Some Topics on Nonlinear Moving-Horizon Observers

Mazen Alamir

CNRS-University of Grenoble

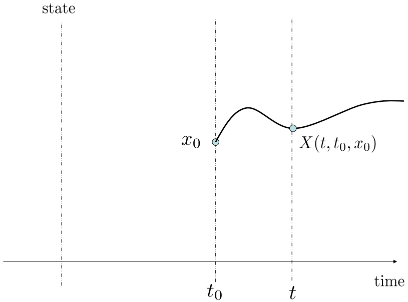
Email: mazen.alamir@grenoble-inp.fr
Web: http://www.mazenalamir.fr

Séminaire École des Mines, Mai 2010, Paris.

Moving Horizon Observer

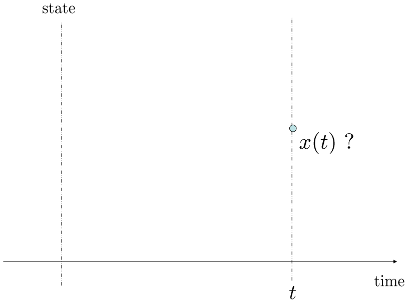
Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



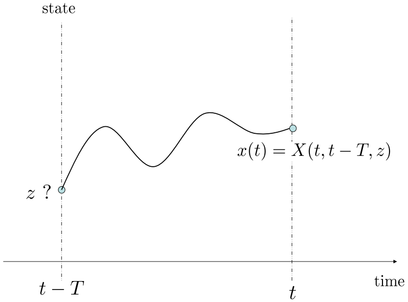
Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



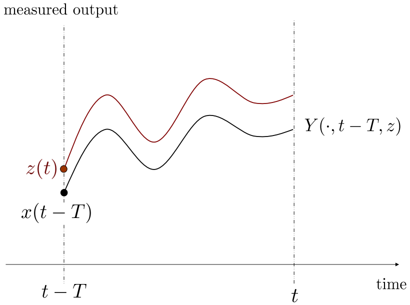
Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



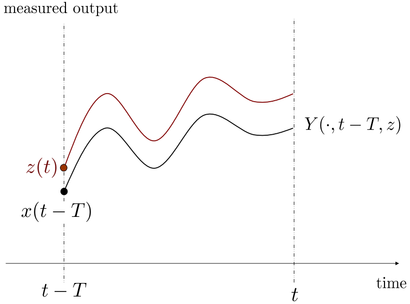
Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



Moving Horizon Observer

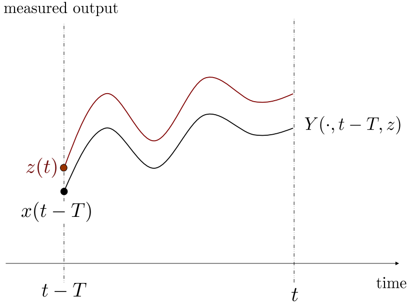
$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



$$z_{opt}(t) \leftarrow arg\min_z \quad J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$



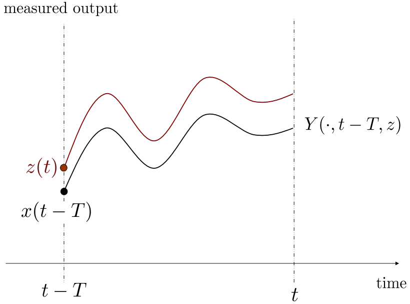
$$z_{opt}(t) \leftarrow \arg\min_z \quad J_0(z, t, y_{t-T}^t) = \int_{t-T}^{t} \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

$$\hat{x}(t) := X(t, t-T, z_{opt}(t))$$

Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$

- *Generally a non convex optimization problem*
- *many local minima*



$$z_{opt}(t) \leftarrow \arg\min_z \quad J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

$$\hat{x}(t) := X(t, t-T, z_{opt}(t))$$

Overview

State estimation: An optimization problem

Overview

State estimation: An optimization problem

(Dynamic & non convex)

Overview

State estimation: An optimization problem

(Dynamic & non convex)

Local minima

Computation time

Overview

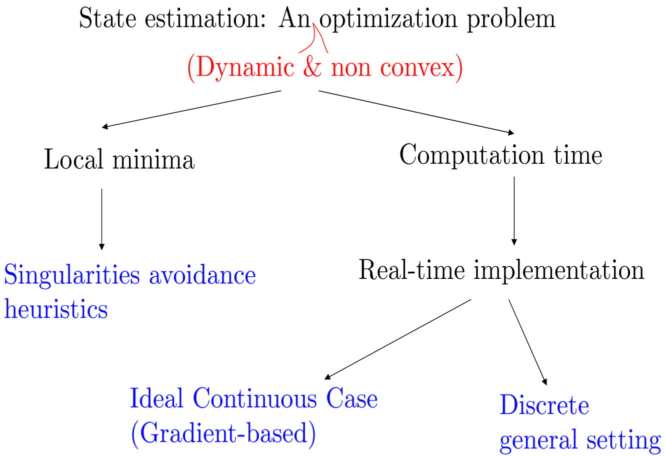
State estimation: An optimization problem

(Dynamic & non convex)

Local minima → Singularities avoidance heuristics

Computation time

Overview


State estimation: An optimization problem
(Dynamic & non convex)
Local minima
Computation time
Singularities avoidance heuristics
Real-time implementation
Ideal Continuous Case (Gradient-based)
Discrete general setting

Some definitions and notation (1): The System

(Uncertainty & noise)-free system

$$\begin{aligned} x(t) &= X(t, t_0, x_0) \\ y(t) &= h(t, x(t)) \end{aligned}$$

Uncertain and noisy system

$$\begin{aligned} x(t) &= X(t, t_0, x_0, w_{t_0}^t) \\ y(t) &= h(t, x(t)) + v(t) \end{aligned}$$

Constraints

- $x(t) \in \mathbb{X}(t) \subset \mathbb{R}^n$
- $w(t) \in \mathbb{W}(t) \subset \mathbb{R}^{n_w}$ Uncertainties/Disturbances.
- $v(t) \in \mathbb{V}(t) \subset \mathbb{R}^{n_y}$ Measurement noise

Definitions and notation (2): Measurements-compatible configurations

Consider

- Time interval $[t-T, t]$
- Measurement profile y_{t-T}^{t}
- $(\xi, \mathbf{w}) \in \mathbb{X}(t-T) \times [\mathbb{R}^{n_w}]^{[t-T,t]}$

Definitions and notation (2): Measurements-compatible configurations

Consider

- Time interval $[t-T, t]$
- Measurement profile y_{t-T}^{t}
- $(\xi, \mathbf{w}) \in \mathbb{X}(t-T) \times [\mathbb{R}^{n_w}]^{[t-T,t]}$

$(\xi, \mathbf{w})$ is (y_{t-T}^{t})-compatible

if for all $\sigma \in [t-T, t]$:

1. $w(\sigma) \in \mathbb{W}(\sigma)$,
2. $X(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{X}(\sigma)$,
3. $y_{t-T}^{t}(\sigma) - Y(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{V}(\sigma)$.

Definitions and notation (2): Measurements-compatible configurations

Consider

- Time interval $[t-T,t]$
- Measurement profile y_{t-T}^t
- $(\xi,\mathbf{w}) \in \mathbb{X}(t-T) \times [\mathbb{R}^{n_w}]^{[t-T,t]}$

$(\xi,\mathbf{w})$ is (y_{t-T}^t)-compatible

if for all $\sigma \in [t-T,t]$:

1. $w(\sigma) \in \mathbb{W}(\sigma)$,
2. $X(\sigma,t-T,\xi,\mathbf{w}) \in \mathbb{X}(\sigma)$,
3. $y_{t-T}^t(\sigma) - Y(\sigma,t-T,\xi,\mathbf{w}) \in \mathbb{V}(\sigma)$.

Notation

$$(\xi,\mathbf{w}) \in \mathbb{C}(t,y_{t-T}^t)$$

Definitions and notation (2): Measurements-compatible configurations

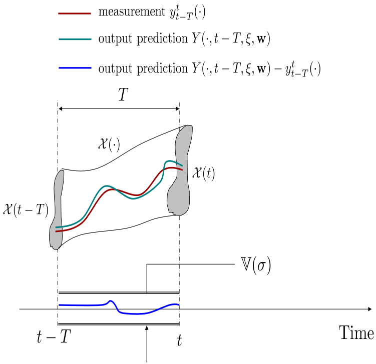
Consider

- Time interval $[t-T,t]$
- Measurement profile y_{t-T}^t
- $(\xi,\mathbf{w}) \in \mathbb{X}(t-T) \times [\mathbb{R}^{n_w}]^{[t-T,t]}$

$(\xi,\mathbf{w})$ is (y_{t-T}^t)-compatible

if for all $\sigma \in [t-T,t]$:

1. $w(\sigma) \in \mathbb{W}(\sigma)$,
2. $X(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{X}(\sigma)$,
3. $y_{t-T}^t(\sigma) - Y(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{V}(\sigma)$.



Notation

$$(\xi,\mathbf{w}) \in \mathbb{C}(t, y_{t-T}^t)$$

Definitions and notation (2): Measurements-compatible configurations

Consider

- Time interval $[t-T, t]$
- Measurement profile y_{t-T}^{t}
- $(\xi, \mathbf{w}) \in \mathbb{X}(t-T) \times [\mathbb{R}^{n_w}]^{[t-T,t]}$

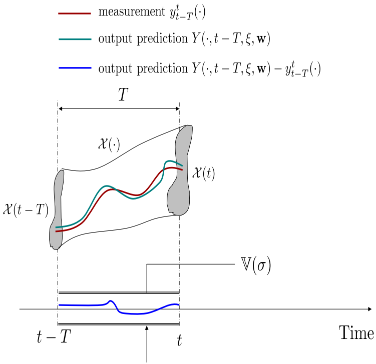
$(\xi, \mathbf{w})$ is (y_{t-T}^{t})-compatible

if for all $\sigma \in [t-T, t]$:

1. $w(\sigma) \in \mathbb{W}(\sigma)$,
2. $X(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{X}(\sigma)$,
3. $y_{t-T}^{t}(\sigma) - Y(\sigma, t-T, \xi, \mathbf{w}) \in \mathbb{V}(\sigma)$.

Notation

$$(\xi, \mathbf{w}) \in \mathbb{C}(t, y_{t-T}^{t})$$



$(\xi, \mathbf{w}) \in \mathbb{C}(t, y_{t-T}^{t})$ if the corresponding trajectory

1. meets the constraints
2. explains the measurements

The finite horizon observation problem

The finite horizon observation problem

Choose a finite $T > 0$ and use at each t, the available information:

1. System equations
2. Past measurements y_{t-T}^{t},
3. Constraints and
4. *Some additional exogenous knowledge*.

in order to produce an estimation $\hat{x}(t)$ of the current state $x(t)$. ♡

The finite horizon observation problem

Choose a finite $T > 0$ and use at each t, the available information:

1. System equations
2. Past measurements y_{t-T}^{t},
3. Constraints and
4. *Some additional exogenous knowledge*.

in order to produce an estimation $\hat{x}(t)$ of the current state $x(t)$. ♡

$$\text{Find } (\xi, \mathbf{w}) \quad \longrightarrow \quad \hat{x}(t) = X(t, t-T, \xi, \mathbf{w})$$

The finite horizon observation problem

The finite horizon observation problem

Choose a finite $T > 0$ and use at each t, the available information:

1. System equations
2. Past measurements y_{t-T}^{t},
3. Constraints and
4. *Some additional exogenous knowledge*.

in order to produce an estimation $\hat{x}(t)$ of the current state $x(t)$. ♡

Find $(\xi, \mathbf{w}) \quad \longrightarrow \quad \hat{x}(t) = X(t, t-T, \xi, \mathbf{w})$

The set of candidate estimates $\hat{x}(t)$:

$$\Omega_t = \Big\{ X(t, t-T, \xi, \mathbf{w}) \quad | \quad (\xi, \mathbf{w}) \in \mathbb{C}(t, y_{t-T}^{t}) \Big\}.$$

The need for additional knowledge

$$\Omega_t = \Big\{X(t, t-T, \xi, \mathbf{w}) \quad | \quad (\xi, \mathbf{w}) \in \mathbb{C}(t, y_{t-T}^t)\Big\}.$$

- Either $\Omega_t = \{x(t)\}$, for instance because
  - $\mathbb{W} = \{0\}$, $\mathbb{V} = \{0\}$ and
  - The system has no indistinguishable states

$$\int_{t-T}^{t} \|Y(\sigma, t-T, x^{(1)}) - Y(\sigma, t-T, x^{(2)})\|^2 d\sigma \geq \alpha(\|x^{(1)} - x^{(2)}\|)$$

for all $t \geq 0$ and all $(x^{(1)}, x^{(2)}) \in \mathbb{X}(t-T) \times \mathbb{X}(t-T)$.

The need for additional knowledge

$$\Omega_t = \Big\{ X(t, t-T, \xi, \mathbf{w}) \quad | \quad (\xi, \mathbf{w}) \in \mathbb{C}(t, y_{t-T}^{t}) \Big\}.$$

- Or $\Omega_t \neq \{x(t)\}$, and a *selection* must be made by solving

$$P(t) \quad : \quad \min_{(\xi,\mathbf{w})\in\Omega_t} \Phi(\xi) + J(t,\xi,\mathbf{w}) \quad \rightarrow (\hat{\xi}(t), \hat{\mathbf{w}}(t))$$

  - Estimation: $\hat{x}(t) = X(t, t-T, \hat{\xi}(t), \hat{\mathbf{w}}(t))$
  - $\Phi(\xi)$ is the cost to go
    (summary of the past measurement related information)
  - In the sequel $\Phi(\xi) + J(t,\xi,\mathbf{w})$ is shortly denoted by $J(t,\xi,\mathbf{w})$

Temporal Parametrization (1)

Solve $P(t) \quad : \quad \min_{(\xi,\mathbf{w})\in\mathbb{C}(t)} J(t,\xi,\mathbf{w})$

In many textbooks, the following parametrization is suggested for **w**:

$$p_w := \{\mathbf{w}(k\tau)\}_{k=k_0}^{k_0+N-1} \in \mathbb{W}(k_0)\times\cdots\times\mathbb{W}(k_0+N-1)\subset\mathbb{R}^{n_w\cdot N}$$

- Decision variable (ξ, p_w) of dimension $n + N\cdot n_w$
- Too rich spectral content *increasing* uselessly Ω_t
- High sensitivity to the knowledge of $\mathbb{W}(\cdot)$.

Unrealistically too many possible interpretations of the measurements

Temporal parametrization (2)

Solve $P(t) \quad : \quad \min_{(\xi,\mathbf{w})\in\mathbb{C}(t)} J(t,\xi,\mathbf{w})$

Use a reduced dimensional parametrization

$$w(t) = \mathcal{W}(t,p_w) \quad ; \quad p_w \in \mathbb{P}.$$

Solve $P(t) \quad : \quad \min_{(\xi,p_w)\in\mathbb{C}(t)} J(t,\xi,\mathcal{W}(\cdot,p_w)) =: J(t,\xi,p_w) \quad \rightarrow (\hat{\xi}(t),\hat{p}_w(t))$

$$\hat{x}(t) = X(t,t-T,\hat{\xi}(t),\mathcal{W}(\cdot,\hat{p}_w(t)))$$

- $\bar{x} := (x^T, p_w^T)^T \in \mathbb{R}^n \times \mathbb{R}^{n_p}$
- $\dot{p}_w = 0$

New uncertainty-free extended state estimation problem.

