

Crossing Singularities in Dynamic Inverse Optimization Problems

Mazen Alamir

CNRS, University of Grenoble



Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Analytic nonlinear observers

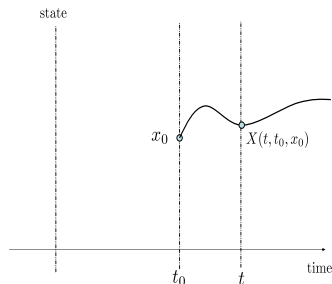
- Need structural properties
 - Unable to handle constraints
- Optimization based observers

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Analytic nonlinear observers

- Need structural properties
 - Unable to handle constraints
- Optimization based observers



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

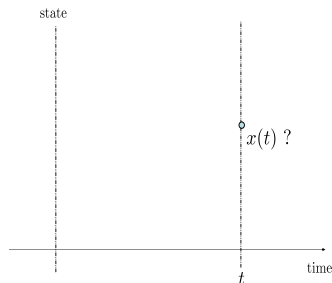
$$y = h(x)$$

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Analytic nonlinear observers

- Need structural properties
 - Unable to handle constraints
- Optimization based observers



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

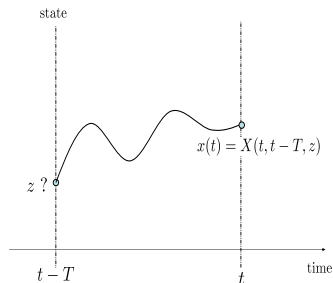
$$y = h(x)$$

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Analytic nonlinear observers

- Need structural properties
 - Unable to handle constraints
- Optimization based observers



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

$$y = h(x)$$

Many systems

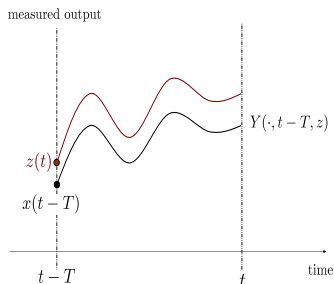
- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

Analytic nonlinear observers

- Need structural properties
- Unable to handle constraints

→ Optimization based observers

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



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

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

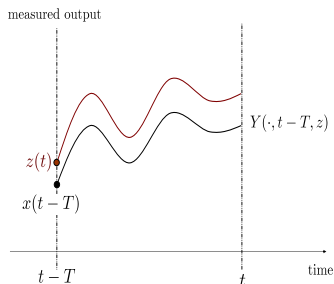
Analytic nonlinear observers

- Need structural properties
- Unable to handle constraints

→ Optimization based observers

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

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



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

Many systems

- ▶ Highly nonlinear dynamic
 - ▶ Badly known parameters
 - ▶ Lack of sensors
- Need for nonlinear observers

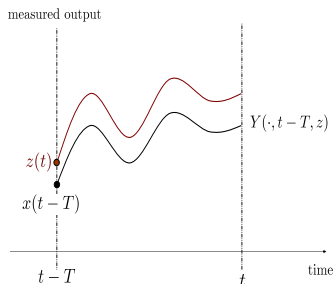
Analytic nonlinear observers

- Need structural properties
- Unable to handle constraints

→ Optimization based observers

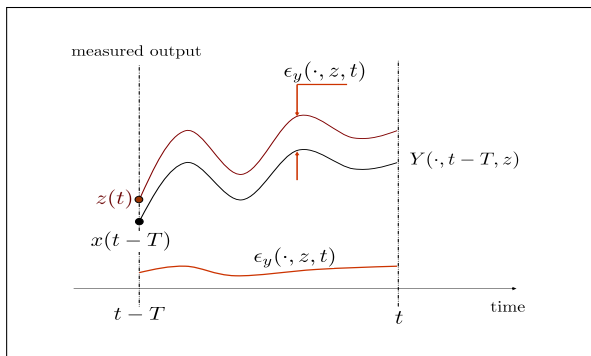
$$z_{opt}(t) \leftarrow \arg \min_z J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

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



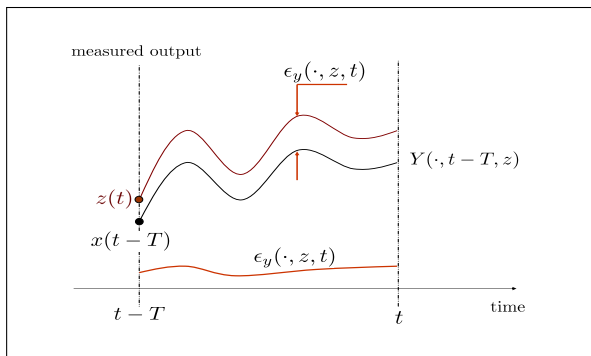
- ▶ *Generally a non convex optimization problem*
- ▶ *many local minima*

A very particular problem ...!



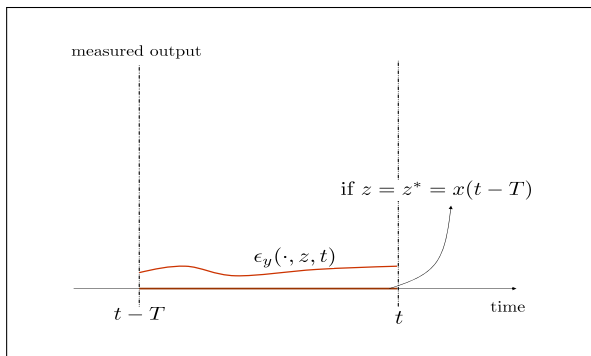
$$z_{opt}(t) \leftarrow \arg \min_z J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|Y(\tau, t-T, z) - y(\tau)\|_Q^2 d\tau$$

A very particular problem ...!



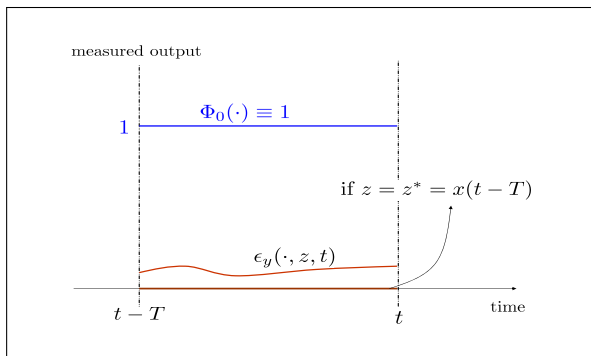
$$z_{opt}(t) \leftarrow \arg \min_z J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|\epsilon_y(\tau, z, t)\|_Q^2 d\tau$$

A very particular problem ...!



