Identification-Based Nonlinear Moving-Horizon Observers

Mazen Alamir

† *French National Research Center (CNRS)*
Gipsa-lab, Control Systems Department. University of Grenoble, France

Back to the origins

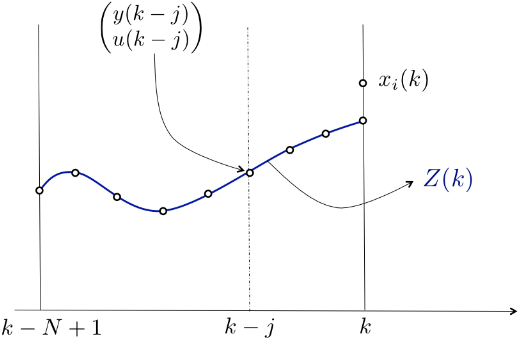
Observability

$\exists\ n$ maps F_i, $i \in \{1,\dots,n\}$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big)$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$



$\Rightarrow$ Identify the maps F_i

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

$$\hat{x} = F\big(Z, \hat{x}^{-}\big)$$

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

$$\hat{x} = F\big(Z, \hat{x}^-\big)$$

1. Generate a grid of $\mathbf{N}$ values $(Z, \hat{x}^-)^{(j)} \in \mathbb{R}^{N(n_u+n_y)} \times \mathbb{R}^n$

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

$$\hat{x} = F\big(Z, \hat{x}^-\big)$$

1. Generate a grid of $\mathbf{N}$ values $(Z, \hat{x}^-)^{(j)} \in \mathbb{R}^{N(n_u+n_y)} \times \mathbb{R}^n$
2. Form the resulting set of N NL-MHO Programming Problems

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

$$\hat{x} = F\big(Z, \hat{x}^-\big)$$

1. Generate a grid of $\mathbf{N}$ values $(Z, \hat{x}^-)^{(j)} \in \mathbb{R}^{N(n_u+n_y)} \times \mathbb{R}^n$
2. Form the resulting set of N NL-MHO Programming Problems
3. Solve them to obtain the N values of $\hat{x}^{(j)}$

Related works

Related works

[**Alessandri et al**. Moving-Horizon State Estimation For NL Discrete-Time Systems: New Stability Results and Approximation Schemes, **Automatica**, **2008**]

[**Alessandri et al**. Movinh-Horizon State Estimation For Nonlinear Systems using Neural Network. **IEEE Trans. Neural Networks** 22(5): 768-780, **2011**]

$$\hat{x} = F\big(Z, \hat{x}^-\big)$$

1. Generate a grid of $\mathbf{N}$ values $(Z, \hat{x}^-)^{(j)} \in \mathbb{R}^{N(n_u+n_y)} \times \mathbb{R}^n$
2. Form the resulting set of N NL-MHO Programming Problems
3. Solve them to obtain the N values of $\hat{x}^{(j)}$
4. Identify a **NN-model** F based on the LS problem:

$$\left\{\hat{x}^{(j)}, (Z, \hat{x}^-)^{(j)}\right\}_{i=j}^{N}$$

Related works

Related works

- High number of n-**dimensional** non convex optimization.
- $\rightarrow$ Data **not reliable** (even for correct models . . . !!)
- What definition for $\hat{x}^-$?
- NN: **not suitable** for high dimensional regressors.

$$\hat{x} = F\big(Z, \hat{x}^-\big)$$

1. Generate a grid of $\mathbf{N}$ values $(Z, \hat{x}^-)^{(j)} \in \mathbb{R}^{N(n_u+n_y)} \times \mathbb{R}^n$
2. Form the resulting set of N NL-MHO Programming Problems
3. Solve them to obtain the N values of $\hat{x}^{(j)}$
4. Identify a **NN-model** F based on the LS problem:

$$\Big\{\hat{x}^{(j)}, (Z, \hat{x}^-)^{(j)}\Big\}_{i=j}^{N}$$

Pushing the idea to its limits

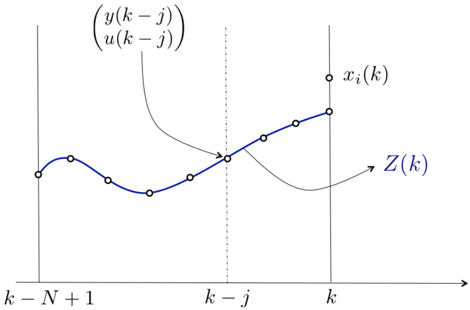
Observability

$\exists\ n$ maps F_i, $i \in \{1, \dots, n\}$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big)$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$