Analytic vs optimization based observer

Analytic observers

$$\begin{array}{ll}\text{(System)} & \dot x = f(x)\ ;\ y = h(x)\\ \text{(Observ)} & \dot{\hat x} = f(\hat x) + K(\hat x, y)\end{array}$$

Try to show asymptotic convergence of $e = x - \hat x$ governed by

$$\begin{aligned}\dot x &= f(x)\\ \dot e &= f(x) - f(x-e) - K(x-e, h(x))\end{aligned}$$

Very Hard Task

- Need for structural properties
- Coordinate transformation
- Constructive assumptions
- Observability $\neq$ Existence of observer

Analytic vs optimization based observer

Analytic observers

$$\begin{array}{ll}\text{(System)} & \dot{x} = f(x)\ ;\ y = h(x) \\ \text{(Observ)} & \dot{\hat{x}} = f(\hat{x}) + K(\hat{x}, y)\end{array}$$

Try to show asymptotic convergence of $e = x - \hat{x}$ governed by

$$\begin{aligned}\dot{x} &= f(x) \\ \dot{e} &= f(x) - f(x-e) - K(x-e, h(x))\end{aligned}$$

Very Hard Task

- Need for structural properties
- Coordinate transformation
- Constructive assumptions
- Observability $\neq$ Existence of observer

Optimization based observers

Rely on the implication

$$\Big\{J(t,\xi) \to 0\Big\} \Rightarrow \Big\{\underbrace{X(t, t-T, \xi)}_{\hat{x}(t)} \to x(t)\Big\}$$

+ No need to study the dynamic of e
+ No need for structural assumptions
+ Observability $\Leftrightarrow$ Observer
+ Handling constraints on the state

Potential problems

- Global convergence ?
- Computation time ?

Example: State estimation of terpolymerization reactors

- Produce polymer from multi-monomer
- Controlling the final properties need the state to be estimated
- State: Polymer composition $\leftrightarrow$ Monomers concetrations
- Complex equations
- Unknown dynamics
- High gain observers need tremendous simplifications to give rather poor performance

Coll. Nida Sheibat-Othman & Sami Othman (LAGEP,Lyon)

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad i = 1, 2, 3$$

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad\qquad i = 1,2,3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad\qquad i = 1, 2, 3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

where

$$P_1^P = \frac{\alpha}{\alpha+\beta+\gamma} \quad ; \quad P_2^P = \frac{\beta}{\alpha+\beta+\gamma} \quad ; \quad P_3^P = 1 - P_1^P - P_2^P$$

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad i = 1, 2, 3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

where

$$P_1^P = \frac{\alpha}{\alpha+\beta+\gamma} \quad ; \quad P_2^P = \frac{\beta}{\alpha+\beta+\gamma} \quad ; \quad P_3^P = 1 - P_1^P - P_2^P$$

in which

$$\begin{aligned}
\alpha &= [M_1^P](k_{p21}k_{p31}[M_1^P] + k_{p21}k_{p32}[M_2^P] + k_{p31}k_{p23}[M_3^P]) \\
\beta &= [M_2^P](k_{p12}k_{p31}[M_1^P] + k_{p12}k_{p32}[M_2^P] + k_{p13}k_{p32}[M_3^P]) \\
\gamma &= [M_3^P](k_{p13}k_{p21}[M_1^P] + k_{p21}k_{p23}[M_2^P] + k_{p13}k_{p23}[M_3^P])
\end{aligned}$$

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad i = 1, 2, 3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

where

$$P_1^P = \frac{\alpha}{\alpha+\beta+\gamma} \quad ; \quad P_2^P = \frac{\beta}{\alpha+\beta+\gamma} \quad ; \quad P_3^P = 1 - P_1^P - P_2^P$$

The $[M_i^P]$ depend in the state according to:

$$[M_i^P] = \begin{cases} \dfrac{(1-\phi_p^p)N_i}{\sum_j \dfrac{N_j MW_j}{\rho_j}}, & \text{(Phase II)} \\[2ex] \dfrac{N_i}{\sum_j MW_j(\dfrac{N_j^T - N_j}{\rho_{j,h}} + \dfrac{N_j}{\rho_j})} & \text{(Phase III)} \end{cases}$$

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad i = 1, 2, 3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

- μ plays a crucial role
- The dynamic of μ is unknown

Example: terpolymerization reactors (The mathematical model)

$$\dot{N}_i = Q_i - R_{Pi} \qquad i = 1, 2, 3$$

$$R_{P_i} = \mu[M_i^P](k_{p1i}P_1^P + k_{p2i}P_2^P + k_{p3i}P_3^P)$$

- μ plays a crucial role
- The dynamic of μ is unknown

- Measurement

The overall monomer conversion measured by calorimetry:

$$y = \frac{\sum_{i=1}^{3} MW_i(N_i^T - N_i)}{\sum_{j=1}^{3} MW_j N_j^T}$$

Example: State reconstruction of terpolymerization reactors (Validation)

1. Simulation results
2. Experimental results

Example: State reconstruction of terpolymerization reactors (Validation)

Simulation results

$$
\begin{aligned}
\dot{N} &= \begin{pmatrix} 1+d_1 & 0 & 0 \\ 0 & 1+d_2 & 0 \\ 0 & 0 & 1+d_3 \end{pmatrix} \cdot f(x,u) \\
\dot{\mu} &= 0 \\
y &= (1+\nu)\cdot h(x)
\end{aligned}
$$

- The state $x := \begin{pmatrix} N_1 & N_2 & N_3 & \mu \end{pmatrix} \in \mathbb{R}_+^4$
- The uncertainties

$$
\begin{aligned}
d_i(k) &= d_{max} \cdot r_i(k) \\
\nu(k) &= \nu_{max} \cdot r_\nu(k)
\end{aligned}
$$

- r_i and ν randomly chosen in $[-1,+1]$

Simulation results

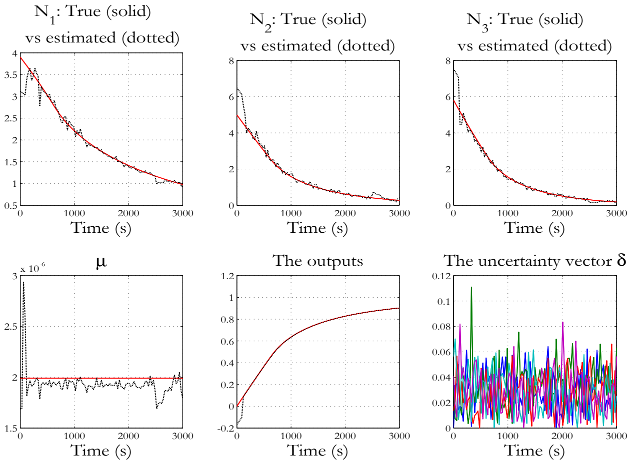


Figure: *Observer behavior under model uncertainty given by (1)-(1) with* $d_{max} = 10\%$ *and no measurement noise* ($\nu_{max} = 0$)*. The observation horizon is* $N = 10$ *and the number of trials for the singularity crossing scheme is* $N_{\text{trials}} = 4$*. Initial state of the observer is* $\hat{x}(0) = diag(0.8, 1.3, 1.3) \cdot x(0)$ *and* $\mu_{obs}(0) = 0.8\mu_{model}$*.*

Simulation results

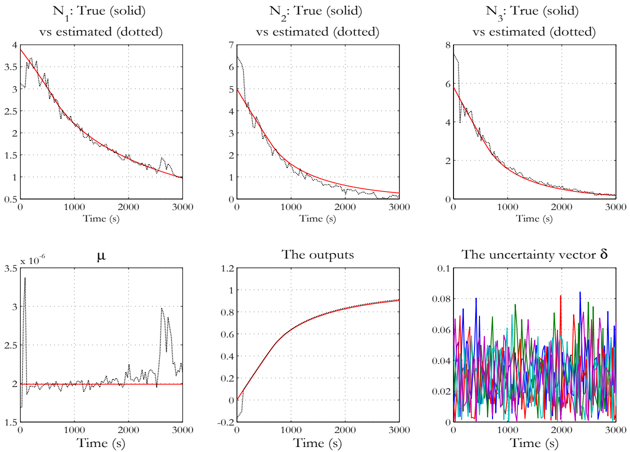


Figure: *Observer behavior under model uncertainty given by (1)-(1) with $d_{max} = 10\%$ and in the presence of measurement noise ($\nu_{max} = 0.01$). The observation horizon is $N = 15$ and the number of trials for the singularity crossing scheme is $N_{\text{trials}} = 4$. $\mu_{obs}(0) = 0.8\mu_{model}$. Note that concerning the output, only the true output and the estimated one are shown, measurement noise is not presented. This scenario uses a tolerance $\epsilon = 10^{-8}$ for the optimization subroutine.*

Experimental results: $N_{trials} = 10$

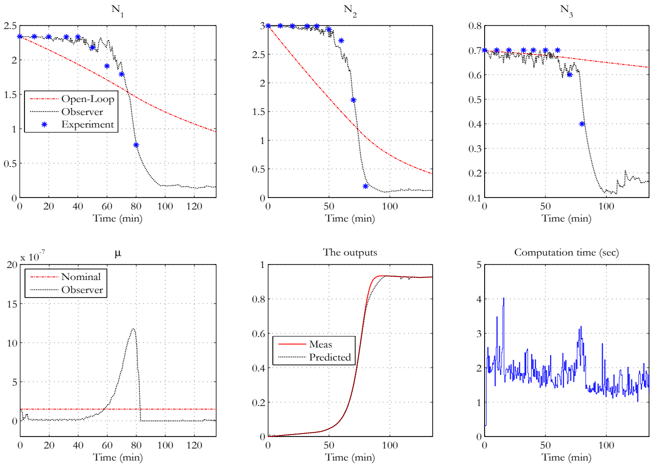


Figure: *Experimental validation with* $N_{trials} = 10$ *and tolerance threshold* $\varepsilon = 10^{-3}$. *The same scenario is depicted on figure 6 where* $N_{trials} = 1$ *is used. The computation time is given in seconds.*

Forthcoming issues

Global convergence ?

- No generic and definitive solution . . . !
- Heuristics for singularities avoidance

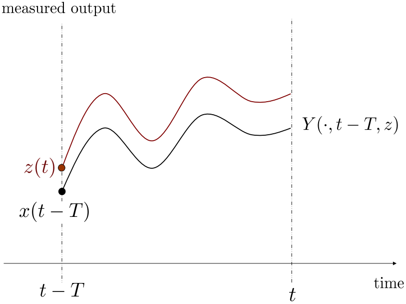
Computation time ?

- Differential form of optimization based observer
- Real-Time iterations / Optimal choice of updating period

Moving Horizon Observer

$$x(t) = X(t, t_0, x_0)$$
$$y = h(x)$$

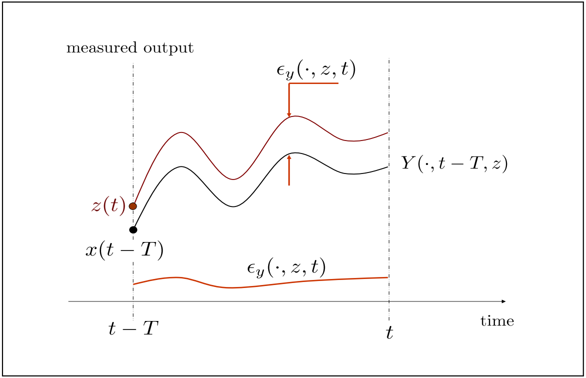
- *Generally a non convex optimization problem*
- *many local minima*



$$z_{opt}(t) \leftarrow \arg\min_{z} \quad J_0(z, t, y_{t-T}^{t}) = \int_{t-T}^{t} \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

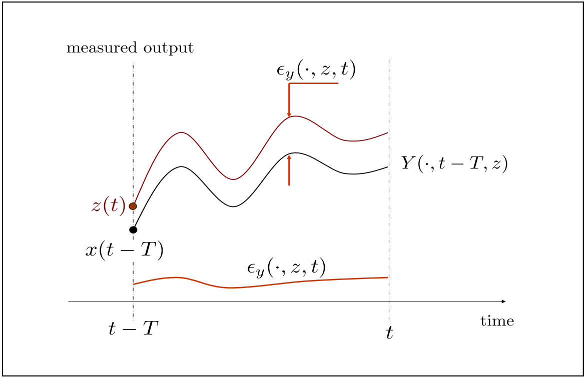
$$\hat{x}(t) := X(t, t-T, z_{opt}(t))$$

A very particular problem . . . !



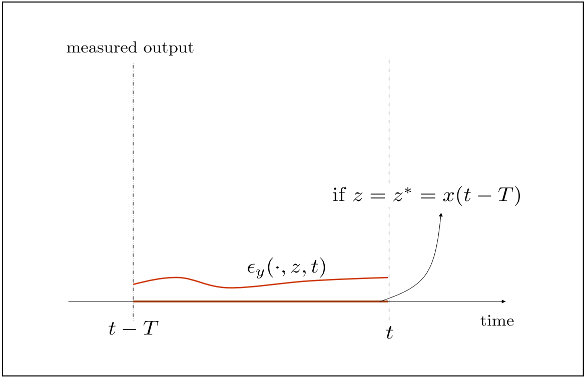
$$z_{opt}(t) \leftarrow arg \min_{z} \quad J_0(z, t, y_{t-T}^{t}) = \int_{t-T}^{t} \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

A very particular problem . . . !



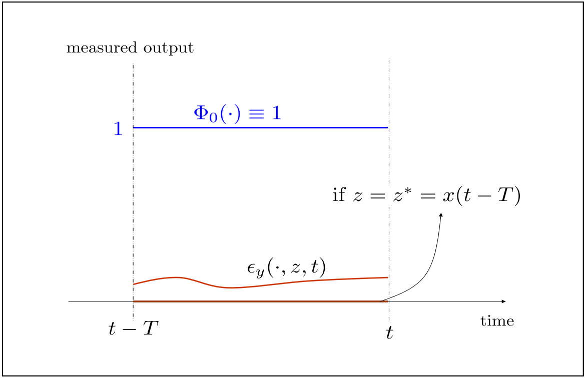
$$z_{opt}(t) \leftarrow arg \min_{z} \quad J_0(z, t, y_{t-T}^{t}) = \int_{t-T}^{t} \|\epsilon_y(\tau, z, t)\|_Q^2 d\tau$$

A very particular problem . . . !



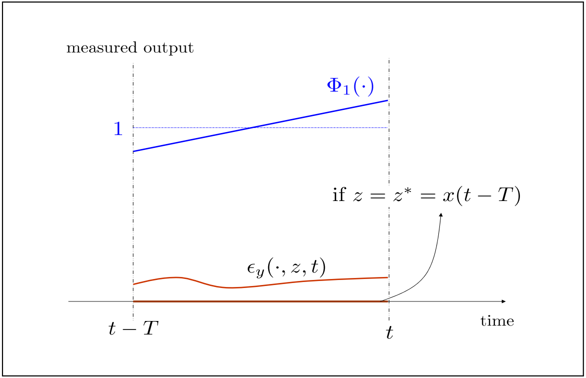
$$z_{opt}(t) \leftarrow \arg\min_z \quad J_0(z, t, y_{t-T}^t) = \int_{t-T}^{t} \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem . . . !



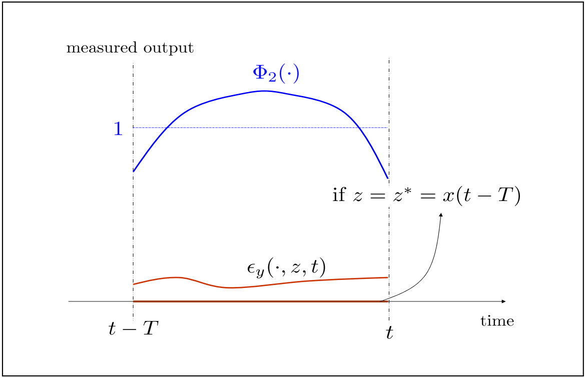
$$z_{opt}(t) \leftarrow \arg\min_z \quad J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \Phi_0(\tau) \cdot \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem . . . !



$$z_{opt}(t) \leftarrow \arg\min_{z} \quad J_1(z, t, y_{t-T}^{t}) = \int_{t-T}^{t} \Phi_1(\tau) \cdot \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem . . . !