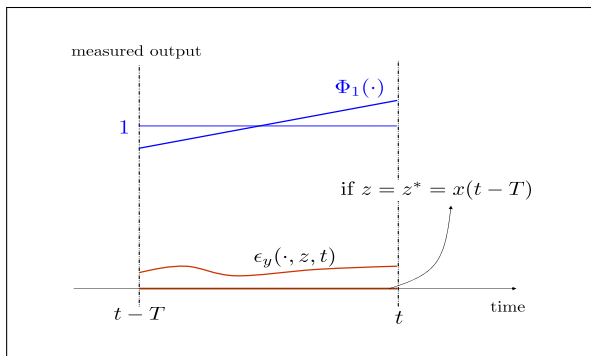
$$z_{opt}(t) \leftarrow \arg \min_z J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem ...!



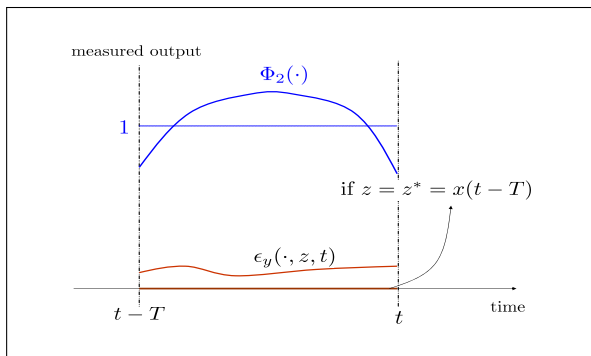
$$z_{opt}(t) \leftarrow \arg \min_z J_0(z, t, y_{t-T}^t) = \int_{t-T}^t \Phi_0(\tau) \cdot \|\epsilon_y(\tau, z^*, t)\|_Q^2 d\tau = 0$$

A very particular problem ...!



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

A very particular problem ...!



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

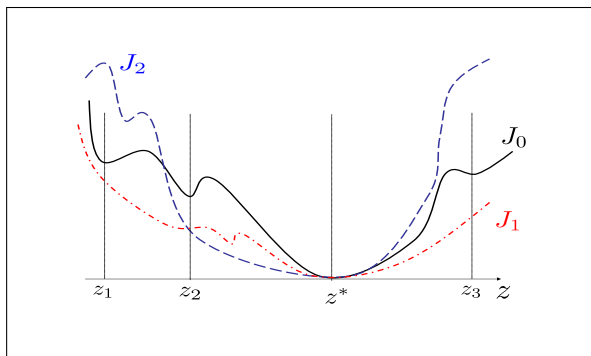
A very particular problem ...!



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

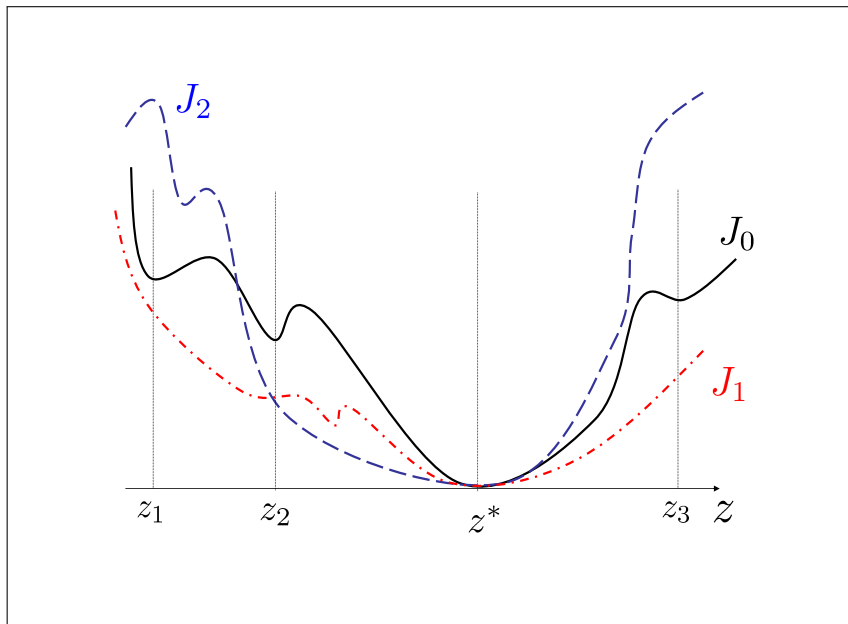
$$J_i(z, t, y_{t-T}^t) = \int_{t-T}^t \Phi_i(\tau) \cdot \Psi(\epsilon(\tau, z, t)) d\tau$$

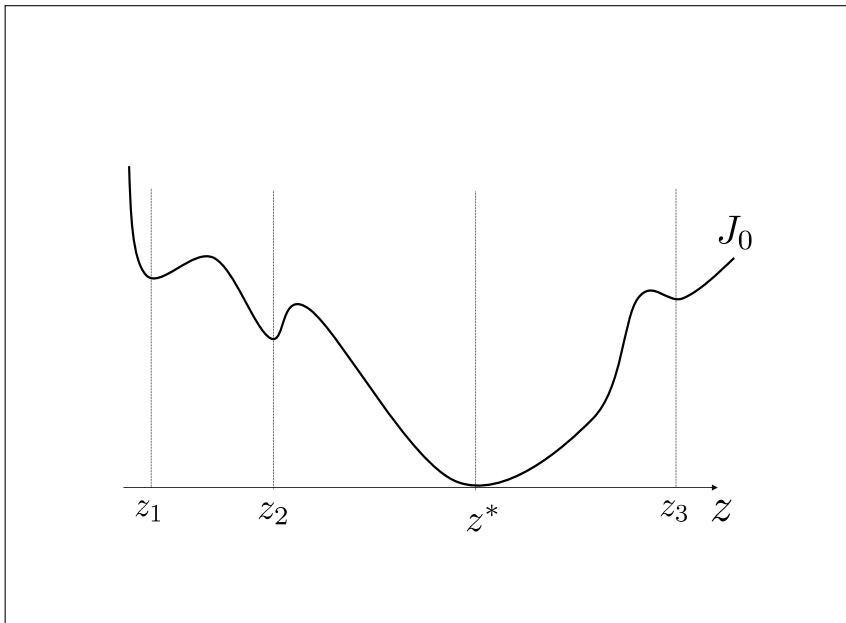
A very particular problem ...!

