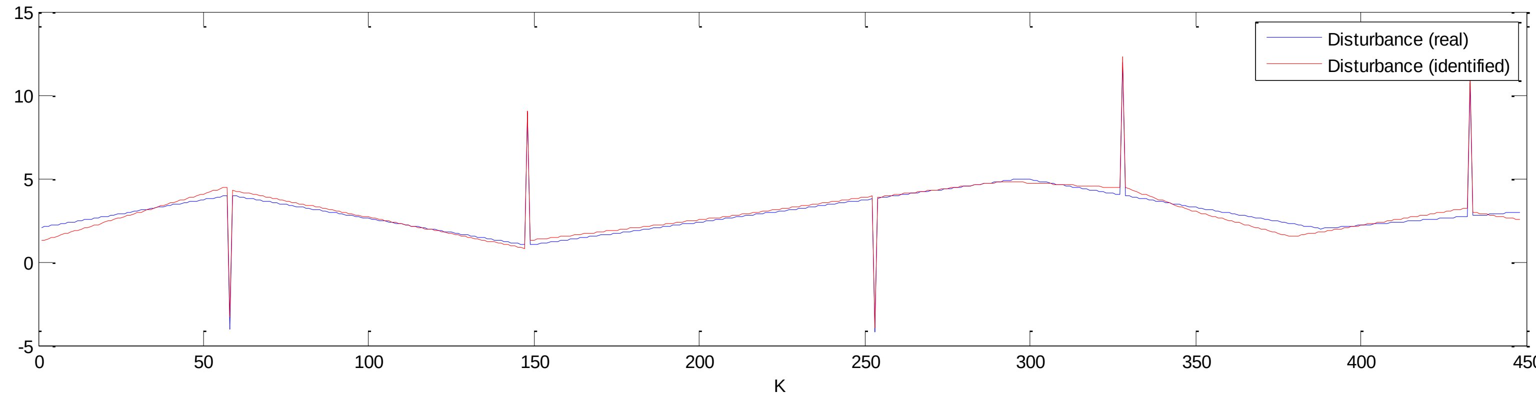
$$d_0(k) = \begin{cases} 8, & k = 150, 330, 435 \\ -8, & k = 60, 255 \\ 0, & \text{otherwise} \end{cases}$$

Piecewise trends:

$$d_2(k) = \begin{cases} 2 \ldots 4, & \text{k=1-60} \\ 4 \ldots \text{-}2, & \text{k=61-150} \\ \text{-}2 \ldots 7, & \text{k=151-300} \\ 7 \ldots 2, & \text{k=301-390} \\ 2 \ldots 3, & \text{k=391-450} \end{cases}$$

Example1





θ	1.5000	-0.7000	1.0000	0.5000	RMSE = 0.3296
$\hat{\theta}$	1.4902	-0.6907	1.0000	0.5152	RMSE = 0.3246
$\hat{\theta}_{LS}$	1.2247	-0.3063	1.0192	0.8883	RMSE = 1.6500

Example2,3,4

Example 2: structured disturbance consisting of outliers (spikes) and level shifts:

$$d(k) = d_0(k) + d_1(k)$$

Outliers:

$$d_0(k) = \begin{cases} 8, & k = 150, 330, 435 \\ -8, & k = 60, 255 \\ 0, & \text{otherwise} \end{cases}$$

Level shifts:

$$d_1(k) = \begin{cases} -3, & \text{k=1-150} \\ 5, & \text{k=151-250} \\ 1, & \text{k=251-450} \end{cases}$$