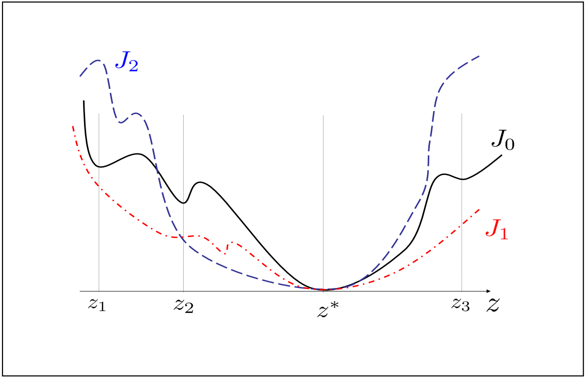
$$z_{opt}(t) \leftarrow arg \min_z \quad J_2(z, t, y_{t-T}^t) = \int_{t-T}^{t} \Phi_2(\tau) \cdot \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem . . . !

z^* is THE *global minimum of* ALL *the cost functions J_i s.t:*

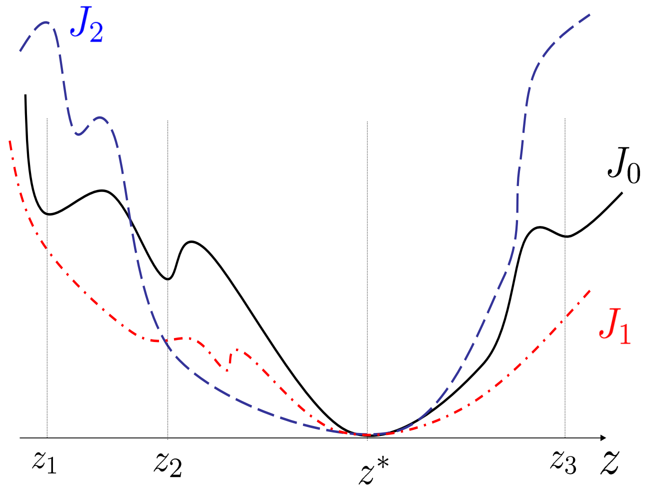
$$J_i(z, t, y_{t-T}^t) = \int_{t-T}^{t} \Phi_i(\tau) \cdot \Psi\big(\epsilon(\tau, z, t)\big)\, d\tau$$

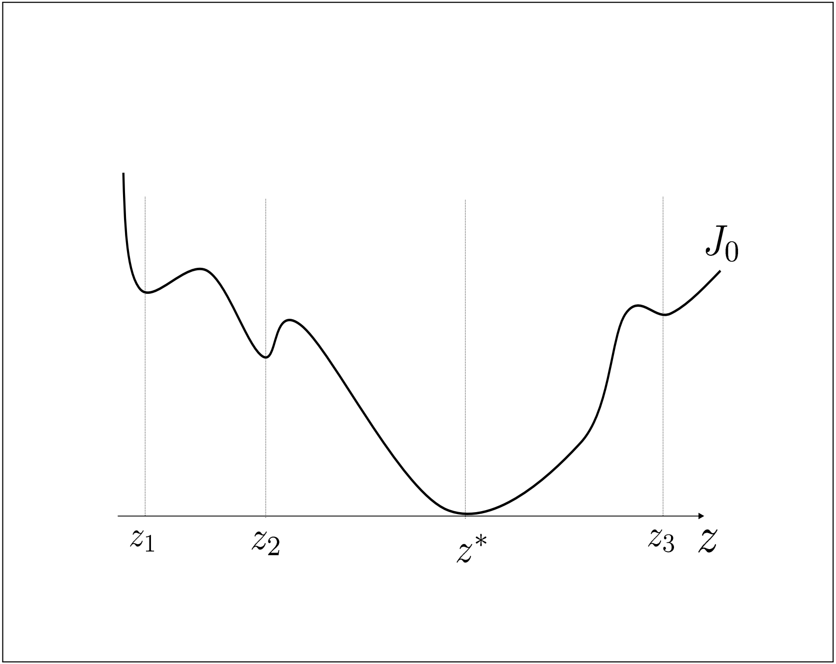
A very particular problem . . . !

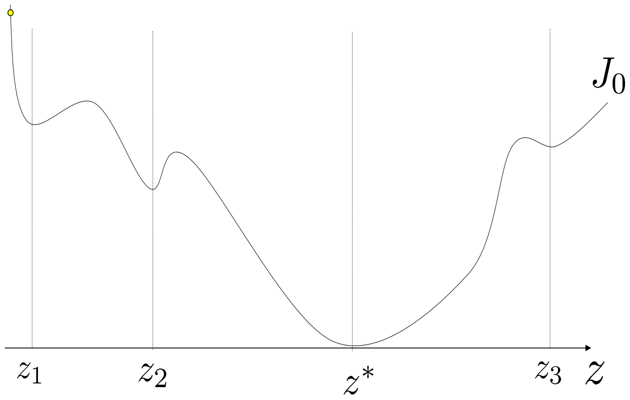


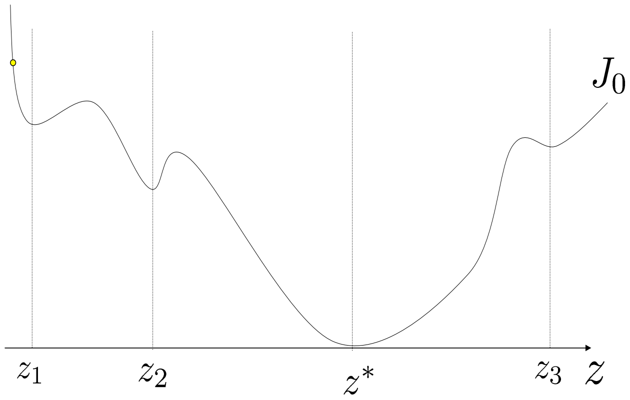
z^* is THE global minimum of ALL the cost functions J_i s.t:

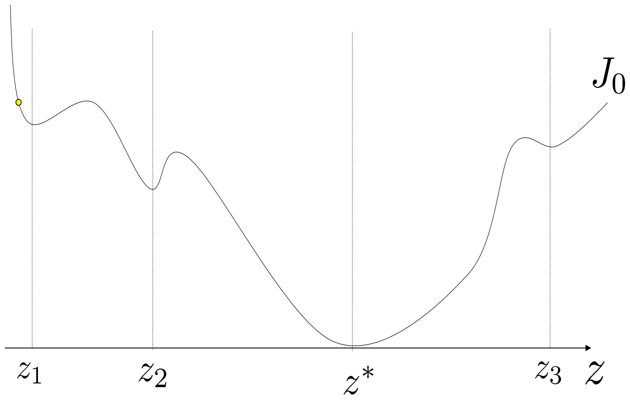
$$J_i(z, t, y_{t-T}^t) = \int_{t-T}^{t} \Phi_i(\tau) \cdot \Psi\big(\epsilon(\tau, z, t)\big)\, d\tau$$