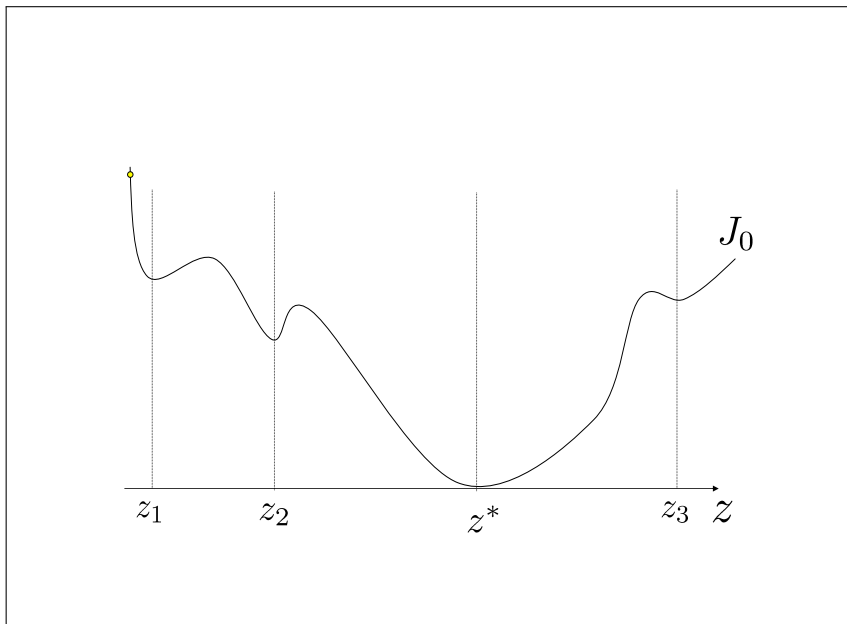


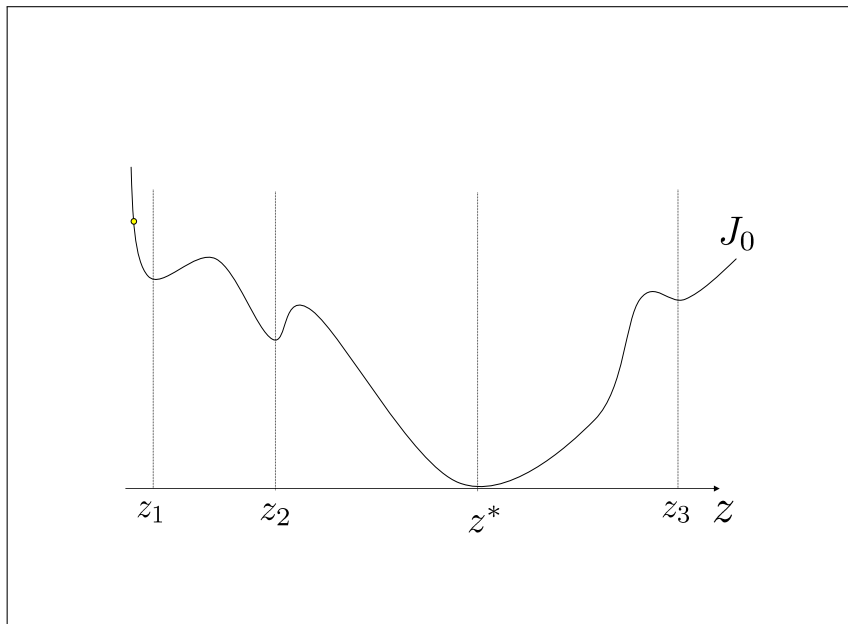
z^* is **THE** global minimum of **ALL** the cost functions J_i s.t:

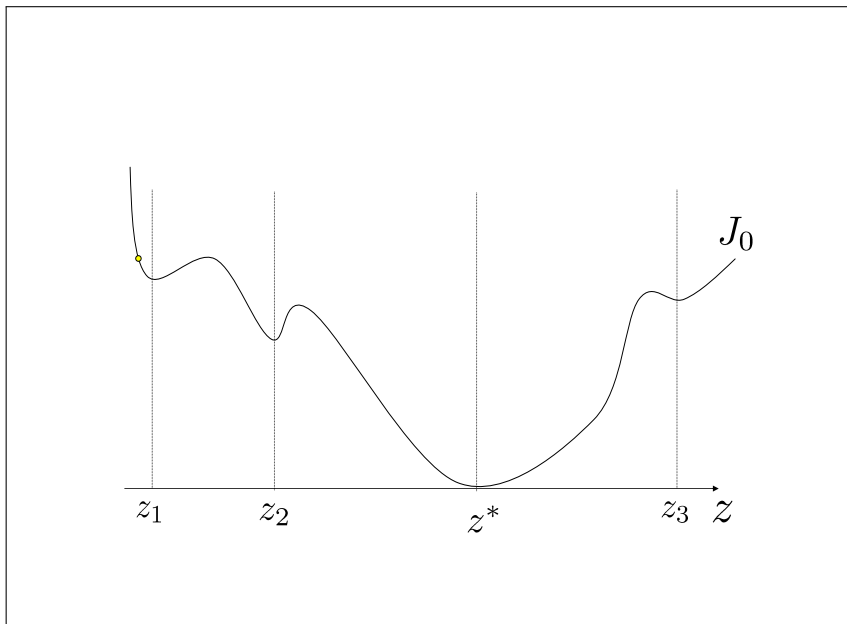
$$J_i(z, t, y_{t-T}^t) = \int_{t-T}^t \Phi_i(\tau) \cdot \Psi(\epsilon(\tau, z, t)) d\tau$$