Pushing the idea to its limits

Observability

$\exists\ m < n$ maps F_i, $i \in I$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big)$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$

- **Partial observers**
  (dedicated to components)

Pushing the idea to its limits

Observability

$\exists\ m < n$ maps F_i, $i \in I$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big)$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$

- **Partial observers**
  (dedicated to components)
- **Supervision-oriented observers**
  (Dedicated to risk-relevant functions)

Pushing the idea to its limits

Observability

$\exists\ n_u$ maps F_i, $i \in \{1, \dots, n_u\}$

$$\mathbf{k_i}(x(k)) \approx F_i\Big(Z(k)\Big)$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$

- **Partial observers**
  (dedicated to components)
- **Supervision-oriented observers**
  (Dedicated to risk-relevant functions)
- **Control-oriented observer**
  (Dedicated to the $\mathbf{k}(\cdot)$ components)

Pushing the idea to its limits

Observability

$\exists\ n$ maps F_i, $i \in I$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big) \quad \forall x(k) \in \mathbb{X}$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$

- **Partial observers**
  (dedicated to components)
- **Supervision-oriented observers**
  (Dedicated to risk-relevant functions)
- **Control-oriented observer**
  (Dedicated to the $\mathbf{k}(\cdot)$ components)
- **Context-relevant observers**
  (Dedicated to relevant domains)

Road Map

Observability

$\exists\ n$ maps F_i, $i \in I$ s.t.

$$x_i(k) \approx F_i\Big(Z(k)\Big) \quad \forall x(k) \in \mathbb{X}$$

where

$$Z(k) := \begin{pmatrix} y(k) \\ u(k) \\ \vdots \\ y(k-N+1) \\ u(k-N+1) \end{pmatrix}$$

1. Choose $(\mathbb{X}, \mathbb{U})$ of interest.
2. **Simulate the noisy** model over an $(\mathbb{X}, \mathbb{U})$-grid to obtain the id-data:
   $$\Big\{x^{(j)}, Z(x^{(j)}, u^{(j)})\Big\}_{j=1}^{N}$$
3. $\forall i$, find F_i that solves the LS problem **related to the component** i, i.e.
   $$\min_{F_i \in \mathbb{F}} \sum_{j=1}^{N} \Big\| x_i^{(j)} - F_i(Z^{(j)}) \Big\|^2$$

[Noise trade-off is implicitly learned !]

The Class of Nonlinear Relationships

Problem Statement

Find a sampling period, an integer N and a NL map F such that:

$$\bar{r}(k) \approx F(Z(k)) \in \mathbb{R}$$

Interest is focused on the following class:

$$F(Z) = \Gamma(Z^T L) \quad ; \quad \Gamma(\cdot) \text{ strictly increasing} \tag{1}$$

The Class of Nonlinear Relationships

Problem Statement

Find a sampling period, an integer N and a NL map F such that:

$$\bar{r}(k) \approx F(Z(k)) \in \mathbb{R}$$

Interest is focused on the following class:

$$F(Z) = \Gamma(Z^T L) \quad ; \quad \Gamma(\cdot) \text{ strictly increasing} \tag{1}$$

since in this case, (1) leads to the following problem:

$$\Gamma^{-1}(\bar{r}) \approx Z^T L \tag{2}$$

The Class of Nonlinear Relationships

Problem Statement

Find a sampling period, an integer N and a NL map F such that:

$$\bar{r}(k) \approx F(Z(k)) \in \mathbb{R}$$

Interest is focused on the following class:

$$F(Z) = \Gamma(Z^T L) \quad ; \quad \Gamma(\cdot) \text{ strictly increasing} \tag{1}$$

since in this case, (1) leads to the following problem:

$$\Gamma^{-1}(\bar{r}) \approx Z^T L \tag{2}$$

and by choosing $\Gamma^{-1}(\bar{r}) := B(\bar{r})\mu$, the problem becomes:

$$(Z^T, -B(\bar{r}))\begin{pmatrix} L \\ \mu \end{pmatrix} \approx 0 \quad ; \quad \text{(A linear least squares problem)}$$

The Class of Nonlinear Relationships

Problem Statement

Find a sampling period, an integer N and a NL map F such that:

$$\bar{r}(k) \approx F(Z(k)) \in \mathbb{R}$$

Interest is focused on the following class:

$$F(Z) = \Gamma(Z^T L) \quad ; \quad \Gamma(\cdot) \text{ strictly increasing} \tag{1}$$

since in this case, (1) leads to the following problem:

$$\Gamma^{-1}(\bar{r}) \approx Z^T L \tag{2}$$

and by choosing $\Gamma^{-1}(\bar{r}) := B(\bar{r})\mu$, the problem becomes:

$$(Z^T, -B(\eta(\bar{r}))) \begin{pmatrix} L \\ \mu \end{pmatrix} \approx 0 \quad ; \quad \text{(A linear least squares problem)}$$

A Constrained QP Identification Problem

Consider the following constrained **LLS** problem:

$$\min_{L,\mu} \sum_{k\in\mathbb{K}} \|Z^T(k)L - B(\eta(k))\mu\|^2 \tag{3}$$

under the following **constraints**:

1. Γ is strictly increasing, namely:

$$\forall \eta \in [0,1], \quad \Big[\frac{dB}{d\eta}(\eta)\Big]\mu \geq \varepsilon \tag{4}$$

2. Normalization constraint

$$\Big[\int_0^1 B(\eta)d\eta\Big] \cdot \mu = \frac{1}{2}\Big[\bar{r}_{min} + \bar{r}_{max}\Big] \tag{5}$$

Inequality (4) $\Leftrightarrow A\mu \leq B$ using a grid $0 = \eta_1 < \eta_2 < \dots < \eta_r = 1$

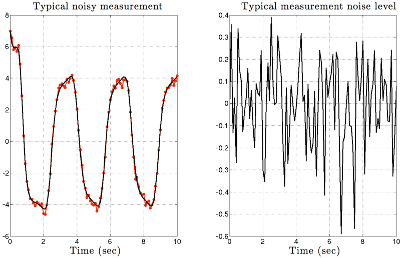
Related Works

- Wiener structure in System Identification literature
- Work by A. Hagenblad and L. Ljung [1999]
  - Parametrization of Γ^{-1}
  - QP solution as initial guess for descent methods
  - No monotonicity enforcement
- Recent work by A. Wills and L. Ljung [2010]
  - maximum likelihood method
  - Descent method / gradient computation
  - Biased/Unbiased ...
- ...

The Van der Pol Oscillator

$$\dot{x}_1 = -x_2$$
$$\dot{x}_2 = 4x_1 + (1 - x_1^2)x_2$$
$$y = x_1 + x_2 + w$$

- $\tau = 0.1$, $\sigma = 0.2$
- $\mathbb{X} = [-2, +2] \times [-5, +5]$
- $n_g = 4^2 = 16$ grid points
- 10 sec simulation
- $card$(learning set)$= 1456$
- $N = 10$, $n_\mu = 10$
- Compression ratio $= 73$