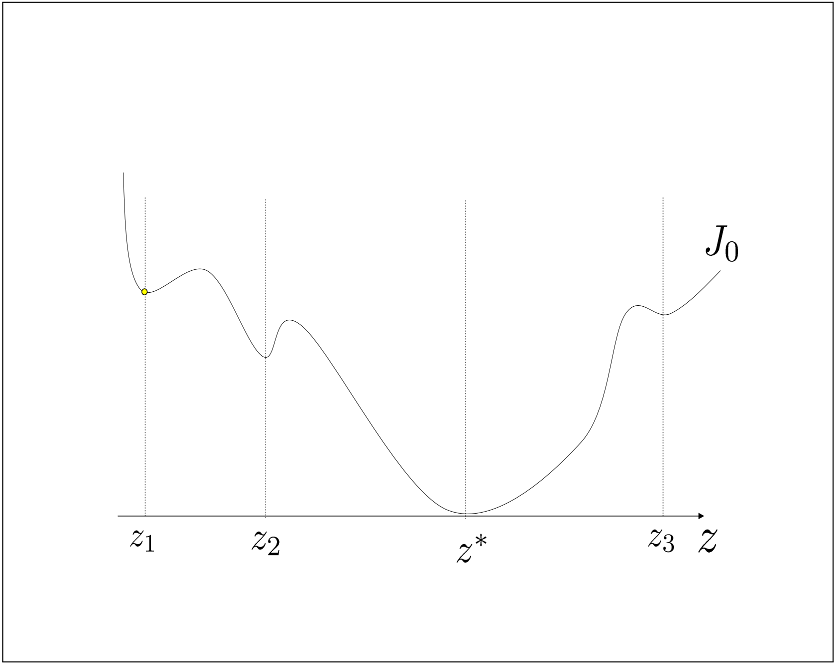

J_2
J_0
J_1
z_1
z_2
z^*
z_3
Z

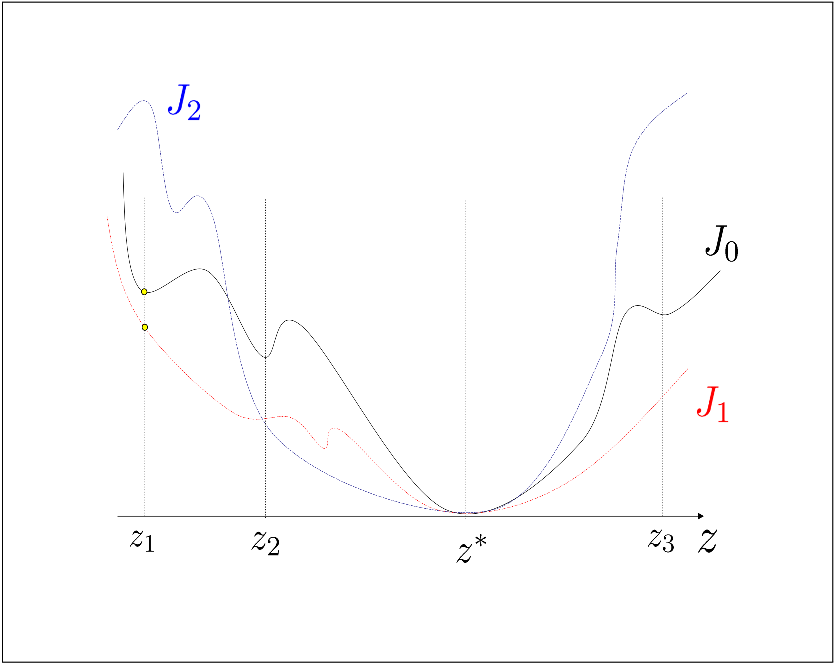

J_0
z_1
z_2
z^*
z_3
z

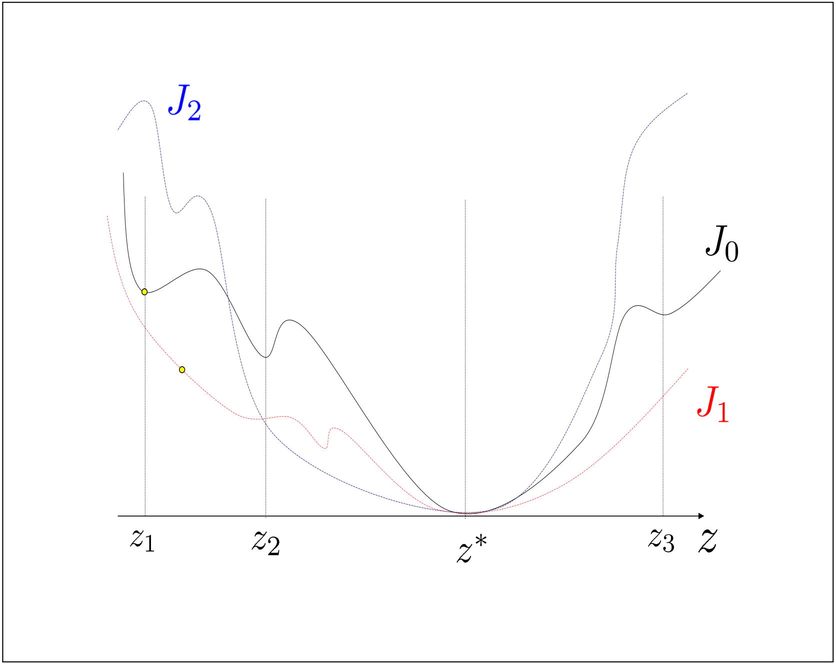

J_0
z_1
z_2
z^*
z_3
Z

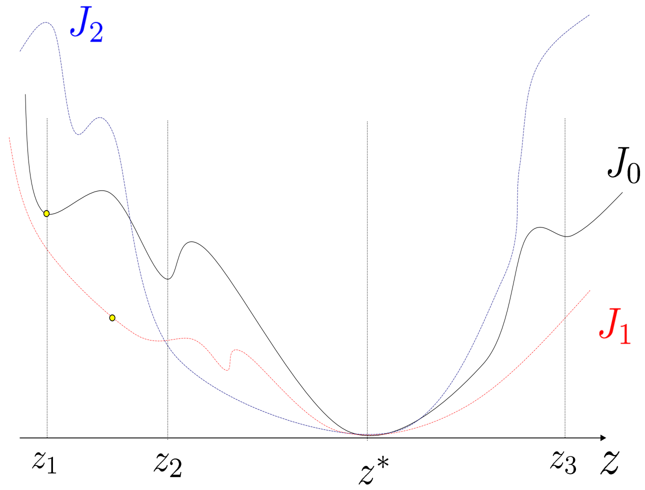

J_0
z_1
z_2
z^*
z_3
Z

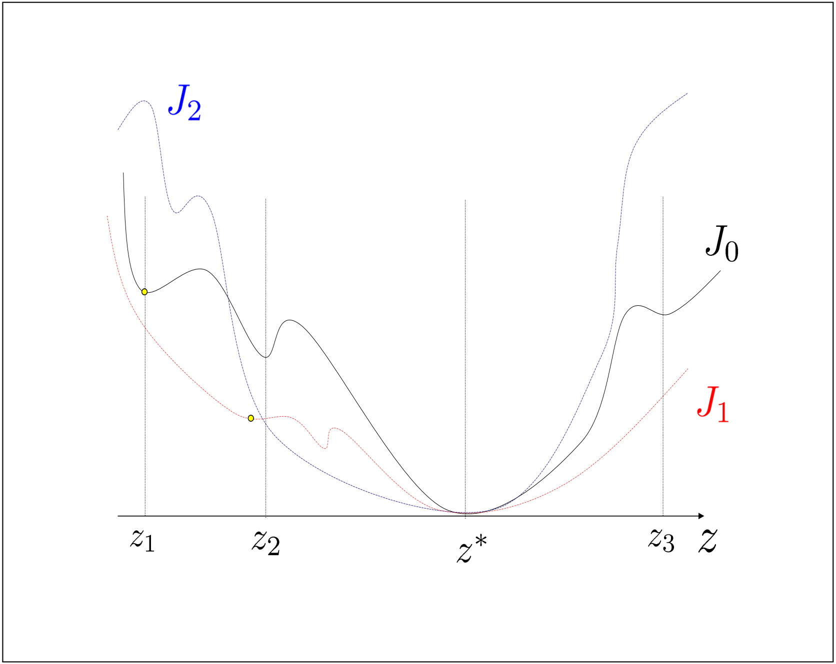

J_0
z_1
z_2
z^*
z_3
z

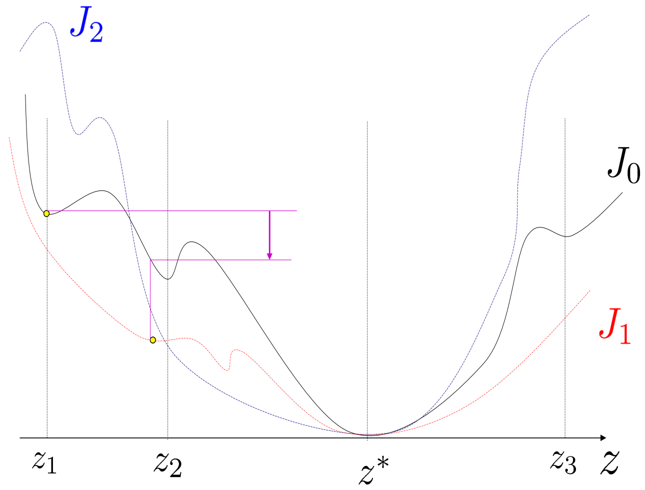

J_0
z_1
z_2
z^*
z_3
Z

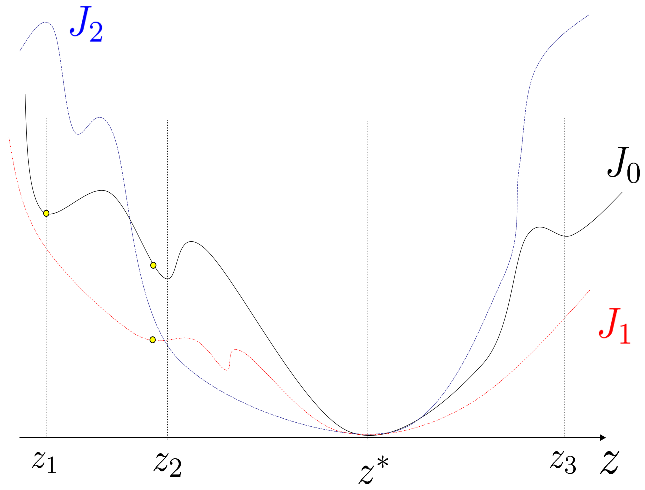

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

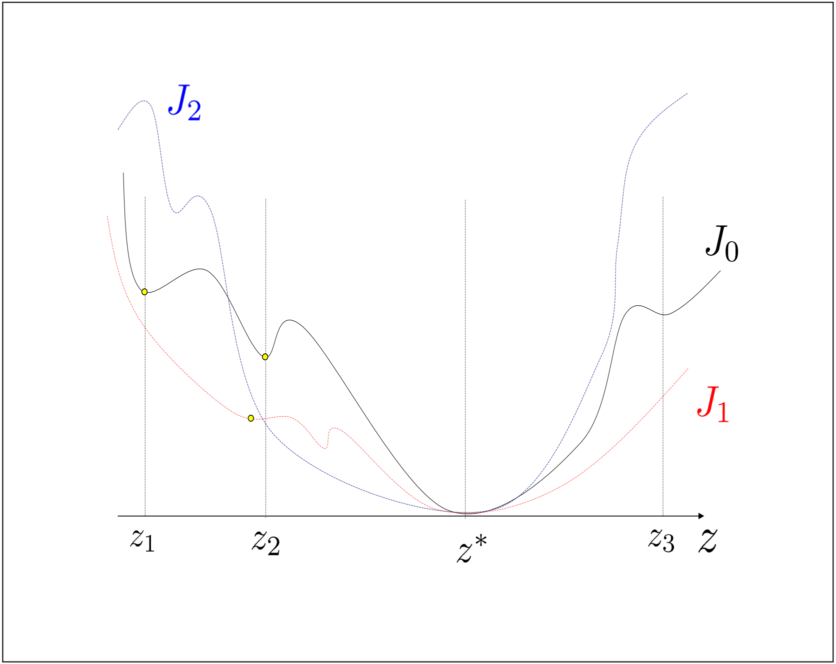

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

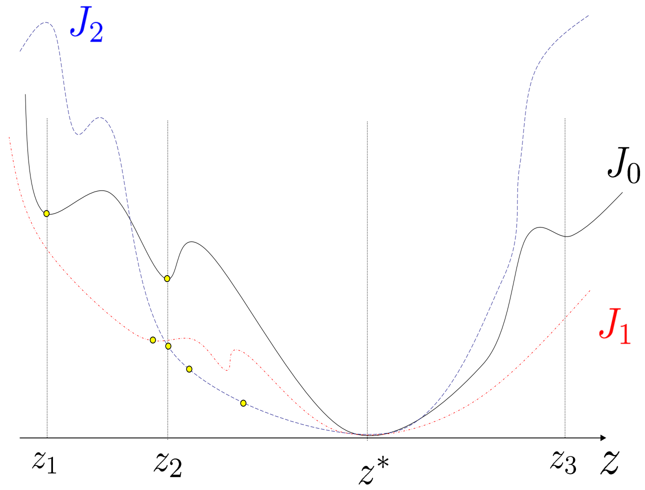

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

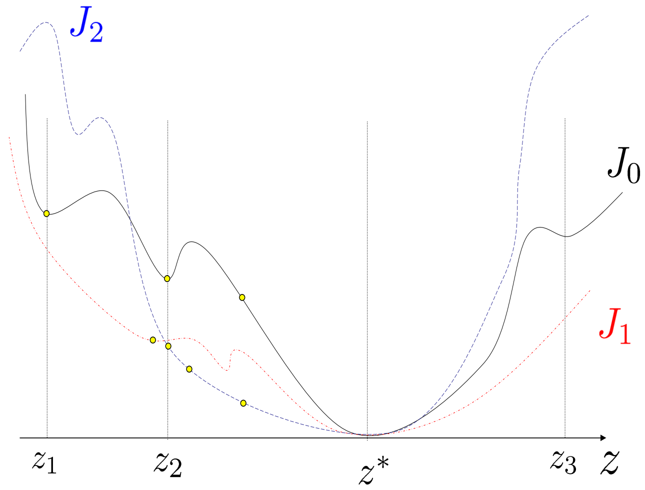

J_2
J_0
J_1
z_1
z_2
z^*
z_3
Z

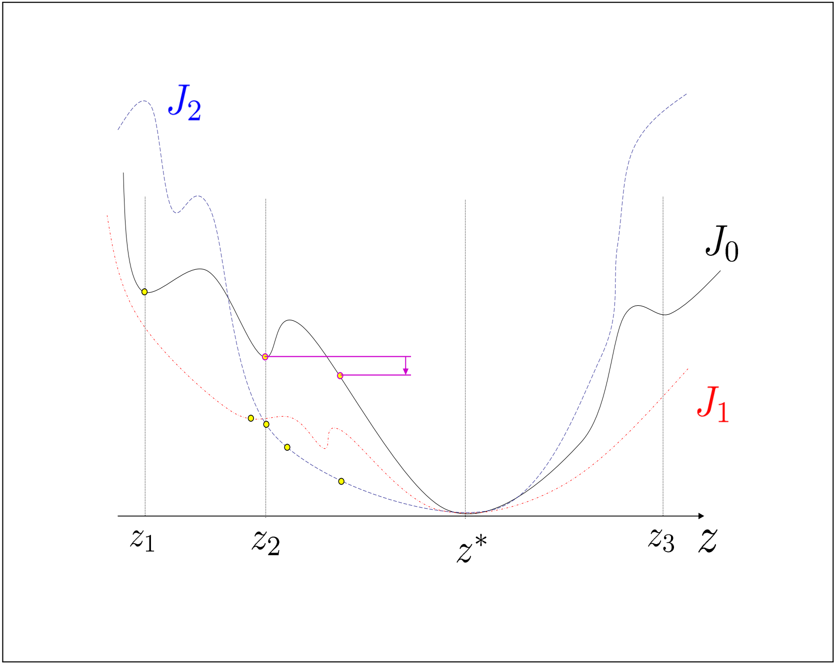

J_2
J_0
J_1
z_1
z_2
z^*
z_3
Z

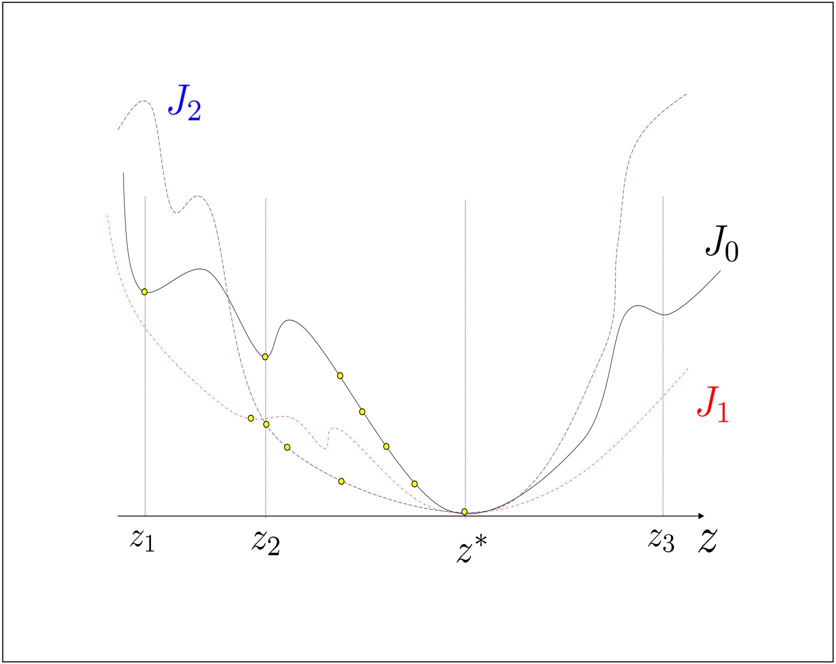

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

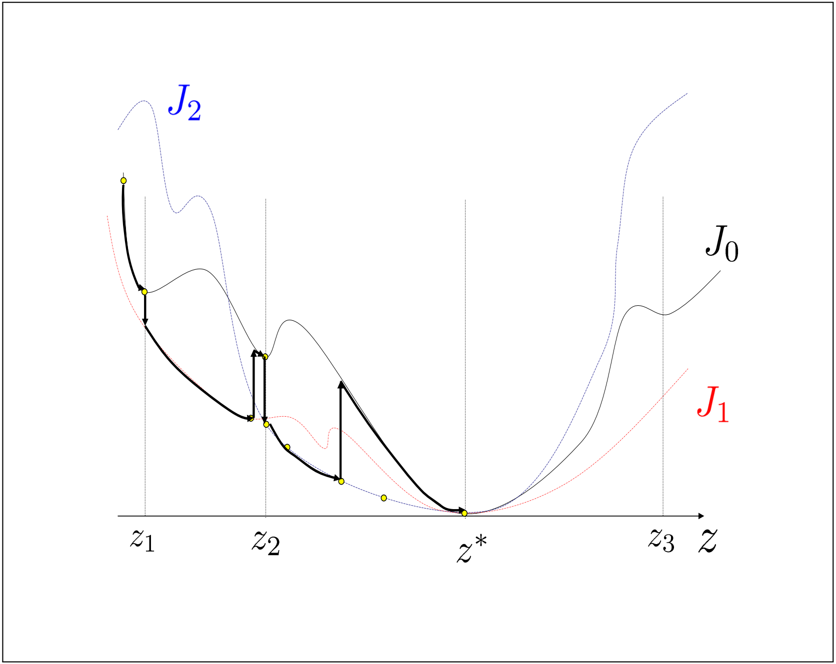

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

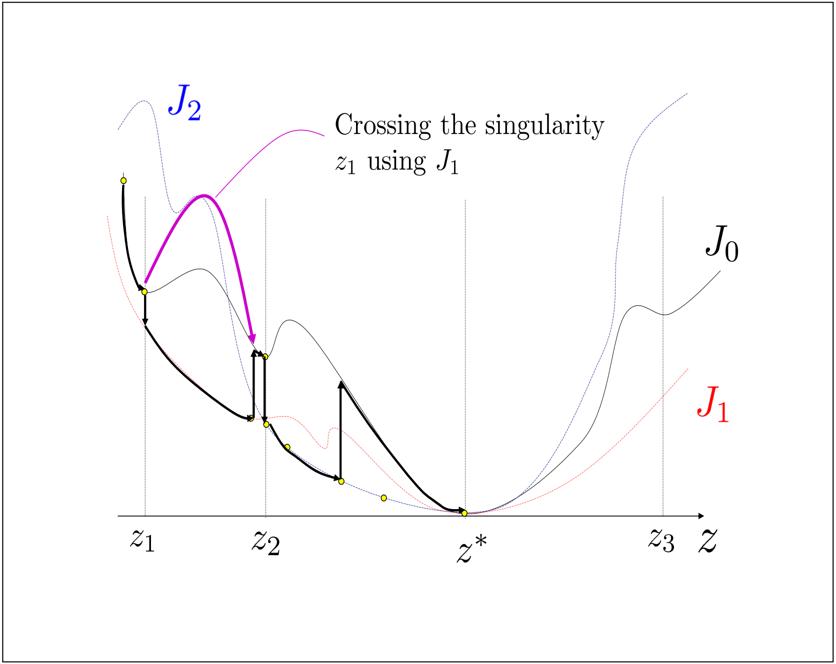

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

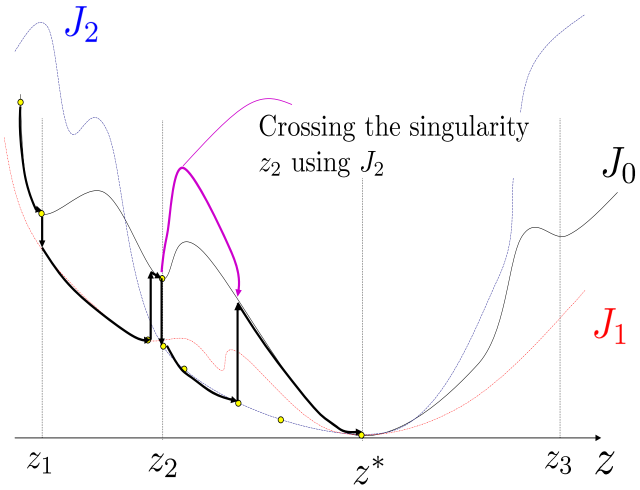

J_2
J_0
J_1
z_1
z_2
z^*
z_3
z


J_2
J_0
J_1
z_1
z_2
z^*
z_3
z


J_2
J_0
J_1
z_1
z_2
z^*
z_3
z


J_2
J_0
J_1
z_1
z_2
z^*
z_3
z


J_2
Crossing the singularity
z_1 using J_1
J_0
J_1
z_1
z_2
z^*
z_3
z

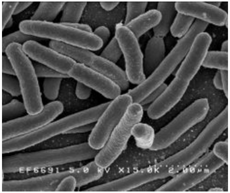

J_2
Crossing the singularity
z_2 using J_2
J_0
J_1
z_1
z_2
z^*
z_3
z

Example 1: Recombinant Escherichia Coli

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

- X: *E. Coli* strain
- S: substrate glycerol
- P: intracellular product β-galactosidase protein
- μ is the growth rate

$$\mu(S) = \frac{\mu_m S}{k_s + S}$$

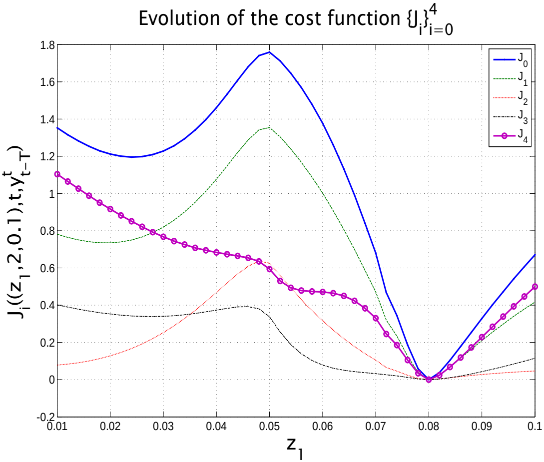


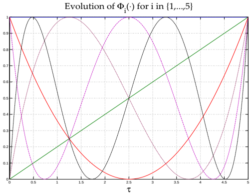
Escherichia coli under 15000 magnification factor