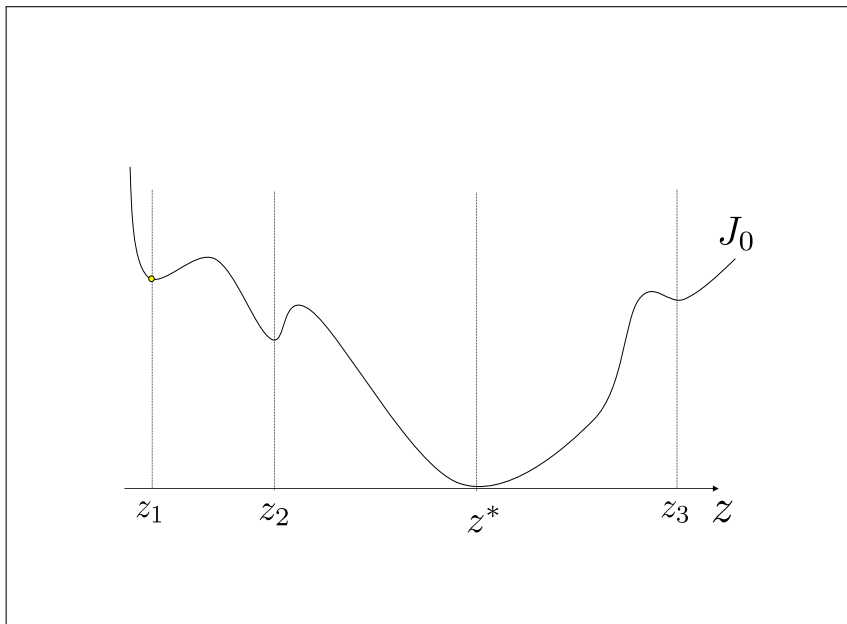


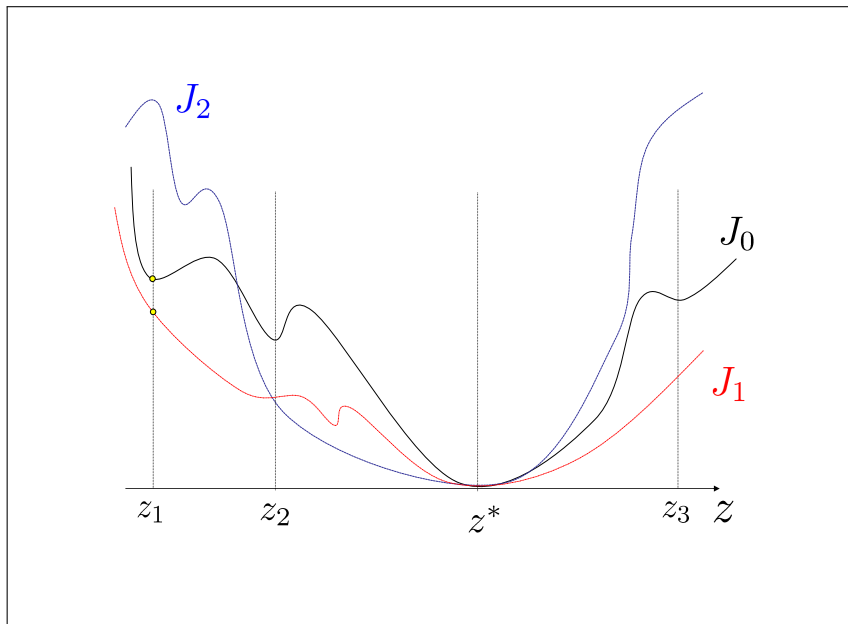


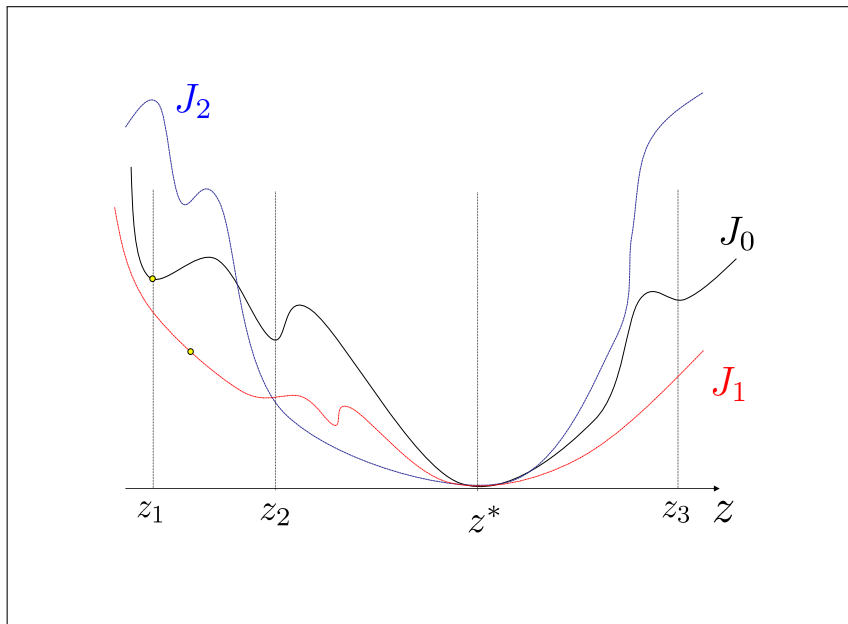


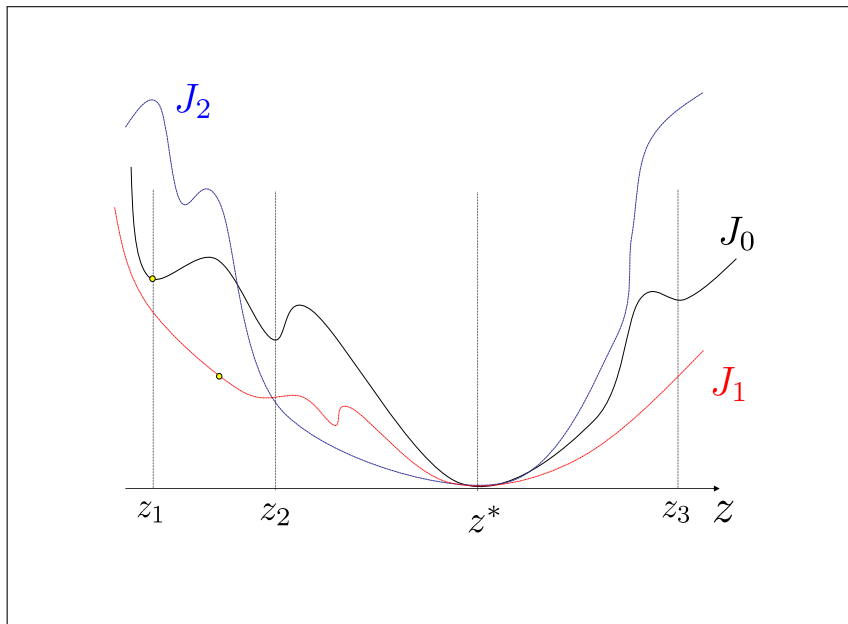