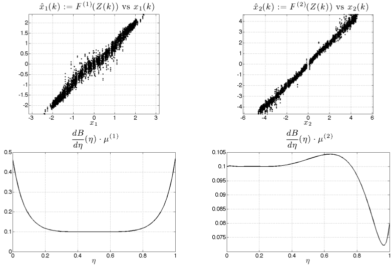


Typical noisy scenario

The Van der Pol Oscillator

$$
\begin{aligned}
\dot{x}_1 &= -x_2 \\
\dot{x}_2 &= 4x_1 + (1 - x_1^2)x_2 \\
y &= x_1 + x_2 + w
\end{aligned}
$$

- $\tau = 0.1$, $\sigma = 0.2$
- $\mathbb{X} = [-2, +2] \times [-5, +5]$
- $n_g = 4^2 = 16$ grid points
- 10 sec simulation
- $card$(learning set)$= 1456$
- $N = 10$, $n_\mu = 10$
- Compression ratio $= 73$

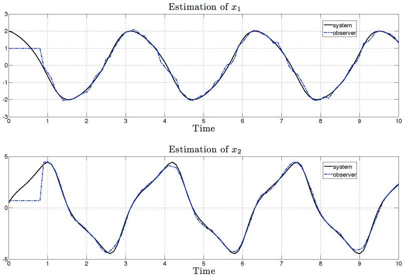


Identification results

The Van der Pol Oscillator

$$\begin{aligned}\dot{x}_1 &= -x_2 \\ \dot{x}_2 &= 4x_1 + (1 - x_1^2)x_2 \\ y &= x_1 + x_2 + w\end{aligned}$$

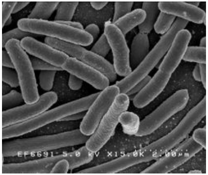
- $\tau = 0.1$, $\sigma = 0.2$
- $\mathbb{X} = [-2, +2] \times [-5, +5]$
- $n_g = 4^2 = 16$ grid points
- 10 sec simulation
- $card$(learning set)$= 1456$
- $N = 10$, $n_\mu = 10$
- Compression ratio $= 73$



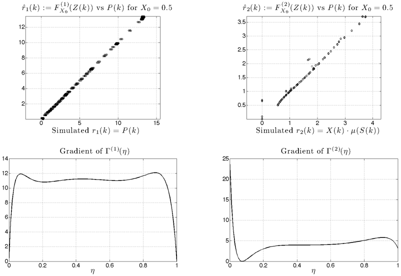
Typical validation scenario

E-coli Strain State Estimation Problem

$$
\begin{aligned}
\dot{X} &= \mu(S)X - k_d \exp(-\frac{k_p}{P})X \\
\dot{S} &= -y_s \mu(S)X - k_m X \\
\dot{P} &= y_p \mu(S)\frac{I}{I+k_I}X - k_d \exp(-\frac{k_p}{P})P \\
y &= (X + w_1, L + w_2) \\
L &= y_l \cdot \mu(S)\frac{I}{I+k_I}XP
\end{aligned}
$$



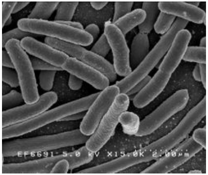
- $\mathbb{X}(X_0) := \{X_0\} \times [0,5] \times [0,0.3]$
- $\tau = 0.2$, $N = 3$, $n_\mu = 10$, $n_g = 25$
- Compression ratio ≈ 58
- $\bar{r}_1 = P$, $\bar{r}_2 = \mu(S) \cdot X$



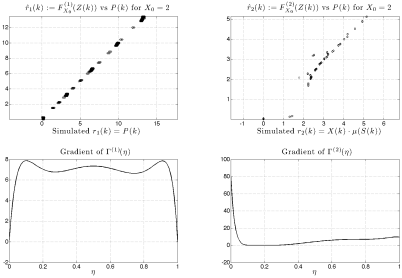
Identification results for $X_0 = 0.5$

E-coli Strain State Estimation Problem

$$
\begin{aligned}
\dot{X} &= \mu(S)X - k_d \exp(-\frac{k_p}{P})X \\
\dot{S} &= -y_s\mu(S)X - k_m X \\
\dot{P} &= y_p\mu(S)\frac{I}{I+k_I}X - k_d \exp(-\frac{k_p}{P})P \\
y &= (X + w_1, L + w_2) \\
L &= y_l \cdot \mu(S)\frac{I}{I+k_I}XP
\end{aligned}
$$