Output measurement:

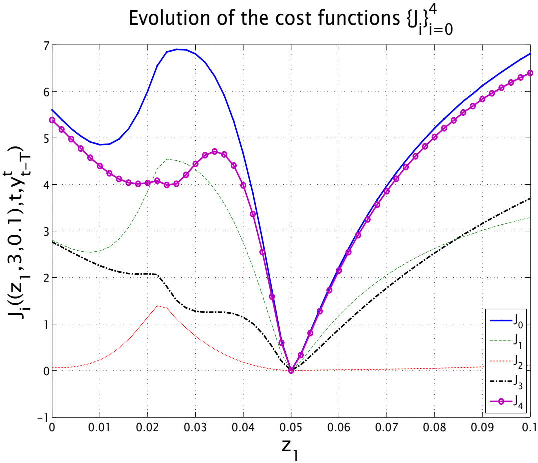
Light produced by the bioluminescence:

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

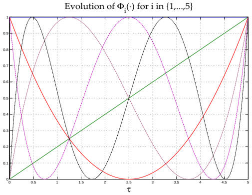


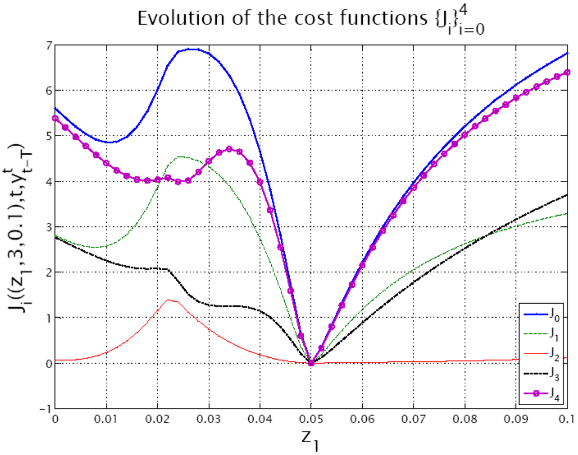


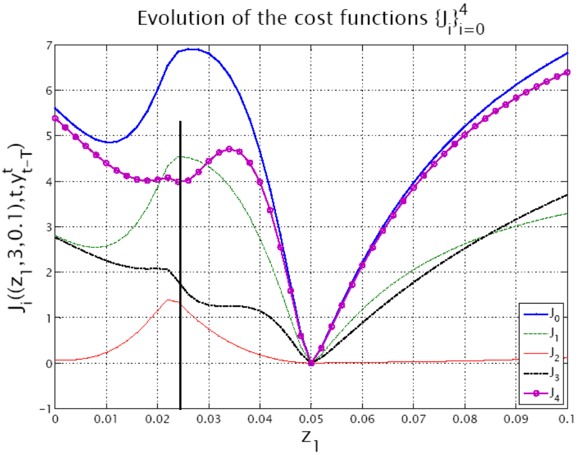
- $T = 10$
- $x(t-T) = (0.08, 2, 0.1)$
- $z_2 = x_2(t-T)$, $z_3 = x_3(t-T)$

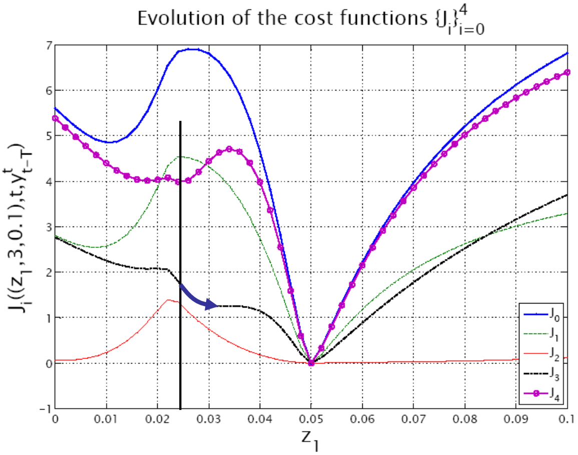


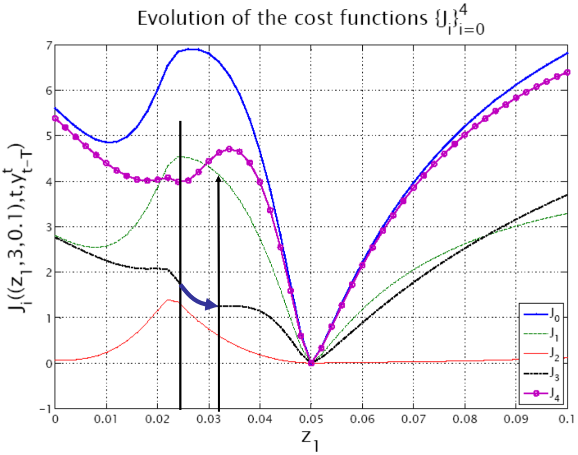
- $T = 15$
- $x(t-T) = (0.05, 3, 0.1)$
- $z_2 = x_2(t-T)$, $z_3 = x_3(t-T)$

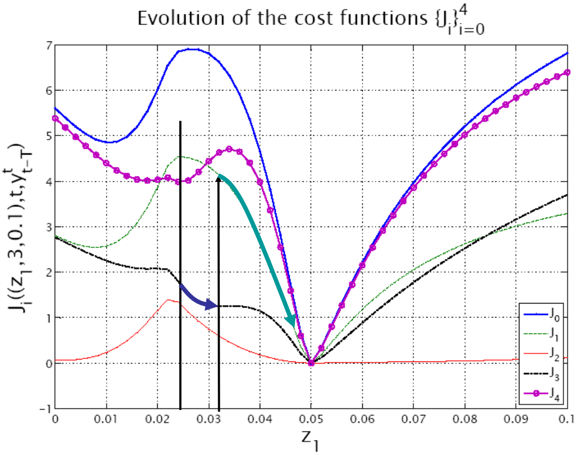



Evolution of the cost functions $\{J_i\}_{i=0}^{4}$
$J_i((z_1,3,0.1),t,y_{t-T}^{t})$
z_1
J_0
J_1
J_2
J_3
J_4
7
6
5
4
3
2
1
0
−1
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1


Evolution of the cost functions $\{J_i\}_{i=0}^4$
$J_i((z_1,3,0.1),t,y_{t-T}^t)$
z_1
J_0
J_1
J_2
J_3
J_4
7
6
5
4
3
2
1
0
−1
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1


Evolution of the cost functions $\{J_i\}_{i=0}^{4}$
$J_i((z_1,3,0.1),t,y_{t-T}^t)$
z_1
J_0
J_1
J_2
J_3
J_4
7
6
5
4
3
2
1
0
−1
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1


Evolution of the cost functions $\{J_i\}_{i=0}^4$
$J_i((z_1,3,0.1),t,y_{t-T}^t)$
z_1
J_0
J_1
J_2
J_3
J_4

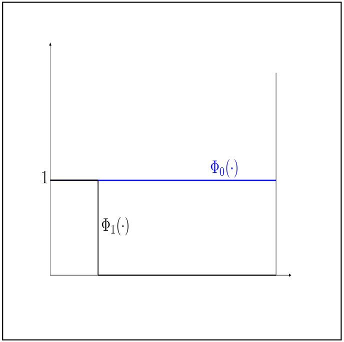

Evolution of the cost functions $\{J_i\}_{i=0}^{4}$
$J_i((z_1,3,0.1),t,y_{t-T}^{t})$
z_1
J_0
J_1
J_2
J_3
J_4
7
6
5
4
3
2
1
0
−1
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1

Example 2: Parameter estimation

$$
\begin{aligned}
\dot{x}_1 &= -p_1 x_2 \\
\dot{x}_2 &= (1+p_2)x_1 + (1-x_1^2)x_2 \\
y &= x_1 + x_2 + \nu
\end{aligned}
$$

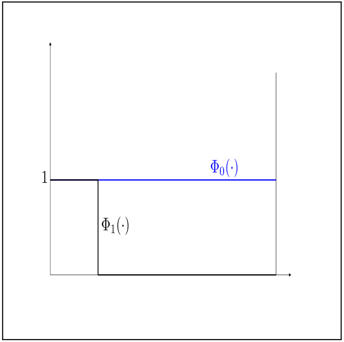
Example 2: Parameter estimation

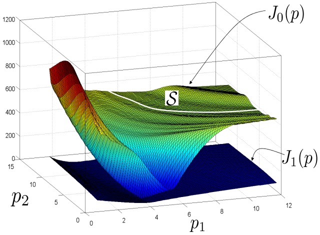
$$
\begin{aligned}
\dot{x}_1 &= -p_1 x_2 \\
\dot{x}_2 &= (1+p_2)x_1 + (1-x_1^2)x_2 \\
y &= x_1 + x_2 + \nu
\end{aligned}
$$



Example 2: Parameter estimation

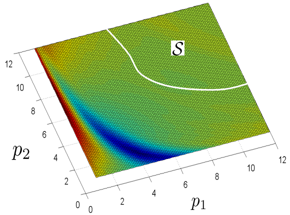
$$\dot{x}_1 = -p_1 x_2$$
$$\dot{x}_2 = (1+p_2)x_1 + (1-x_1^2)x_2$$
$$y = x_1 + x_2 + \nu$$

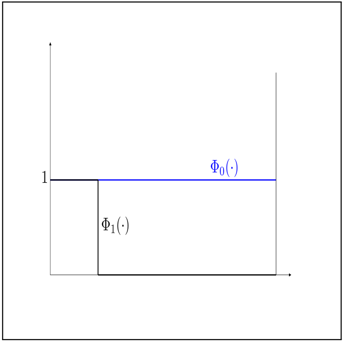




Example 2: Parameter estimation

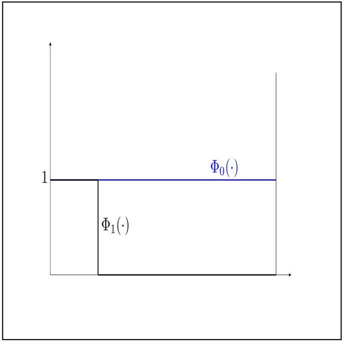
$$\dot{x}_1 = -p_1 x_2$$
$$\dot{x}_2 = (1+p_2)x_1 + (1-x_1^2)x_2$$
$$y = x_1 + x_2 + \nu$$

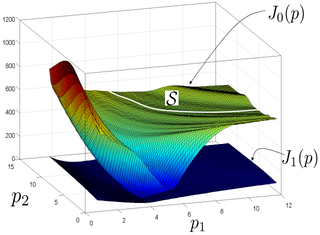




Example 2: Parameter estimation

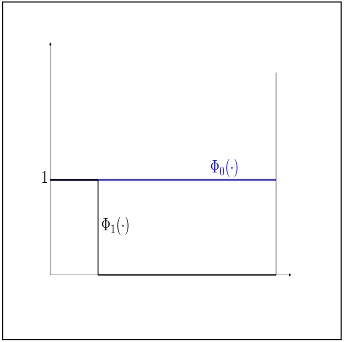
$$\dot{x}_1 = -p_1 x_2$$
$$\dot{x}_2 = (1+p_2)x_1 + (1-x_1^2)x_2$$
$$y = x_1 + x_2 + \nu$$

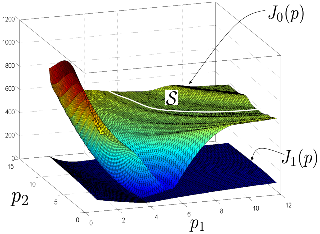


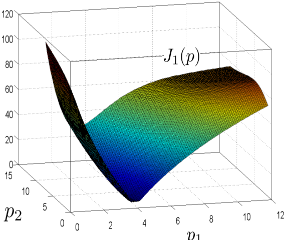


Example 2: Parameter estimation

$$\dot{x}_1 = -p_1 x_2$$
$$\dot{x}_2 = (1+p_2)x_1 + (1-x_1^2)x_2$$
$$y = x_1 + x_2 + \nu$$

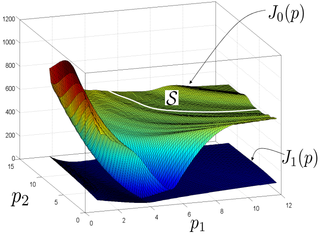


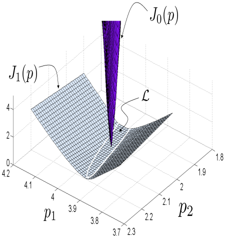


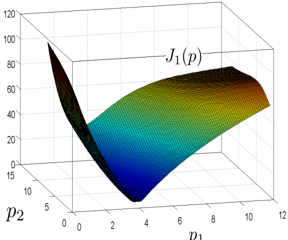


Example 2: Parameter estimation

$$\begin{aligned} \dot{x}_1 &= -p_1 x_2 \\ \dot{x}_2 &= (1+p_2)x_1 + (1-x_1^2)x_2 \\ y &= x_1 + x_2 + \nu \end{aligned}$$

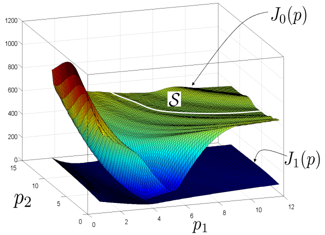


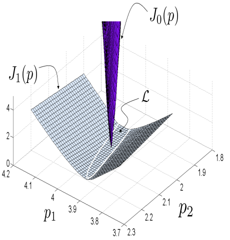




Example 2: Parameter estimation

$$\dot{x}_1 = -p_1 x_2$$
$$\dot{x}_2 = (1+p_2)x_1 + (1-x_1^2)x_2$$
$$y = x_1 + x_2 + \nu$$





i	Algorithm 1	Algorithm 2
1	$Pr\left[J_0(p^{(m,i)}) > 400\right] \geq 0.25$	Conv. to $\mathcal{L}$
2	$Pr\left[J_0(p^{(m,i)}) > 400\right] \geq (0.25)^2$	Conv. to $\{p^r\}$
3	$Pr\left[J_0(p^{(m,i)}) > 400\right] \geq (0.25)^3$	Conv. to $\{p^r\}$
$\vdots$	$\vdots$	$\vdots$

Exemple 3: Back to ter-polymerization (Simulation results)

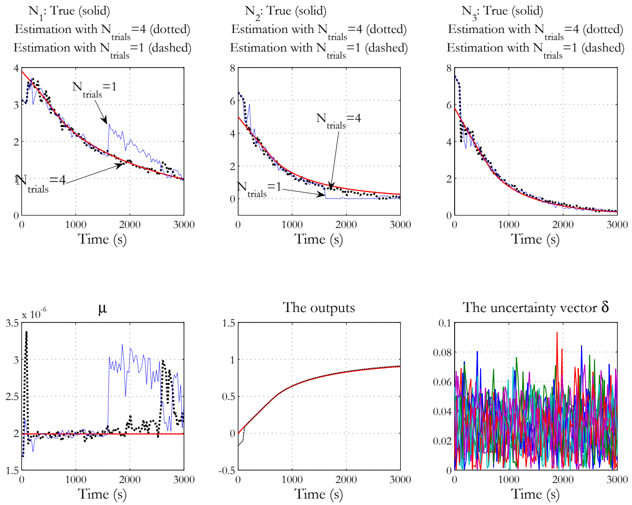


Figure: *Comparison between the observer behavior when $n_\phi = 1$ and $n_\phi = 4$ under the scenario depicted on figure 2. Note how the singularity cross mechanism enables to avoid drops in the estimation quality when the observer encounters a singular situation. This scenario uses a tolerance $\varepsilon = 10^{-8}$ for the optimization subroutine*