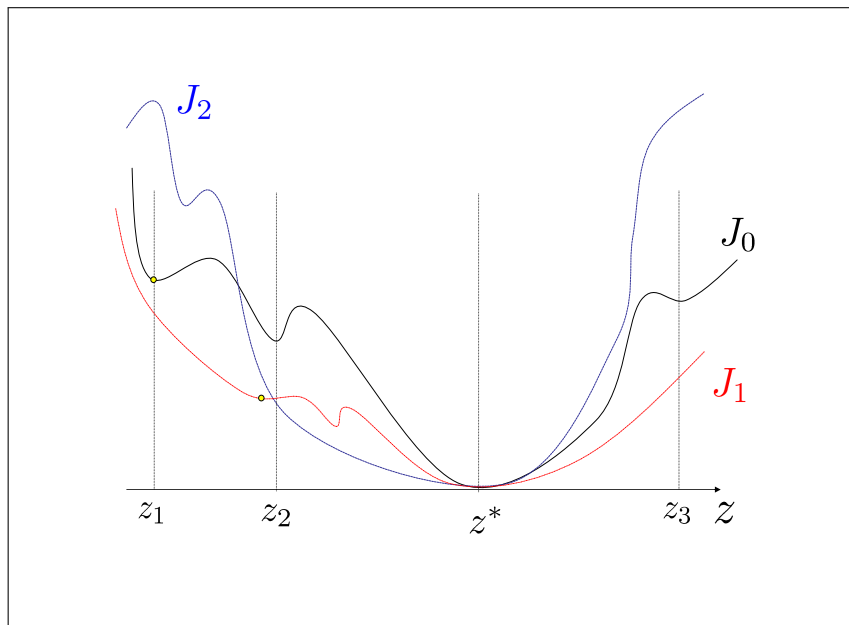


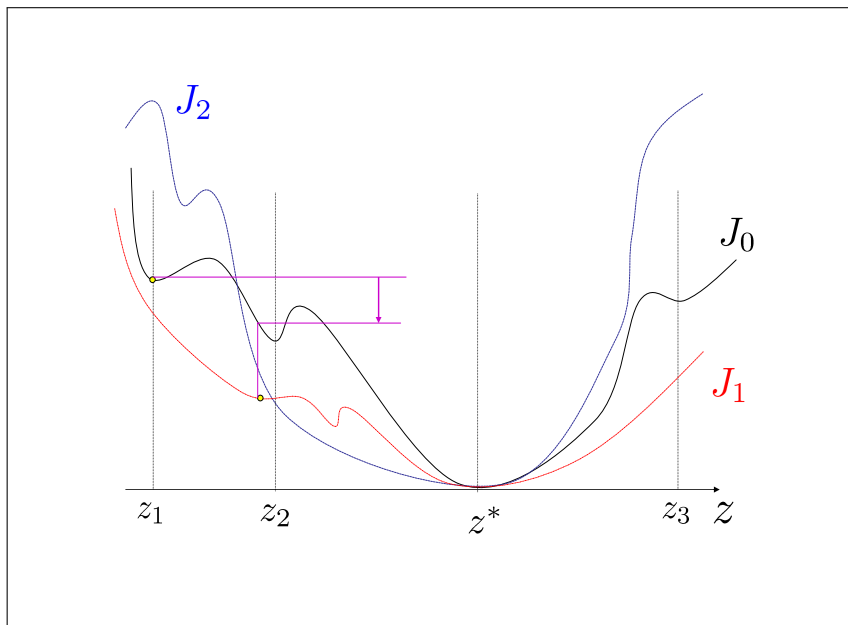


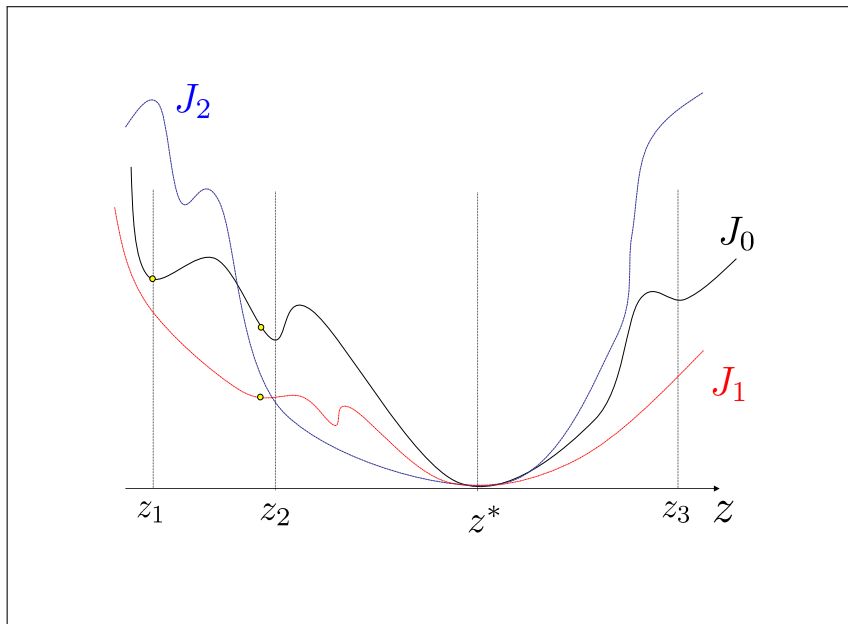


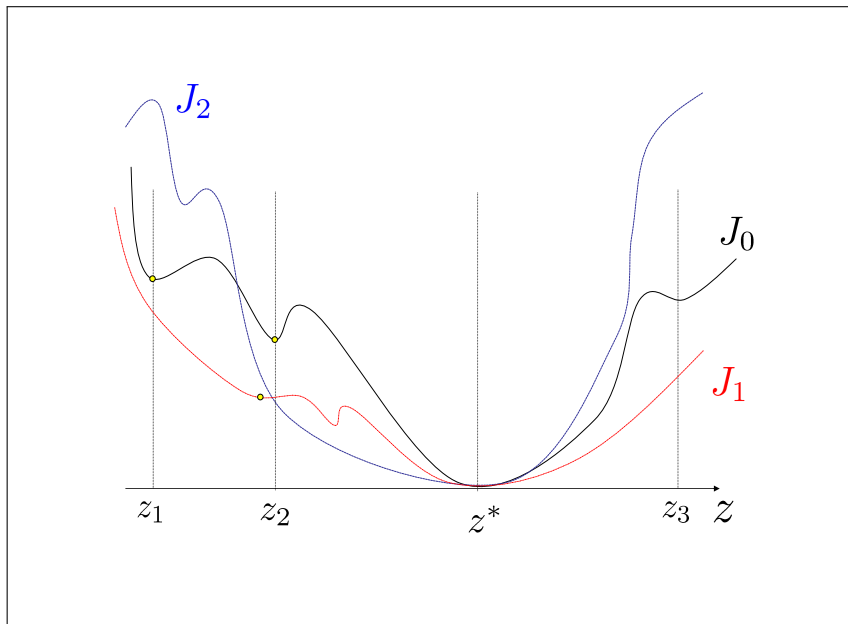