- $\mathbb{X}(X_0) := \{X_0\} \times [0,5] \times [0,0.3]$
- $\tau = 0.2,\ N = 3,\ n_\mu = 10,\ n_g = 25$
- Compression ratio ≈ 58
- $\bar{r}_1 = P,\ \bar{r}_2 = \mu(S) \cdot X$



Identification results for $X_0 = 2.0$

E-coli Strain State Estimation Problem

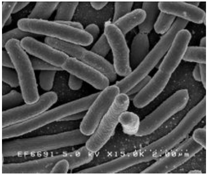
$$\dot{X} = \mu(S)X - k_d \exp(-\frac{k_p}{P})X$$

$$\dot{S} = -y_s\mu(S)X - k_m X$$

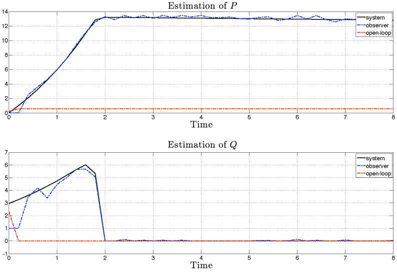
$$\dot{P} = y_p\mu(S)\frac{I}{I+k_I}X - k_d \exp(-\frac{k_p}{P})P$$

$$y = (X + w_1, L + w_2)$$

$$L = y_l \cdot \mu(S)\frac{I}{I+k_I}XP$$



- $\mathbb{X}(X_0) := \{X_0\} \times [0,5] \times [0,0.3]$
- $\tau = 0.2,\ N = 3,\ n_\mu = 10,\ n_g = 25$
- Compression ratio ≈ 58
- $\bar{r}_1 = P,\ \bar{r}_2 = \mu(S) \cdot X$



Typical validation scenario (Run 1)

E-coli Strain State Estimation Problem

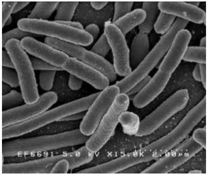
$$\dot{X} = \mu(S)X - k_d \exp(-\frac{k_p}{P})X$$

$$\dot{S} = -y_s \mu(S)X - k_m X$$

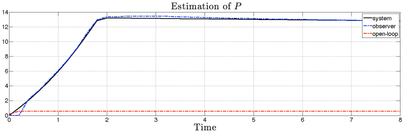
$$\dot{P} = y_p \mu(S)\frac{I}{I+k_I}X - k_d \exp(-\frac{k_p}{P})P$$

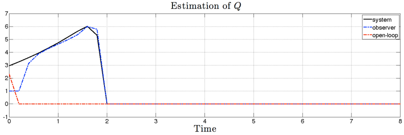
$$y = (X + w_1, L + w_2)$$

$$L = y_l \cdot \mu(S)\frac{I}{I+k_I}XP$$



- $\mathbb{X}(X_0) := \{X_0\} \times [0,5] \times [0,0.3]$
- $\tau = 0.2$, $N = 3$, $n_\mu = 10$, $n_g = 25$
- Compression ratio ≈ 58
- $\bar{r}_1 = P$, $\bar{r}_2 = \mu(S) \cdot X$

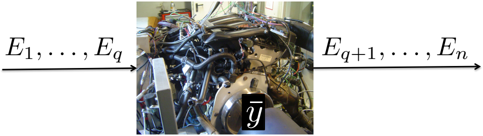




Typical validation scenario (Run 2)

Diesel Engine Emission Estimation

[Coll. D. Alberer and L. Del Re (Johannes Kepler Univ, Linz, Austria)]



- E_i : on-line available measurements (E for ECU)
- $\bar{y}$: emission-related unmeasured quantity (NOx, Opacity)
- $E := (E_1, \ldots, E_n)$
- $Z(k) = (E(k-N+1), \ldots, E(k)) \in \mathbb{R}^{Nn}$