Example 3: Back to ter-polymerization (Experimental results: $n_\phi = 10$)

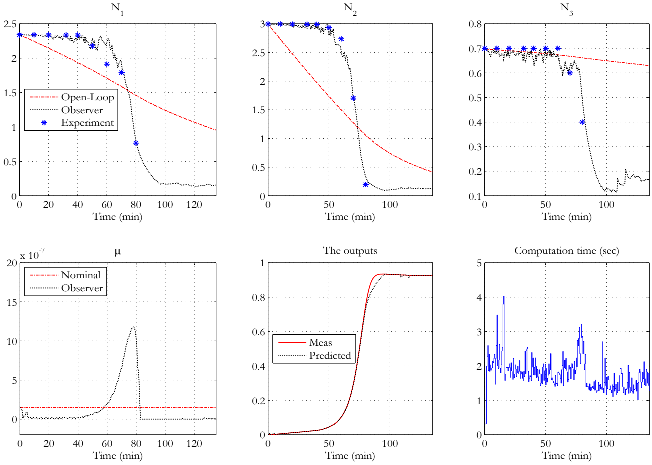


Figure: *Experimental validation with $n_\phi = 10$ and tolerance threshold $\varepsilon = 10^{-3}$. The same scenario is depicted on figure 6 where $n_\phi = 1$ is used. The computation time is given in seconds.*

Example 3: Back to ter-polymerization (Experimental results: $n_\phi = 1$)

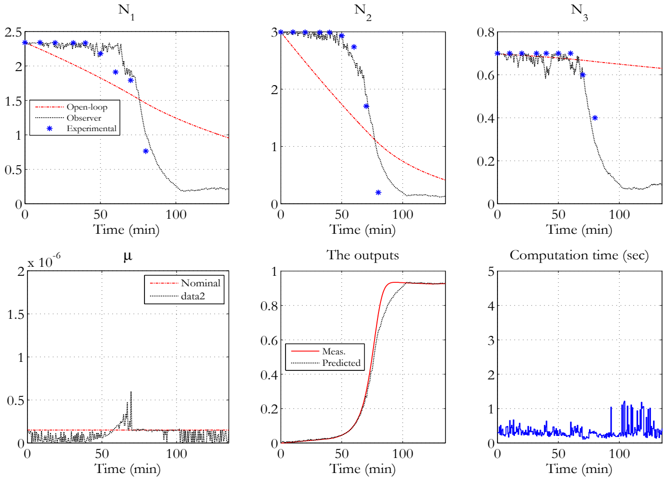


Figure: *Experimental validation with $n_\phi = 1$ and tolerance threshold $\varepsilon = 10^{-3}$. The same scenario is depicted on figure 5 where $n_\phi = 10$ is used. The computation time is given in seconds.*

Discussion

- The scheme holds regardless the optimizer $\mathcal{S}$
  - Gradient-based iteration
  - SQP
  - multiple shooting
  - non smooth (simplex, powell's, etc.)

Discussion

- The scheme holds regardless the optimizer $\mathcal{S}$
  - Gradient-based iteration
  - SQP
  - multiple shooting
  - non smooth (simplex, powell's, etc.)

- Easily usable in a parallel architecture
  [parallel $(z^{(0)}, J_i)$ path-solvers exploration]

Discussion

- The scheme holds regardless the optimizer $\mathcal{S}$
  - Gradient-based iteration
  - SQP
  - multiple shooting
  - non smooth (simplex, powell's, etc.)

- Easily usable in a parallel architecture
  [parallel $(z^{(0)}, J_i)$ path-solvers exploration]

- $\neq$ a multiple initial guess scheme
  [It's the cost function that changes not the present solution]

Discussion

- The scheme holds regardless the optimizer $\mathcal{S}$
  - Gradient-based iteration
  - SQP
  - multiple shooting
  - non smooth (simplex, powell's, etc.)

- Easily usable in a parallel architecture
  [parallel $(z^{(0)}, J_i)$ path-solvers exploration]

- $\neq$ a multiple initial guess scheme
  [It's the cost function that changes not the present solution]

- Price: Loose optimality that is *loosely defined*

Analytic vs optimization based observer

Analytic observers

$$\text{(System)} \quad \dot{x} = f(x) \; ; \; y = h(x)$$
$$\text{(Observ)} \quad \dot{\hat{x}} = f(\hat{x}) + K(\hat{x}, y)$$

Try to show asymptotic convergence of $e = x - \hat{x}$ governed by

$$\dot{x} = f(x)$$
$$\dot{e} = f(x) - f(x-e) - K(x-e, h(x))$$

Very Hard Task

- Need for structural properties
- Coordinate transformation
- Constructive assumptions
- Observability $\neq$ Existence of observer

Optimization based observers

Rely on the implication

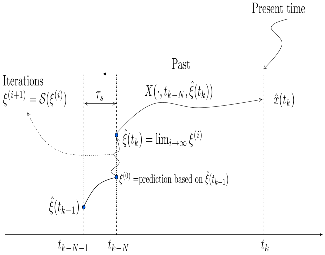
$$\Big\{J(t,\xi) \rightarrow 0\Big\} \Rightarrow \Big\{\underbrace{X(t, t-T, \xi)}_{\hat{x}(t)} \rightarrow x(t)\Big\}$$

+ No need to study the dynamic of e
+ No need for structural assumptions
+ Observability $\Leftrightarrow$ Observer
+ Handling constraints on the state

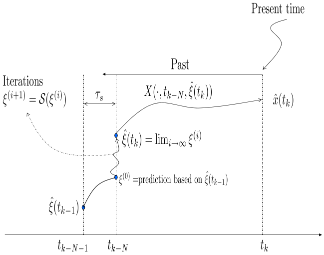
Potential problems

- Global convergence ?
- Computation time ?

Distributing the optimization over the system real-life time

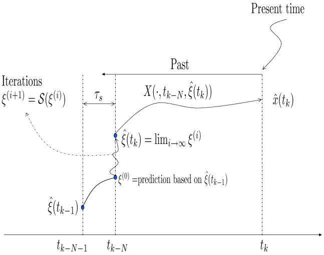


Distributing the optimization over the system real-life time



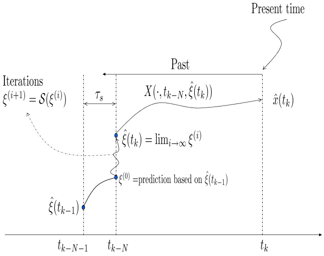
- $$\hat\xi(t_k) = \overbrace{arg \min_{\xi\in\mathbb{X}(t_{k-N})} \Big[J(t_k,\xi)\Big]}^{\text{Initial Guess } \xi^{(0)}}$$

Distributing the optimization over the system real-life time



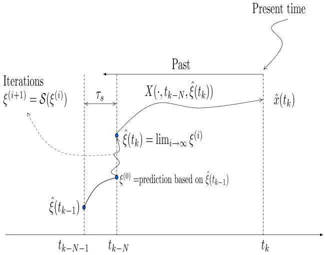
- $\hat\xi(t_k) = \overbrace{arg \min_{\xi\in\mathbb{X}(t_{k-N})} \Big[J(t_k,\xi)\Big]}^{\text{Initial Guess } \xi^{(0)}}$
- $\hat\xi(t_k) = \mathcal{S}^{N_{max}}(\xi^{(0)}, t_k, y_{t_{k-N}}^{t_k})$

Distributing the optimization over the system real-life time



- $\hat{\xi}(t_k) = \overbrace{arg \min_{\xi \in \mathbb{X}(t_{k-N})} \Big[J(t_k, \xi) \Big]}^{\text{Initial Guess } \xi^{(0)}}$
- $\hat{\xi}(t_k) = \mathcal{S}^{N_{max}}(\xi^{(0)}, t_k, y_{t_{k-N}}^{t_k})$
- $\xi^{(0)} = X(t_{k-N}, t_{k-N-1}, \hat{\xi}(t_{k-1}))$

Distributing the optimization over the system real-life time

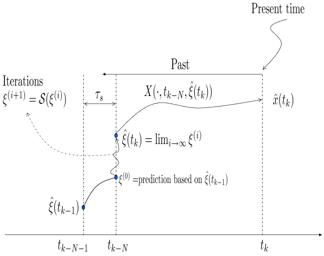


$\Rightarrow$ Implicit updating rule

$$\hat{\xi}(t_k) = F(t_k, \hat{\xi}(t_{k-1}), y_{t_{k-N}}^{t_k})$$

- $\hat{\xi}(t_k) = \overbrace{arg \min_{\xi \in \mathbb{X}(t_{k-N})} \Big[J(t_k, \xi) \Big]}^{\text{Initial Guess } \xi^{(0)}}$
- $\hat{\xi}(t_k) = \mathcal{S}^{N_{max}}(\xi^{(0)}, t_k, y_{t_{k-N}}^{t_k})$
- $\xi^{(0)} = X(t_{k-N}, t_{k-N-1}, \hat{\xi}(t_{k-1}))$

Distributing the optimization over the system real-life time



$\Rightarrow$ Implicit updating rule

$$\hat{\xi}(t_k) = F(t_k, \hat{\xi}(t_{k-1}), y_{t_{k-N}}^{t_k})$$

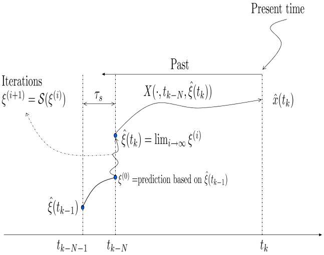
$\downarrow$

Is there a differential form of this updating rule ?

- $\hat{\xi}(t_k) = \overbrace{arg \min_{\xi \in \mathbb{X}(t_{k-N})} \Big[J(t_k, \xi)\Big]}^{\text{Initial Guess } \xi^{(0)}}$
- $\hat{\xi}(t_k) = \mathcal{S}^{N_{max}}(\xi^{(0)}, t_k, y_{t_{k-N}}^{t_k})$
- $\xi^{(0)} = X(t_{k-N}, t_{k-N-1}, \hat{\xi}(t_{k-1}))$

$$\frac{d\xi}{dt}(t) = f(t, \xi, y_{t-T}^t)$$

Distributing the optimization over the system real-life time



- $\hat{\xi}(t_k) = \overbrace{arg \min_{\xi \in \mathbb{X}(t_{k-N})} \Big[J(t_k, \xi) \Big]}^{\text{Initial Guess } \xi^{(0)}}$
- $\hat{\xi}(t_k) = \mathcal{S}^{N_{max}}(\xi^{(0)}, t_k, y_{t_{k-N}}^{t_k})$
- $\xi^{(0)} = X(t_{k-N}, t_{k-N-1}, \hat{\xi}(t_{k-1}))$

$\Rightarrow$ Implicit updating rule

$$\hat{\xi}(t_k) = F(t_k, \hat{\xi}(t_{k-1}), y_{t_{k-N}}^{t_k})$$

$\downarrow$

Is there a differential form of this updating rule ?

$$\frac{d\xi}{dt}(t) = f(t, \xi, y_{t-T}^{t})$$

$\rightarrow$ Differential form for moving horizon observers

Differential Form of Moving-Horizon Observers: Outline

For system of the form $\dot{x} = f(t, x)$:

$$\begin{aligned}\dot{\xi}(t) &= f(t-T, \xi(t)) + c(t, \xi(t))\\ \hat{x}(t) &= X(t, t-T, \xi(t))\end{aligned}$$

- The correction term

$$c(t,\xi) := \gamma\left[\frac{{J_\xi}^T(t,\xi)}{\|J_\xi\|^2+\varepsilon}\right]\left[-\left|\Delta_{t-T}^{t}(\epsilon_y(\cdot,\xi))\right| - \left[1+\phi(t,\xi)\right]\sqrt{J}\right]$$

- Post-Stabilization technique $\rightarrow$ improve (Sampling period)/Precision ratio

- [M. Alamir *Optimization Based Nonlinear Observer Revisited* ... Int. J. of Control 1999]

Differential Form of Moving-Horizon Observers: Outline

For system of the form $\dot{x} = f(t,x)$:

$$\begin{aligned}\dot{\xi}(t) &= f(t-T,\xi(t)) + c(t,\xi(t)) \\ \hat{x}(t) &= X(t,t-T,\xi(t))\end{aligned}$$

Uniform Global Regularity Assumption

There is a K-function $\Upsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following inequality holds:

$$\|J_\xi(t,\xi)\|^2 \geq \Upsilon(J(t,\xi))$$

for all (t,ξ) ♡

- The correction term

$$c(t,\xi) := \gamma\left[\frac{{J_\xi}^T(t,\xi)}{\|J_\xi\|^2+\varepsilon}\right]\left[-\left|\Delta_{t-T}^t(\epsilon_y(\cdot,\xi))\right| - \left[1+\phi(t,\xi)\right]\sqrt{J}\right]$$

- Post-Stabilization technique → improve (Sampling period)/Precision ratio

- [M. Alamir *Optimization Based Nonlinear Observer Revisited* ... Int. J. of Control 1999]

Convergence result

If the following conditions hold:

1. The maps are continuously differentiable
2. The system is uniformly observable
3. The uniform regularity assumption is satisfied

then for any a priori fixed desired precision $\eta > 0$ on the state estimation error, there is a sufficiently high ratio γ/ε such that the dynamic system given by:

$$\begin{aligned} \dot{\xi}(t) &= f(t-T,\xi(t)) + c(t,\xi(t)) \\ \hat{x}(t) &= X(t,t-T,\xi(t)) \end{aligned}$$

where the correction term $c(t,\xi)$ is given by:

$$c(t,J) := \gamma\left[\frac{J_\xi{}^T(t,\xi(t))}{\|J_\xi\|^2+\varepsilon}\right]\left[-\left|\Delta_{t-T}^t(\epsilon_y(\cdot,\xi(t)))\right| - \left[1+\phi(t,\xi(t))\right]\sqrt{J}\right]$$

leads to an estimation error that is asymptotically lower than η. ♡

The Post Stabilization Technique

$$\begin{aligned}
\dot{\xi}(t) &= f_c(t, \xi(t), J_\xi(t)) \\
\hat{x}(t) &= X(t, t-T, \xi(t)) \\
J(t, \xi(t)) &= 0 \quad \text{(Ideally)}
\end{aligned}$$

The Post Stabilization Technique

$$\begin{aligned}
\dot{\xi}(t) &= f_c(t, \xi(t), J_\xi(t)) \\
\hat{x}(t) &= X(t, t-T, \xi(t)) \\
J(t, \xi(t)) &= 0 \quad \text{(Ideally)}
\end{aligned}$$

Bad News: The computation of the r.h.s of the observer equation is expensive.