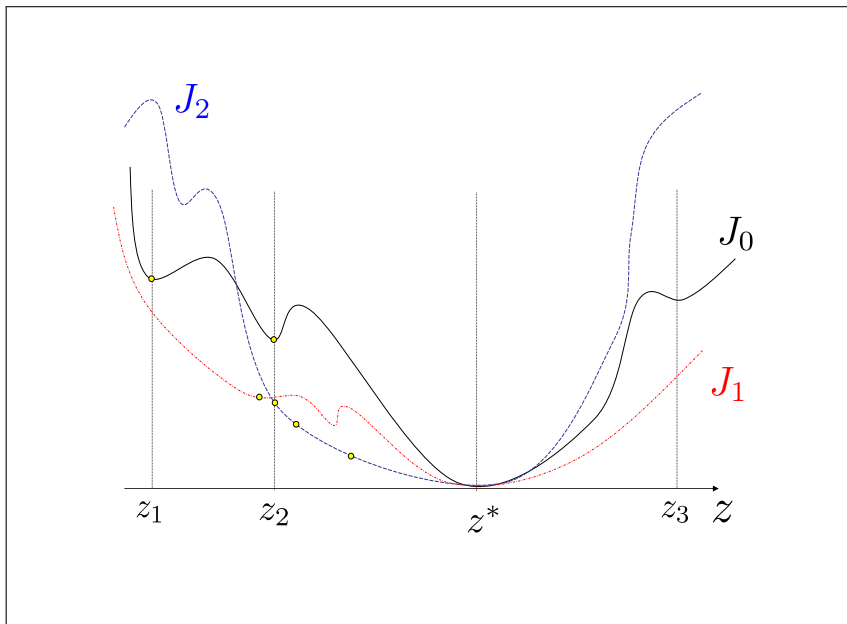


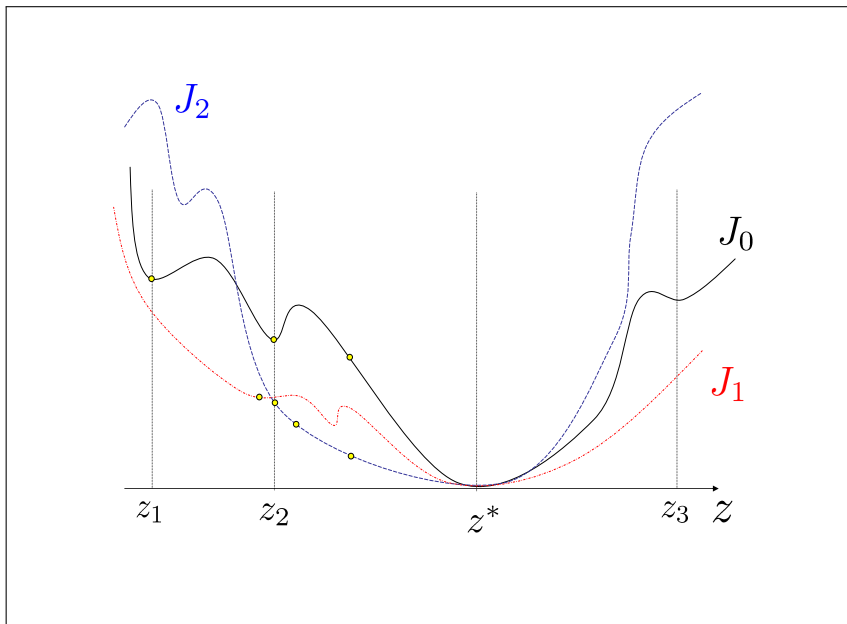


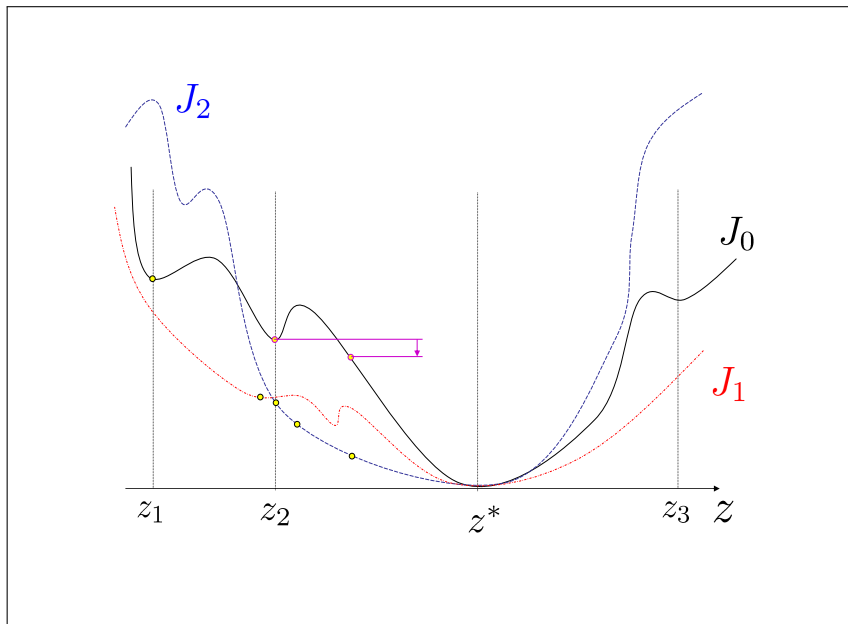