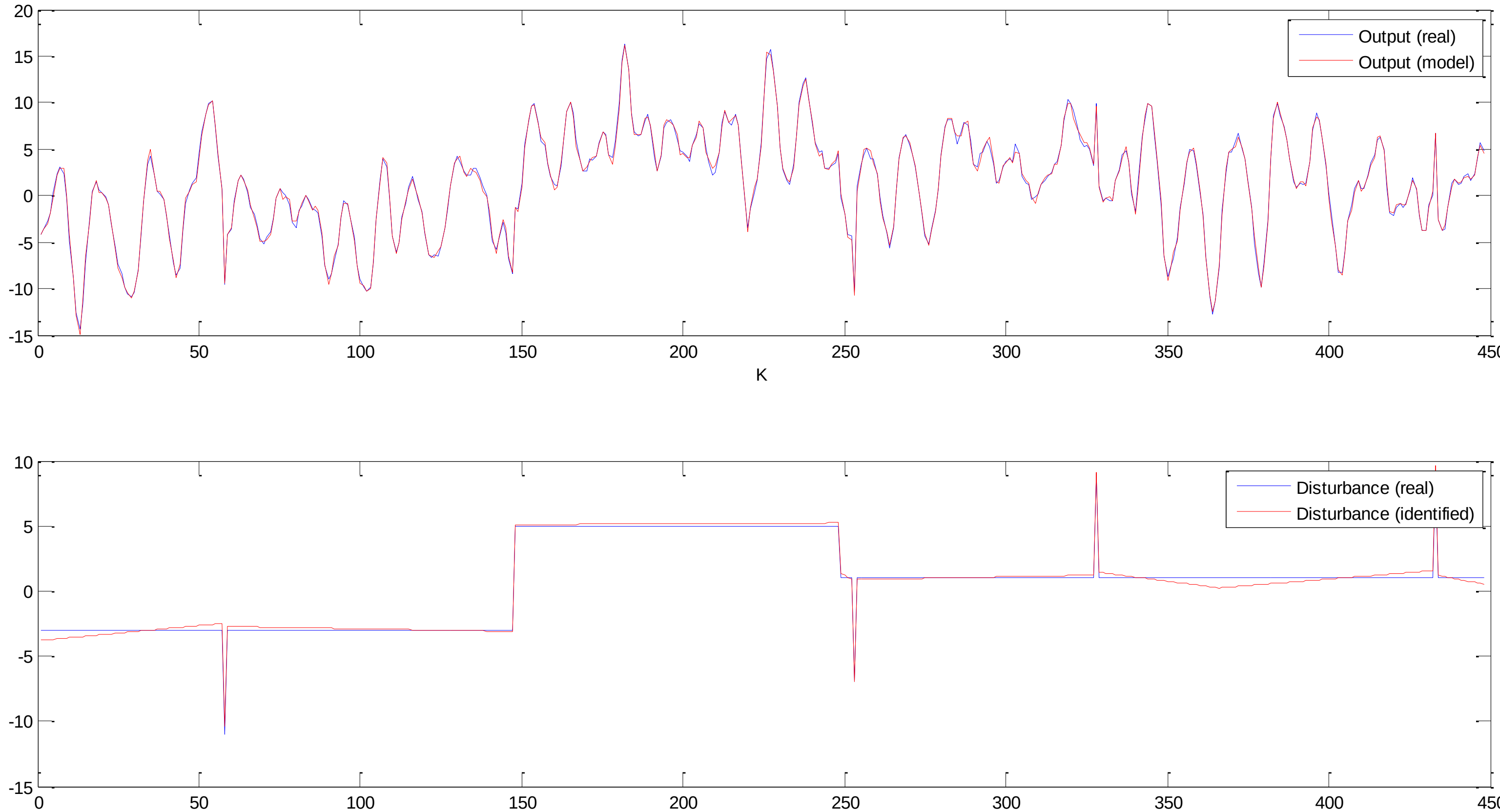
Example 3: structured disturbance consists of level shifts and trends