[Integration of a differential system of order $n(n+1)$]

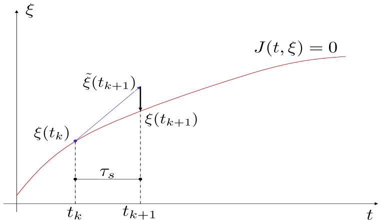
→ For a given sampling period, need for lower order integration methods

The Post Stabilization Technique

$$
\begin{aligned}
\dot{\xi}(t) &= f_c(t,\xi(t),J_\xi(t)) \\
\hat{x}(t) &= X(t,t-T,\xi(t)) \\
J(t,\xi(t)) &= 0 \quad \text{(Ideally)}
\end{aligned}
$$

Good News: Ideally, DAE's with invariant submanifold

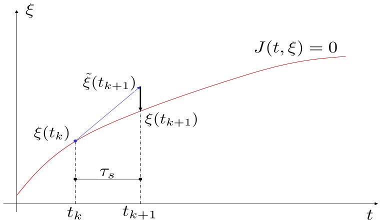
→ There are techniques (Ascher, Num. Alg. 1997) to accurately integrate with lower order methods



1. Integrate over $[t_k, t_{k+1}]$ starting from the initial condition $(t_k, \xi(t_k))$

$$\dot{\xi}(t) = f_c(t, \xi(t), J_\xi(t_k)) \quad ; \quad t \in [t_k, t_{k+1}]$$

to obtain $\tilde{\xi}(t_{k+1})$



1. Integrate over $[t_k, t_{k+1}]$ starting from the initial condition $(t_k, \xi(t_k))$

$$\dot{\xi}(t) = f_c(t, \xi(t), J_\xi(t_k)) \quad ; \quad t \in [t_k, t_{k+1}]$$

   to obtain $\tilde{\xi}(t_{k+1})$

2. Correct the *rough* approximation $\tilde{\xi}(t_{k+1})$ by projection

$$\xi(t_{k+1}) = \tilde{\xi}(t_{k+1}) - \frac{J_\xi(t_{k+1}, \tilde{\xi}(t_{k+1}))}{\|J_\xi(t_{k+1}, \tilde{\xi}(t_{k+1}))\|^2 + \nu} \cdot J\big(t_{k+1}, \tilde{\xi}(t_{k+1})\big)$$

   to obtain the update $\xi(t_{k+1})$

Convergence analysis

Regardless the order of the integration scheme, one has

$$\lim_{k\to\infty} J(\xi(t_k)) = O(\tau_s^4)$$

1. Integrate over $[t_k, t_{k+1}]$ starting from the initial condition $(t_k, \xi(t_k))$

$$\dot{\xi}(t) = f_c(t, \xi(t), J_\xi(t_k)) \quad ; \quad t \in [t_k, t_{k+1}]$$

   to obtain $\tilde{\xi}(t_{k+1})$

2. Correct the *rough* approximation $\tilde{\xi}(t_{k+1})$ by projection

$$\xi(t_{k+1}) = \tilde{\xi}(t_{k+1}) - \frac{J_\xi(t_{k+1}, \tilde{\xi}(t_{k+1}))}{\|J_\xi(t_{k+1}, \tilde{\xi}(t_{k+1}))\|^2 + \nu} \cdot J\big(t_{k+1}, \tilde{\xi}(t_{k+1})\big)$$

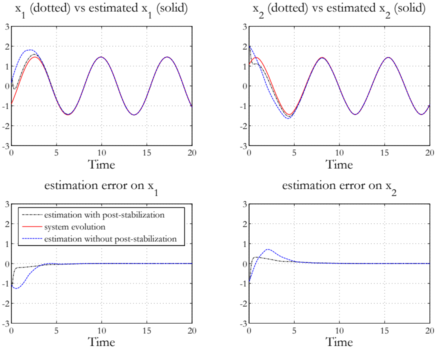
   to obtain the update $\xi(t_{k+1})$

The benefit from the post-stabilization technique

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -\sin(x_1) - 0.2x_1\cos(x_1x_2)\\ y &= x_1 + x_2\end{aligned}$$

$\tau_s = 0.1$

Almost no need for Post-stabilization step

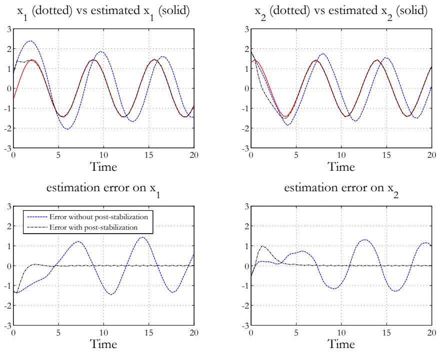


The benefit from the post-stabilization technique

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\sin(x_1) - 0.2x_1\cos(x_1x_2)$$
$$y = x_1 + x_2$$

$\tau_s = 0.4$

Post-stabilization is mandatory to keep precision under high sampling period.



Moving-Horizon Observers with Distributed Optimization

- The system

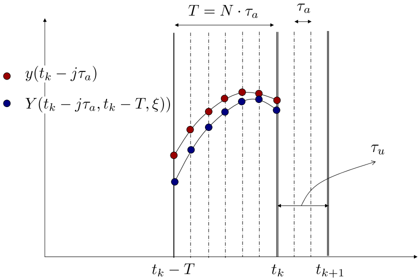
$$\begin{aligned} x(t) &= X(t, t_0, x_0), \\ y(t) &= h(t, x(t)), \end{aligned}$$

- Measurement acquisition period τ_a
- Updating period $\tau_u = N_u \cdot \tau_a$
- Updating instants $t_k = k \cdot \tau_u$
- Observation horizon $T = N \cdot \tau_a$
- Cost function at instant t_k: $J(t_k, \xi)$

Moving-Horizon Observers with Distributed Optimization

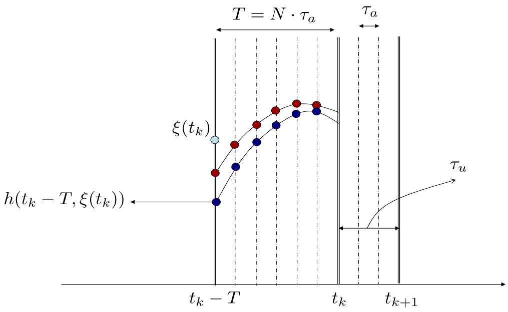
- The system

$$x(t) = X(t, t_0, x_0),$$
$$y(t) = h(t, x(t)),$$

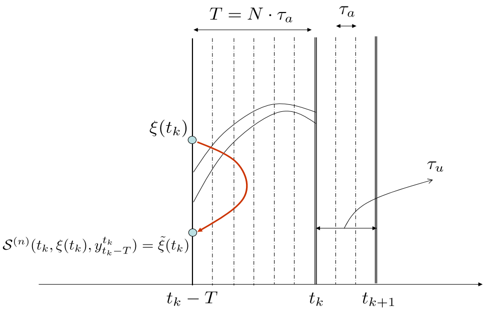


- Measurement acquisition period
- Updating period $\tau_u = N_u \cdot \tau_a$
- Updating instants $t_k = k \cdot \tau_u$
- Observation horizon $T = N \cdot \tau_a$
- Cost function at instant t_k: $J(t_k, \xi)$

Moving-Horizon Observers with Distributed Optimization



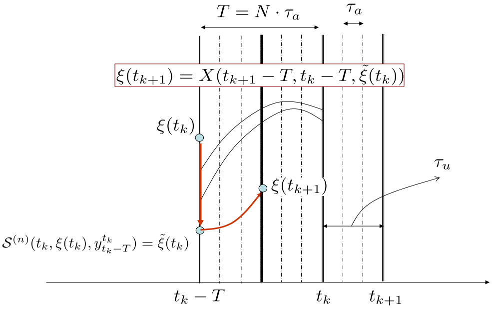
Moving-Horizon Observers with Distributed Optimization



During the interval $[t_k, t_{k+1}]$, perform n iterations of some iterative process $\mathcal{S}$:

$$\tilde{\xi}(t_k) = \mathcal{S}^{(n)}\Big(t_k, \xi(t_k), y_{t_k-T}^{t_k}\Big)$$

Moving-Horizon Observers with Distributed Optimization



During the interval $[t_k, t_{k+1}]$, perform n iterations of some iterative process $\mathcal{S}$:

$$\tilde{\xi}(t_k) = \mathcal{S}^{(n)}\Big(t_k, \xi(t_k), y_{t_k - T}^{t_k}\Big)$$

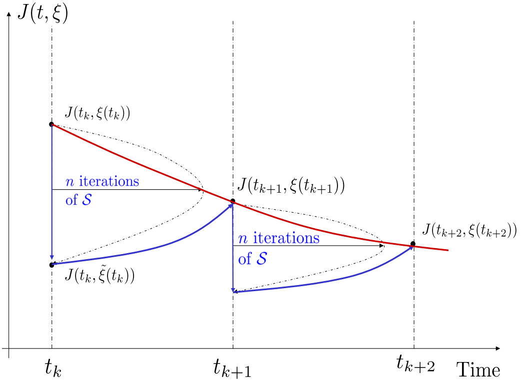
Using $\tilde{\xi}(t_k)$, update the value of $\xi(t_{k+1})$ according to

$$\xi(t_{k+1}) = X(t_{k+1} - T, t_k - T, \tilde{\xi}(t_k))$$

Two opposite processes

The updating mechanism involves two opposite effects on $J(t_k, \xi(t_k))$:

1. A decreasing effect from the n-iterations of the optimization process
2. An increasing effect from the open loop prediction over $\tau_u = N\tau_a$.



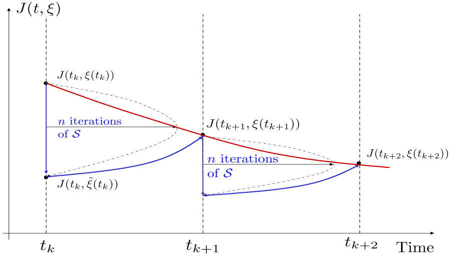
Moving-Horizon Observers with Distributed Optimization

Assumption: Efficiency of the optimizer

The iterative process $\mathcal{S}$ is efficient in the sense that there exists some efficiency map $\alpha_{\mathit{eff}} : \mathbb{N} \rightarrow [0, 1[$ such that for all t and ξ, one has:

$$J\Big(t, \mathcal{S}^{(n)}(t, \xi, y_{t-T}^{t})\Big) \leq \alpha_{\mathit{eff}}(n) \cdot J(t, \xi)$$

where $\alpha(\cdot)$ is a decreasing function such that $\alpha(0) = 1$. ♡



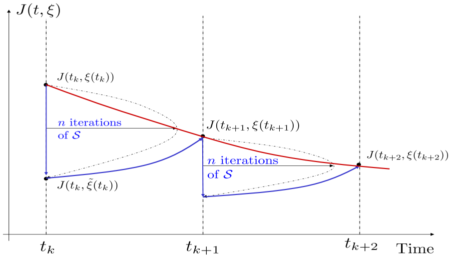
Moving-Horizon Observers with Distributed Optimization

Assumption: Open-loop behavior of the cost function

When using open-loop prediction, the only inequality one can guarantee is given by:

$$J(t+\tau, X(t+\tau-T, t-T, \xi)) \leq \Big[J(t,\xi)\Big] \cdot \vartheta(\tau) \quad (1)$$

♡



Moving-Horizon Observers with Distributed Optimization

A rather qualitative result

Under the assumptions above, the convergence of the distributed in time optimization based observer is guaranteed provided that the following inequality holds:

$$\varpi(N_u) := \alpha_{eff}\left(E\Big(\frac{N_u\tau_a}{\tau_{iter}}\Big)\right)\cdot\vartheta(N_u\tau_a) < 1 \quad (2)$$

Moreover, the *convergence time* is given by:

$$t_r(N_u) \approx \left[\frac{3N_u}{\left|\log\Big(\varpi(N_u)\Big)\right|}\right]\cdot\tau_a \quad (3)$$

where

- ✓ τ_a is the measurement acquisition period
- ✓ $N_u\tau_a$ is the updating period
- ✓ τ_{iter} is the time necessary to perform one iteration of the process $\mathcal{S}$ ♠

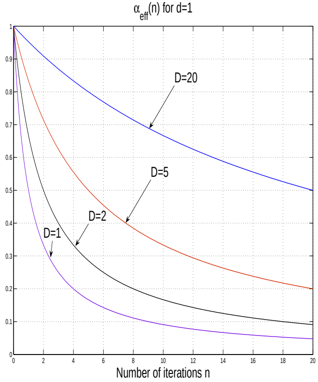
Moving-Horizon Observers with Distributed Optimization

Take the following example for illustration

$$\alpha_{eff}(n) = \frac{D}{n^d + D} \quad ; \quad \vartheta(\tau) = \exp(\beta \cdot \tau)$$

Note that:

- $d \nearrow$ increases the efficiency
- $D \nearrow$ decreases the efficiency
- $\alpha_{eff}(0) = 1$
- $\beta \nearrow$ assumes high model discrepancy



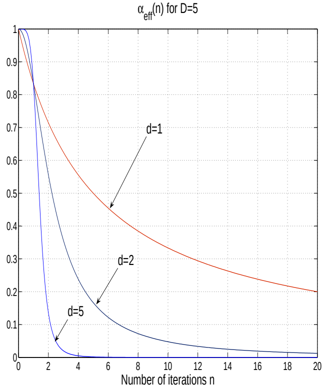
Moving-Horizon Observers with Distributed Optimization

Take the following example for illustration

$$\alpha_{eff}(n) = \frac{D}{n^d + D} \quad ; \quad \vartheta(\tau) = \exp(\beta \cdot \tau)$$

Note that:

- $d \nearrow$ increases the efficiency
- $D \nearrow$ decreases the efficiency
- $\alpha_{eff}(0) = 1$
- $\beta \nearrow$ assumes high model discrepancy



Moving-Horizon Observers with Distributed Optimization

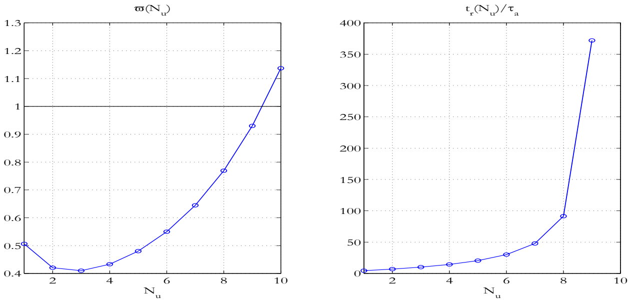


Figure: Evolutions of the stability indicator $\varpi(N_u)$ and the settling time $t_r(N_u)$ vs the number of iterations N_u used to update the state estimation. $D = 3$, $d = 1$, $\beta \cdot \tau_a = 0.3$ and $\tau_a/\tau_{iter} = 5$. Under these conditions, stability cannot be guaranteed when more that 9 iterations are used. The optimal choice (in term of settling time) is the one where only one iteration is used to perform the updating.

Moving-Horizon Observers with Distributed Optimization

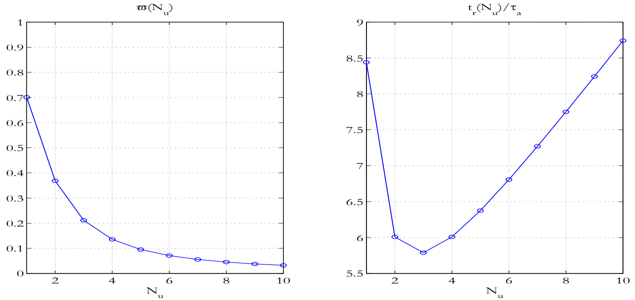
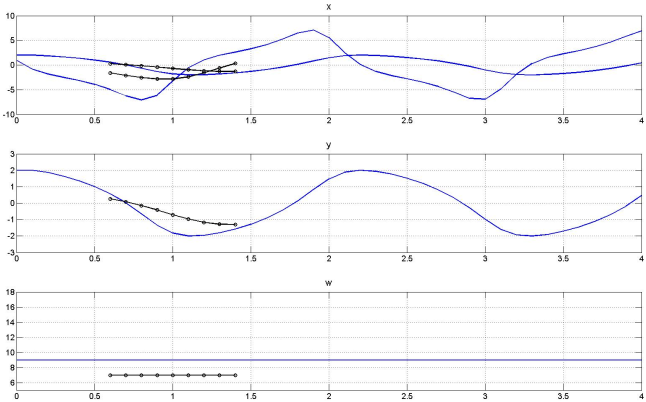


Figure: Evolutions of the stability indicator $\varpi(N_u)$ and the settling time $t_r(N_u)$ vs the number of iterations N_u used to update the state estimation. $D = 50$, $d = 2$, $\beta \cdot \tau_a = 0.05$ and $\tau_a/\tau_{iter} = 5$. Under these conditions, while stability seems guaranteed regardless the number of iterations used to perform the updating, the use of 3 iterations gives the best result in term of settling time.