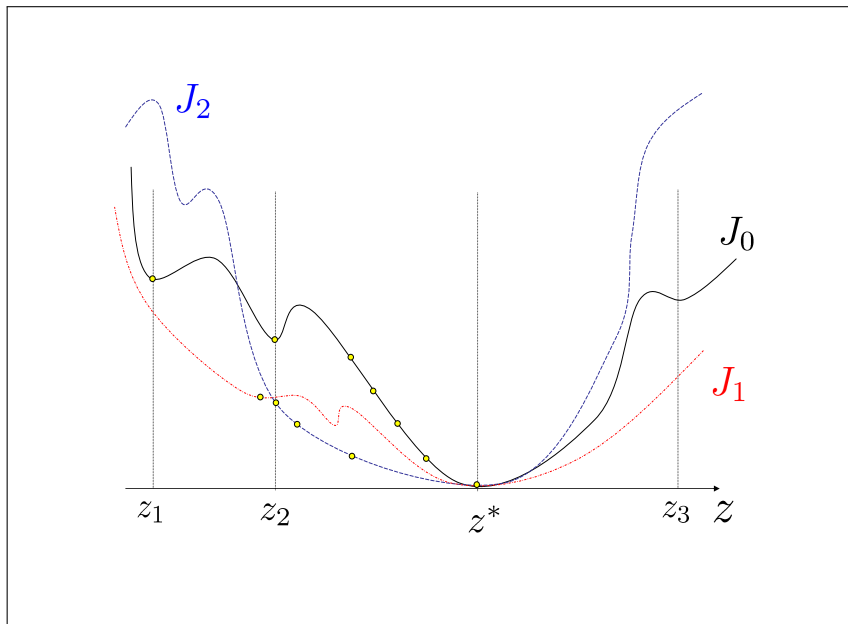


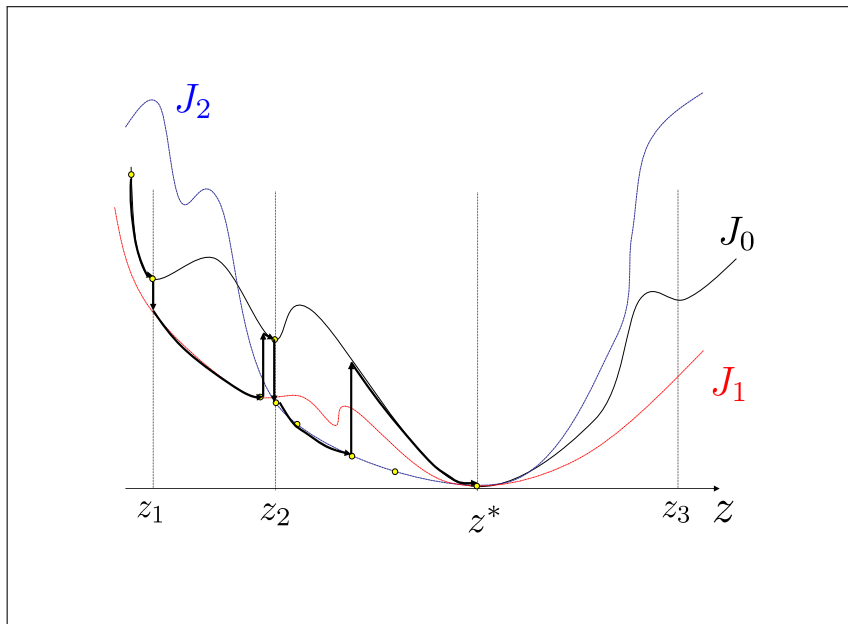


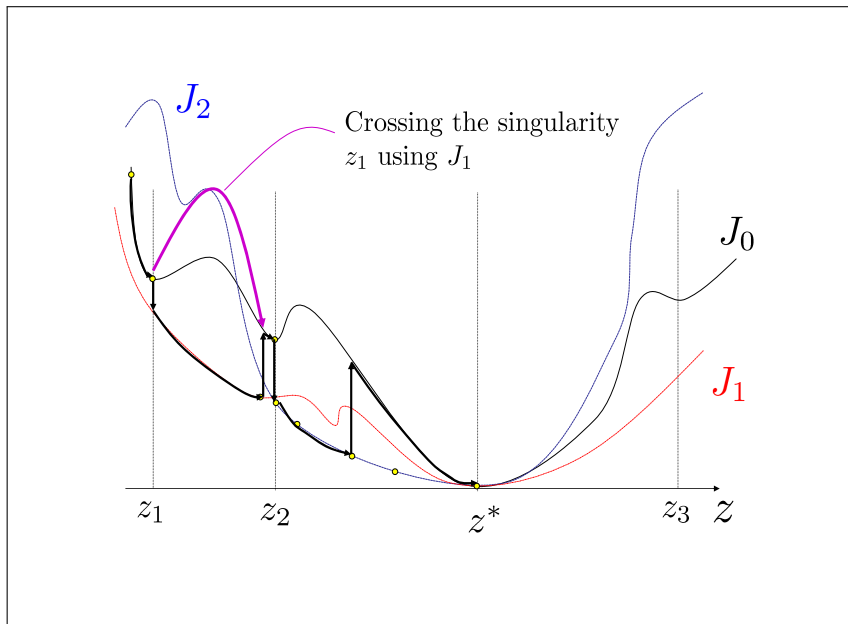


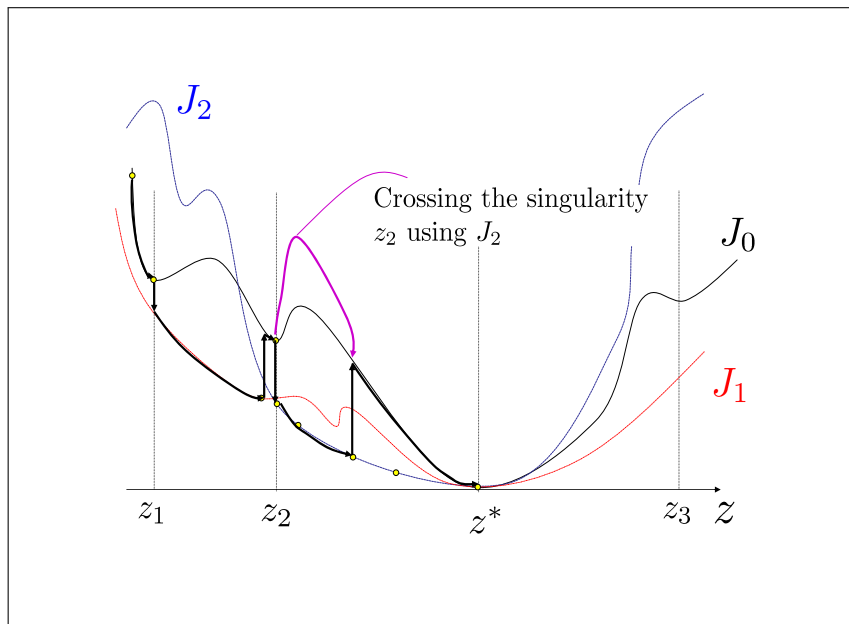