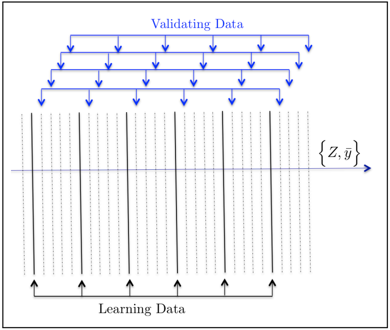
Problem Statement

Find a sampling period, an integer N and a nonlinear map F such that:

$$\bar{y}(k) \approx F(Z(k))$$

Learning/validation data

- BMW M47TUE Diesel Engine (UJK, Linz Test-Bed)
- Acquisition rate (100 Hz)
- 10 ECU on-line available sensors (Boch)
- ≈ 1371600 samples (23 min)



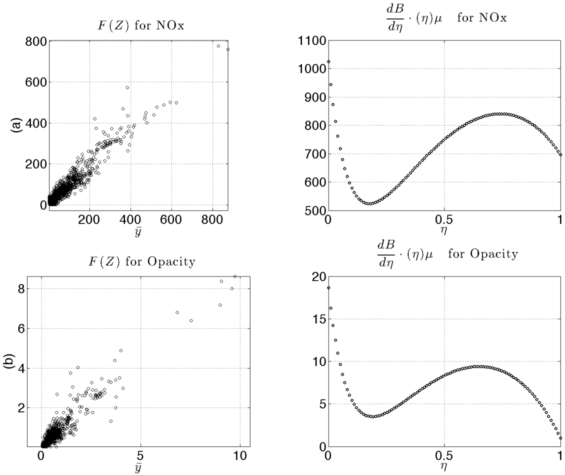
- Under sampling $m = 100$
- $Card(\text{Learning set}) \approx 13716$

> *The learning set is ≈ 100 times smaller than the validating set.*

Parameters of the solution

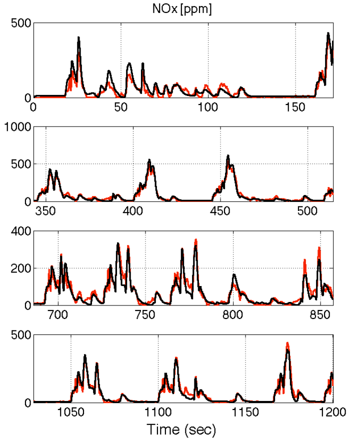
- $\tau = 0.01\ s$, $\mathbf{m = 100}$, $\mathbf{N = 5}$
  - Data over the past 5 seconds are used. (seem to be needed)
  - $Z \in \mathbb{R}^{50}$
- $\mathbf{n}_\mu = \mathbf{6}$
  - $\mu \in \mathbb{R}^6$
- **The same set of parameters** have been found appropriate for the identification problems of
  - NOx
  - Opacity
- No attempt has been done to remove useless sensors.
- **56** parameters and Compression Ratio: $1370000/56 \approx \mathbf{24000}$

Learning Identification Results



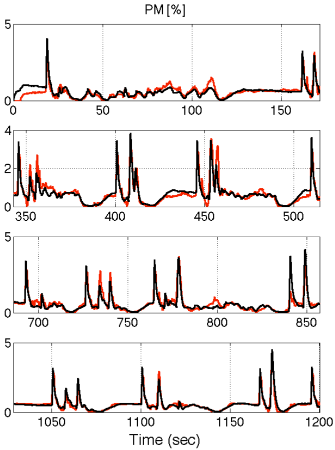
- (a) NOx
- (b) Opacity
- Note the highly nonlinear character of the maps Γ