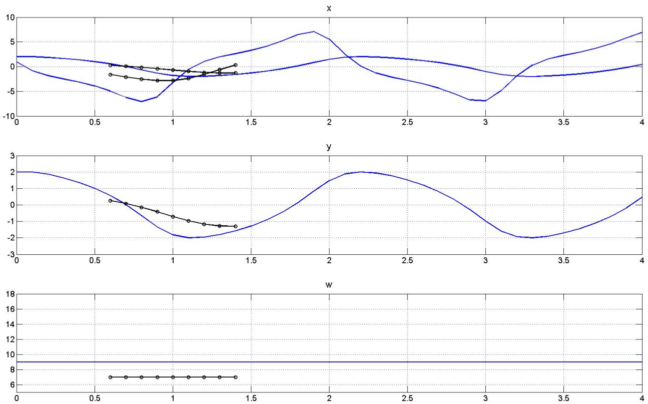
Consider the following state/parameter estimation problem:

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -px_1 + 2(1 - x_1^2)x_2 \\
\\
y &= x_1
\end{aligned}
$$

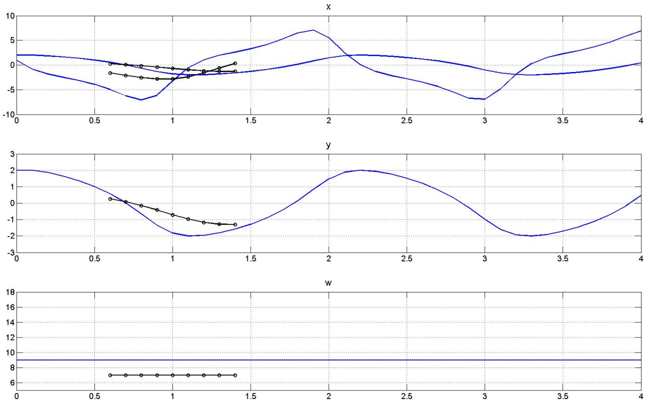
with $p \in [5, 10]$



10 function evaluations



6 function evaluations



3 function evaluations

Moving-Horizon Observers with Distributed Optimization

The success and the quality of the Moving-Horizon-Observer depend on

- The quality of the optimizer (d, D)
- The quality of the model (β)
- The iteration complexity (τ_{iter})
- The problem itself (The very existence of such parameters)

On-line identification of the problem parameters ?

$\Rightarrow$ On-line adaptation of the updating rate ?

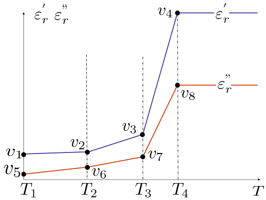
ANR-Blanc: Capteurs Logiciels Plug & Play

- Partners: Gipsa-lab (Grenoble) / GEPEA (Nantes) / LAGEP (Lyon)
- Fast Prototyping of Robust NL Observers
- User Friendly (you need just to know your process)
- Help to detect bad conditioned problems
- Laboratory scale moderate size processes
- Freely available for French Universities (Late 2010)
- http://www.mazenalamir.fr/ANR_CLPP/

Example: Micro-Wave Tempering

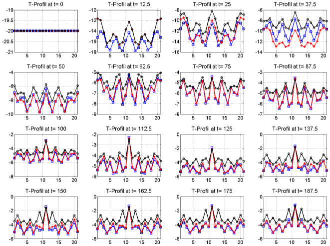
Physics →

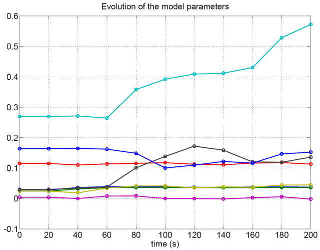
$$\frac{\partial T}{\partial t} = F(T, T_z, u, \underbrace{\varepsilon_r'(T), \varepsilon_r''(T)}_{??})$$



Discretization →

$$\dot{x} = f(x, u, p) \quad (x, p) \in \mathbb{R}^{21} \times [0, 1]^8$$

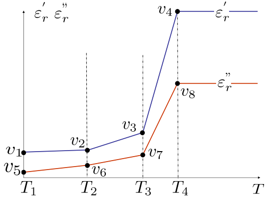




Example: Micro-Wave Tempering

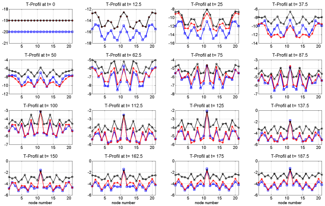
Physics $\rightarrow$

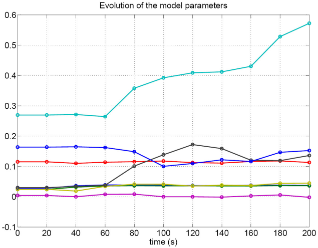
$$\frac{\partial T}{\partial t} = F(T, T_z, u, \underbrace{\varepsilon_r^{'}(T), \varepsilon_r^{"}(T)}_{??})$$



Discretization $\rightarrow$

$$\dot{x} = f(x, u, p) \quad (x, p) \in \mathbb{R}^{21} \times [0, 1]^8$$





General Conclusion

- The progress in MHO $\longleftarrow$ progress of optimization tools
- MHO-related problem is NOT ONLY an optimization problem
- Promising direction:

  Combine Analytic and Optimization Based Observers

  (Let the MHO concentrate on the structure-free part of the problem)

Further readings

- Alamir M., J. S. Welsh and G. C. Goodwin *Redundancy vs multiple starting points in nonlinear systems related inverse problems* Automatica, Vol. 45 pp. 1052-1057 (2009).

- Kuehl P, M. Diehl, T. Kraus, J. Schloeder and G. Bock. *A real-time algorithm for moving horizon state and parameter estimation* (Submitted to Computers and Chemical Engineering), 2010.

- Michalska, H & Mayne D. Q. *Moving Horizon Observers and Observer Based Control* IEEE TAC 1995.

- Zimmer G.*State Observation by On-line Optimization* Int. J. of Control 1994.

- Rao C. V., Rawlings J. B. and Mayne D. Q. *Constrained State Estimation . . .* IEEE TAC 2003.

- Rao C. V. and Rawlings J. B. *Constrained Process Monitoring: Moving-Horizon . . .* AIChE Journal 2004.

- Ohtsuka T. and Fujii H. A. *Nonlinear Receding Horizon . . .* J. Guidance, Control, Dynamics 1996.

- Alvarado I. A., Findeisen R., Kuhl P., . . . *State Estimation for Repetitive . . .* IFAC WC barcelona, 2002.

- Wei Kang *Moving horizon numerical observers of nonlinear control systems* IEEE TAC 2006.

- Alamir M. *Optimization Based Nonlinear Observer Revisited . . .* Int. J. of Control 1999.

- Alamir M. and Calvillo-Corona L. *Further Results on Nonlinear Receding Horizon Observers . . .* IEEE TAC 2002.

- Ascher U M. *Stabilization of Invariants of Discretized Diff. Syst.* Numerical Algorithms, 1997.

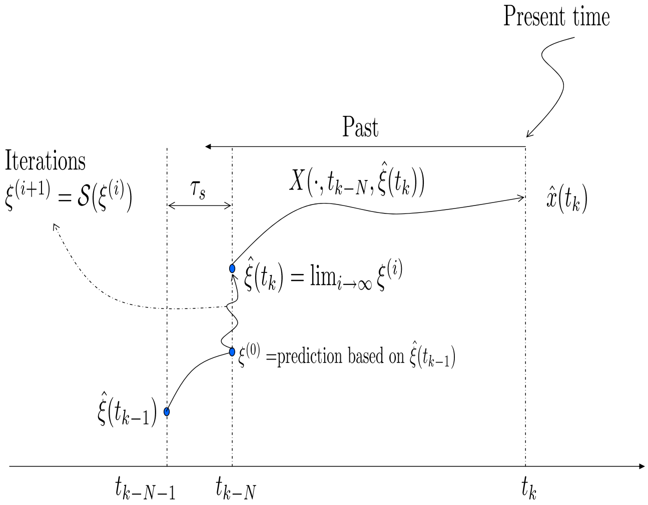
- Othman N., McKenna T. F. and Févotte G. *On-Line Monitoring . . .* Polymer Reaction Engineering, 2001.

- Ferrari-Trecate G. Mignone D. and Morari M. *MHE for Hybrid Systems . . .* ACC 2000.

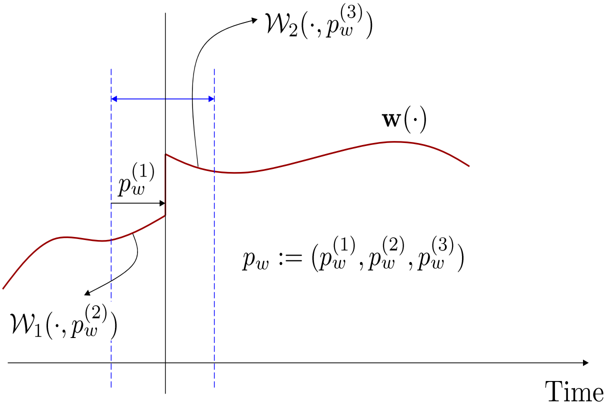
- Zavala V. M., Laird C. D. and Biegler L. T. *A Fast Computation Framework for Large Scale MHE* Int. Symp. On Dyn. and Contr. of Process Syst 2007.

- Alamir, M., Sheibat-Othman, N. and Othman, S. *On the use of Nonlinear Receding Horizon Observers in Batch Terpolymerization Reactors* International Journal of Modeling Identification and Control, Vol 4, No 1 (2008).

- Alamir, M. *Nonlinear Moving Horizon Observers: Theory and Real-Time Implementation* In Nonlinear Observers and Applications. Lecture Notes in Control and Information Sciences, Springer-Verlag, (2007).


Present time
Past
Iterations
$\xi^{(i+1)} = \mathcal{S}(\xi^{(i)})$
τ_s
$X(\cdot, t_{k-N}, \hat{\xi}(t_k))$
$\hat{x}(t_k)$
$\hat{\xi}(t_k) = \lim_{i\to\infty} \xi^{(i)}$
$\xi^{(0)}$ =prediction based on $\hat{\xi}(t_{k-1})$
$\hat{\xi}(t_{k-1})$
t_{k-N-1}
t_{k-N}
t_k

Non smooth behaviors can be parameterized



$$\mathcal{W}(\cdot, p_w) := \begin{cases} \mathcal{W}_1(\tau, p_w^{(2)}) & \text{if } \tau \leq p_w^{(1)} \\ \mathcal{W}_2(\tau, p_w^{(3)}) & \text{otherwise} \end{cases}$$

Definition of the phase II: Existence of monomer droplets

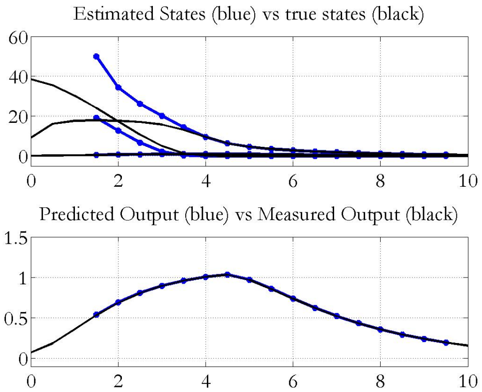
$$N_1\delta_1 + N_2\delta_2 + N_3\delta_3 - \frac{(1-\phi_p^p)}{\phi_p^p}\sigma > 0 \tag{4}$$

where

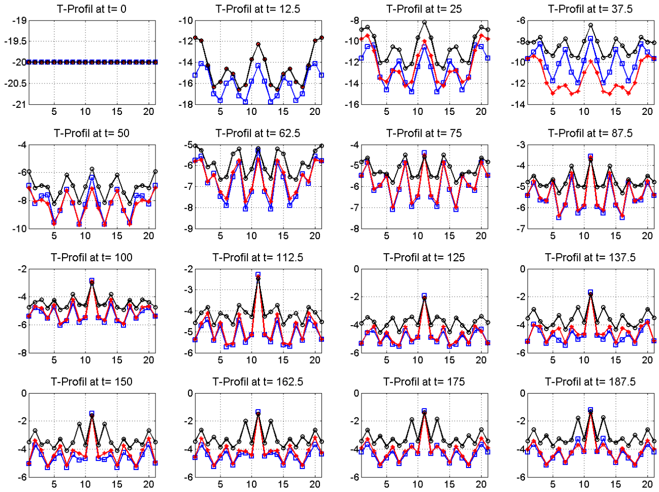
$$\delta_i = MWi(\frac{1}{\rho_i} + \frac{(1-\phi_p^p)}{\rho_{i,h}\phi_p^p}) \ , \quad i = 1,2,3 \tag{5}$$

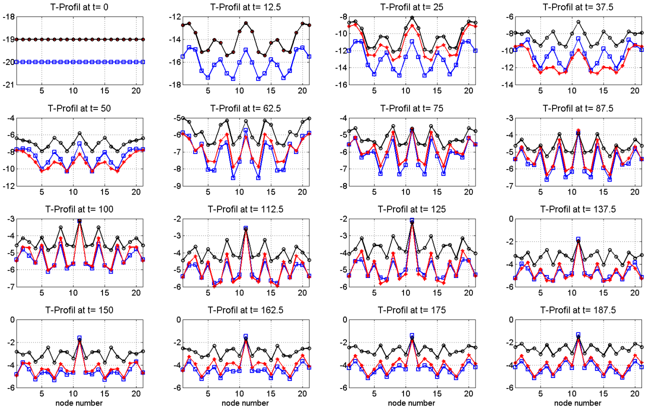
and

$$\sigma = \sum_{j=1}^{3} \frac{MW_j N_j^T}{\rho_j, h} \tag{6}$$



Bad Conditioning of the Photo-Bio-reactor Problem


T-Profil at t= 0
T-Profil at t= 12.5
T-Profil at t= 25
T-Profil at t= 37.5
T-Profil at t= 50
T-Profil at t= 62.5
T-Profil at t= 75
T-Profil at t= 87.5
T-Profil at t= 100
T-Profil at t= 112.5
T-Profil at t= 125
T-Profil at t= 137.5
T-Profil at t= 150
T-Profil at t= 162.5
T-Profil at t= 175
T-Profil at t= 187.5


T-Profil at t= 0
T-Profil at t= 12.5
T-Profil at t= 25
T-Profil at t= 37.5
T-Profil at t= 50
T-Profil at t= 62.5
T-Profil at t= 75
T-Profil at t= 87.5
T-Profil at t= 100
T-Profil at t= 112.5
T-Profil at t= 125
T-Profil at t= 137.5
T-Profil at t= 150
T-Profil at t= 162.5
T-Profil at t= 175
T-Profil at t= 187.5
node number

[illegible]
results vs. closed-form expressions. *AIChE Journal*, 55(6):1569–1583, 2009.

[9] K. W. Farag, J. G. Lyng, D. J. Morgan, and D. A. Cronin. Dielectric and thermophysical properties of different beef meat blends over a temperature range of 19 to +10 ^{o}c. *Meat Science*, 79(4):740., 2008.

[10] J. P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems, application to bioreactors. *IEEE Transactions on Automatic Control*, 37:875–880, 1992.

[11] A. G. Herve, J. Tang, L. Luedeck, and H. Feng. Dielectric properties of cottage cheese and surface treatment using microwaves. *Journal of Food Engineering*, 37(4):389–410, 1998.

[12] X. P. Hu and P. Mallikarjunan. Thermal and dielectric properties of shucked oysters. *Lwt-Food Science and technology*, 38(5):489–494, 2005.