$$d(k) = d_1(k) + d_2(k)$$

Example 4: structured disturbance consists of outliers, level shifts and trends

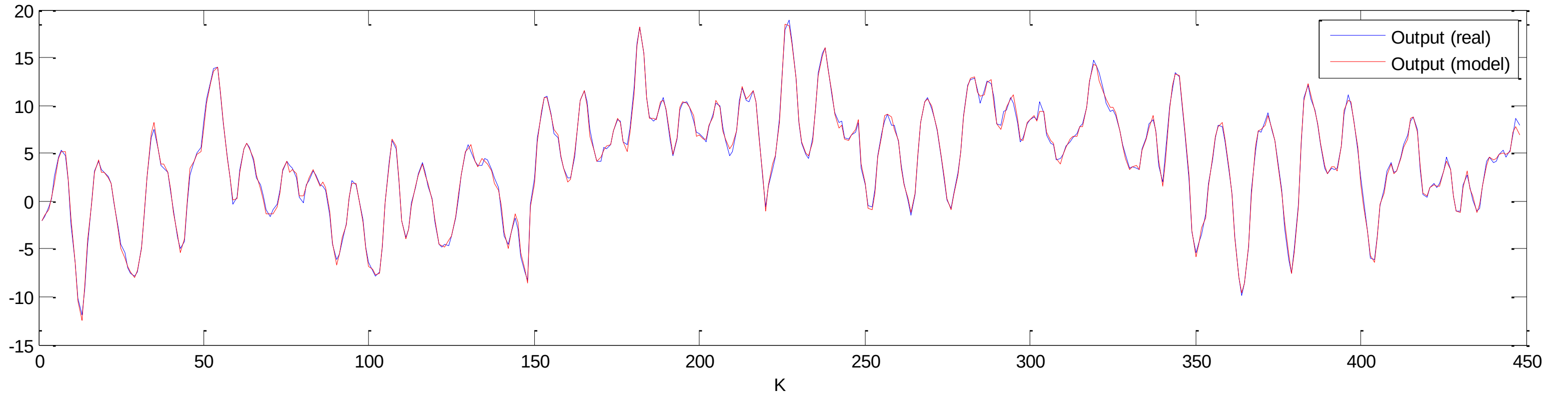
$$d(k) = d_0(k) + d_1(k) + d_2(k)$$

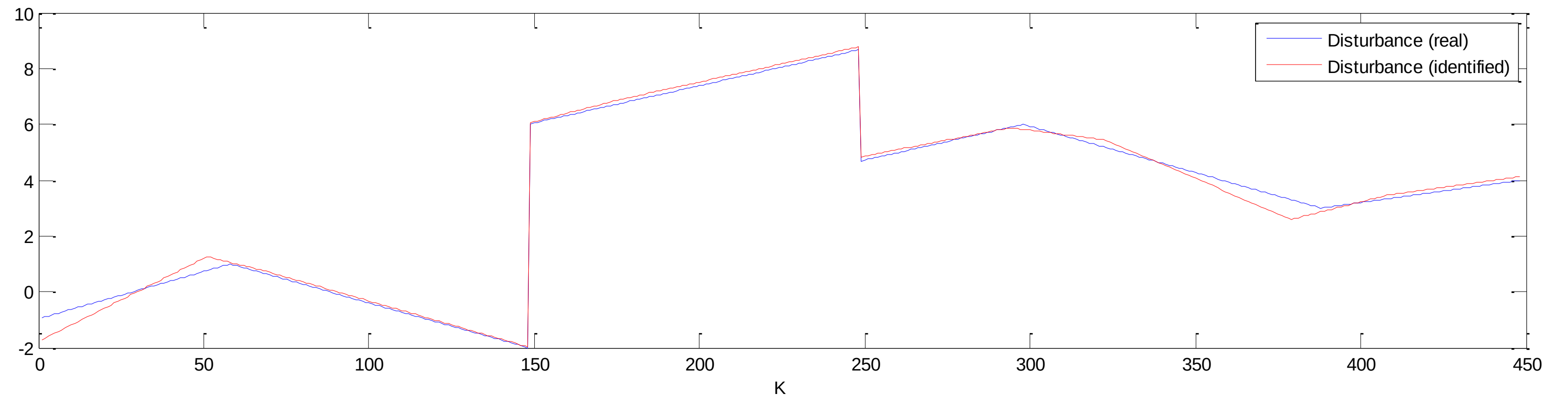
Example2



θ	1.5000	-0.7000	1.0000	0.5000	$RMSE = 0.3296$
$\hat{\theta}$	1.4929	-0.6950	1.0008	0.5086	$RMSE = 0.3222$
$\hat{\theta}_{LS}$	1.2373	-0.3872	0.9611	0.7965	$RMSE = 1.5318$