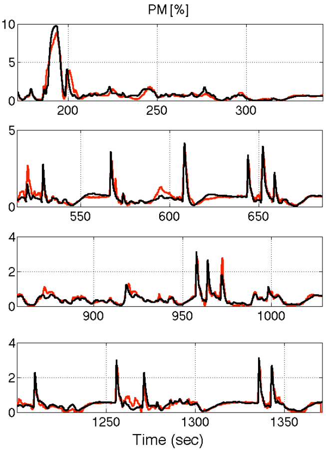
Validation Identification Results (NOx)



Evolution of NOx (Model-based prediction/Measured)

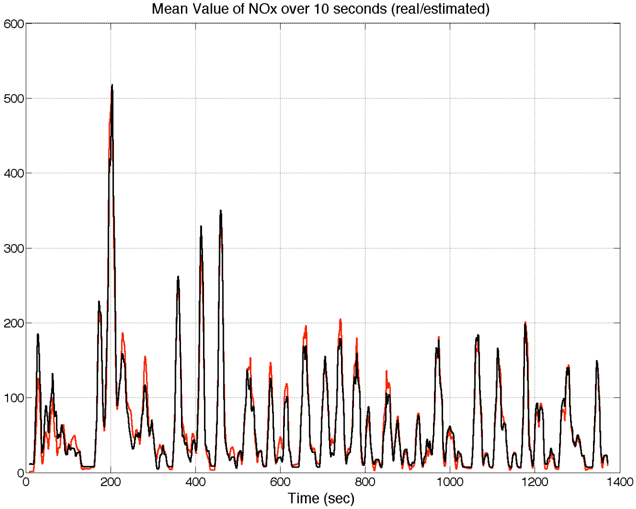
Validation Identification Results (Opacity)





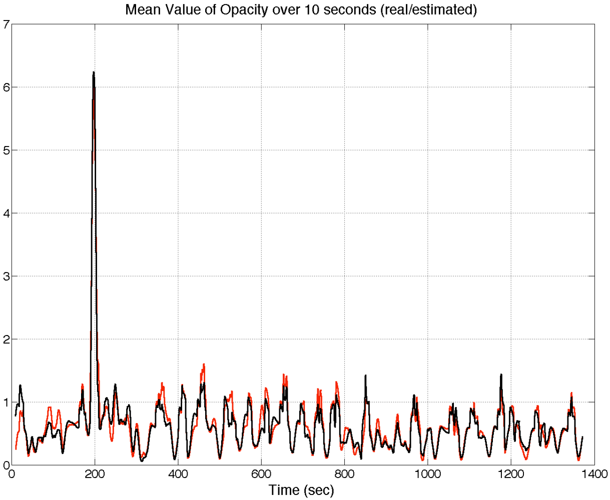
Evolution of opacity (Model-based prediction/Measured)

Validation Identification Results (NOx)



Evolution of NOx (Model-based prediction/Measured)

Validation Identification Results (Opacity)



Evolution of opacity (Model-based prediction/Measured)

Conclusion & Future Work

To summarize . . .

- NL-MHO Problem is a nonlinear high-dimensional ID problem.
- Natural NL-extension to linear identification (QP-solvable)
- Natural one-shot incorporation of classical trade-offs
- Use to reduce the dimension of the on-line optimization problem.

Conclusion & Future Work

To summarize . . .

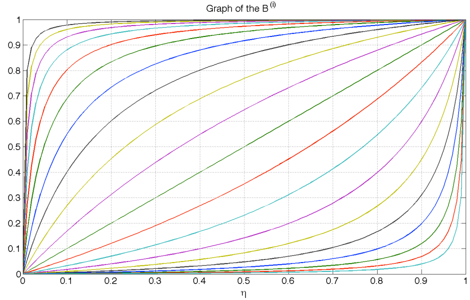
- NL-MHO Problem is a nonlinear high-dimensional ID problem.
- Natural NL-extension to linear identification (QP-solvable)
- Natural one-shot incorporation of classical trade-offs
- Use to reduce the dimension of the on-line optimization problem.

What next ?. . .

- Perform a sparse identification
  [using the QP available lagrange multipliers]

- Extension to Non Linear Parameter-Varying (NLPV . . . !) structure
  [Recall the X_0-dependent E-Coli Identifiability]

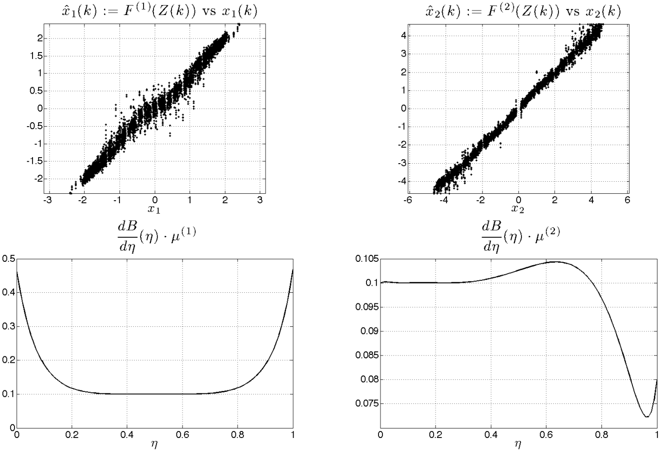
The Basis Functions $B(\eta)$



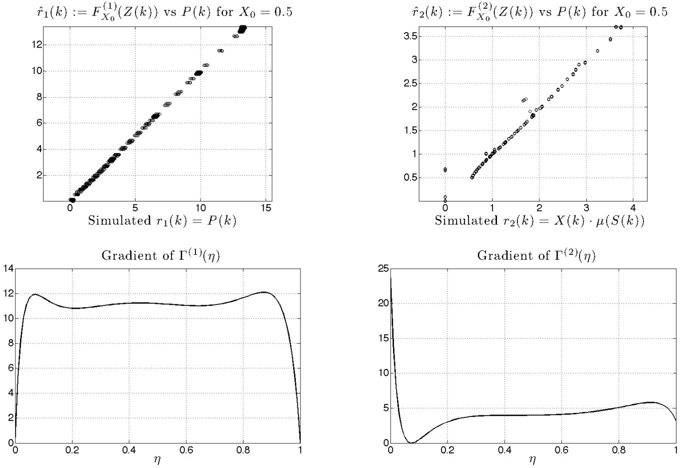
$$\left\{B^{(j)}\right\}_{j=1}^{n_b} := \left\{1\right\} \cup \left\{B_1^{(i)}\right\}_{i=1}^{n_m-1} \cup \left\{B_2^{(i)}\right\}_{i=1}^{n_m}$$

$$B_1^{(i)}(\eta) := (1+\alpha_i)\frac{\eta}{1+\alpha_i\eta} \quad ; \quad B_2^{(i)}(\eta) := \frac{\eta}{1+\alpha_i-\alpha_i\eta}$$

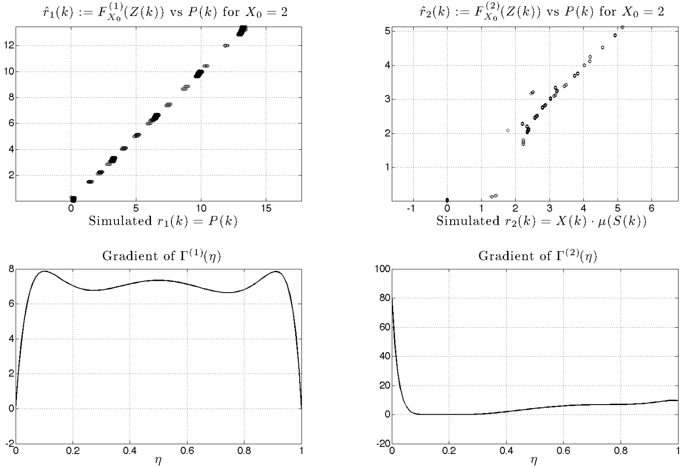
$$\alpha_i = \exp(\beta(1-i)) - 1$$



Identification results for example 1



Identification results for example 2, $X_0 = 0.5$



Identification results for example 2, $X_0 = 2.0$