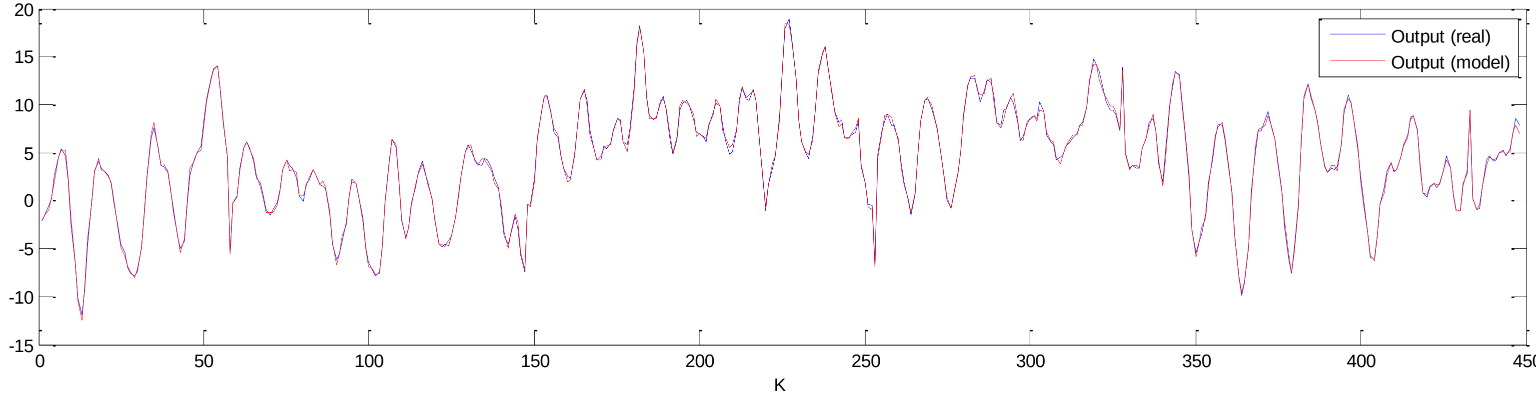
Example3

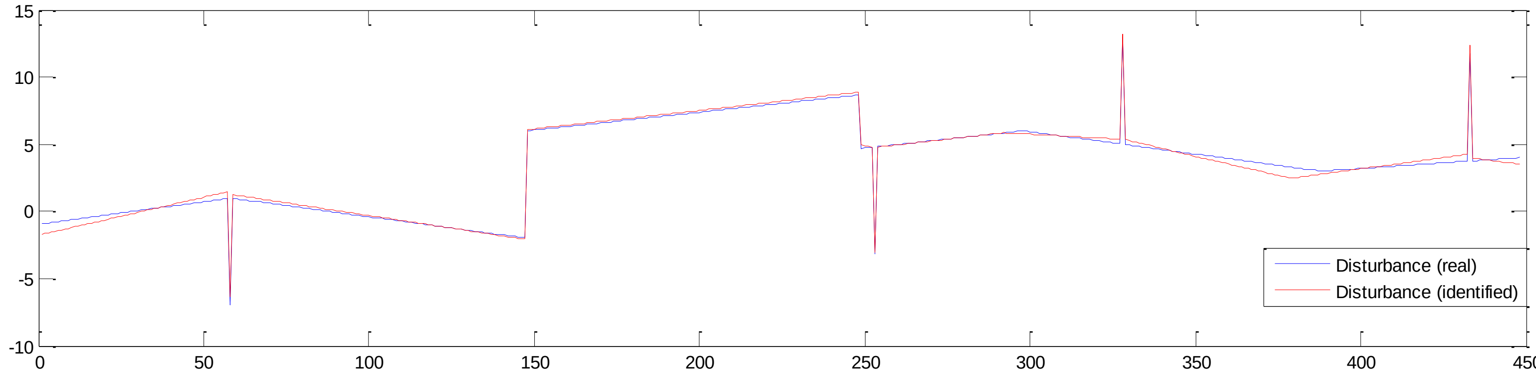




θ	1.5000	-0.7000	1.0000	0.5000	RMSE = 0.3296
$\hat{\theta}$	1.4969	-0.6969	1.0008	0.5078	RMSE = 0.3296
$\hat{\theta}_{LS}$	1.5244	-0.6203	1.0461	0.5013	RMSE = 0.8283

Example4





θ	1.5000	-0.7000	1.0000	0.5000	$RMSE = 0.3296$
$\hat{\theta}$	1.4950	-0.6955	1.0007	0.5093	$RMSE = 0.3271$
$\hat{\theta}_{LS}$	1.2563	-0.3624	0.9996	0.8165	$RMSE = 1.5732$

Summary:

We have presented a method for identification of linear systems in the presence of structured disturbances (outliers, level shifts and trends) by using sparse optimization. The method gives acceptable results for simultaneous identification of the disturbances and the system parameters.

Future work:

- finding trends in nonlinear system and more general disturbances
- real world applications.

Thank you for your attention!

Questions?