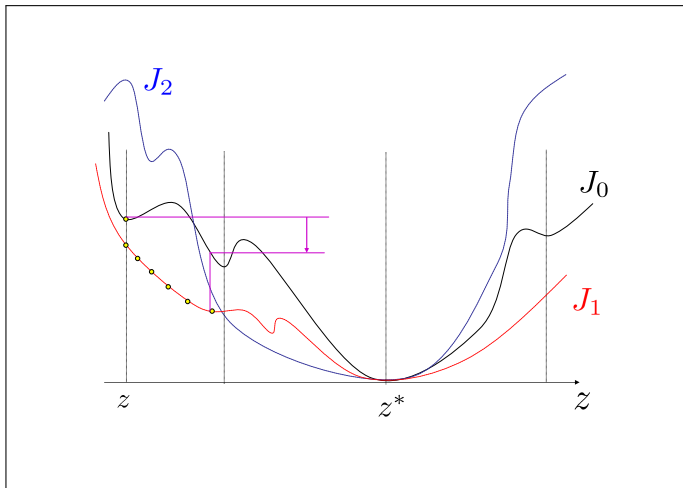


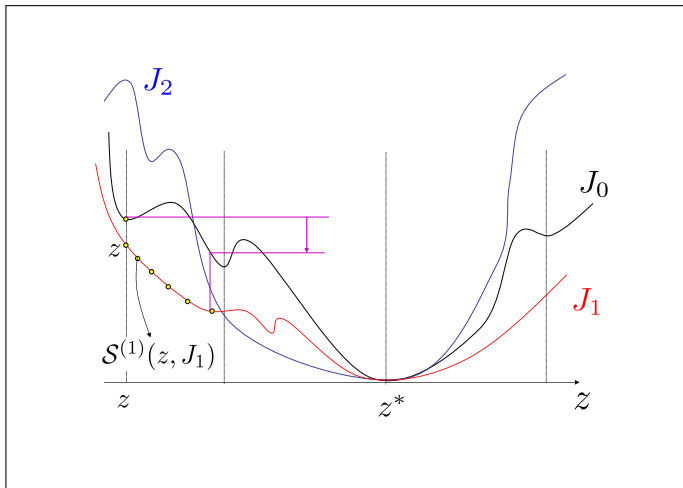


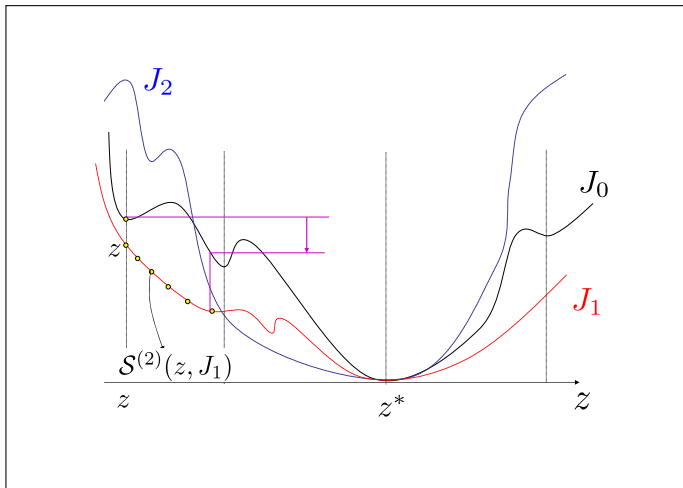


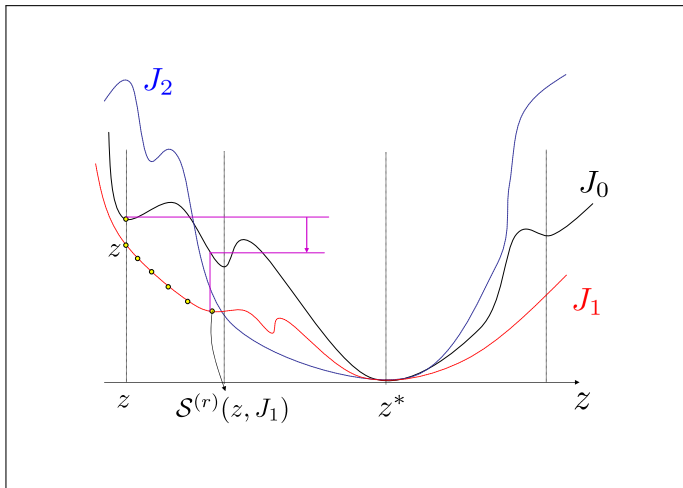


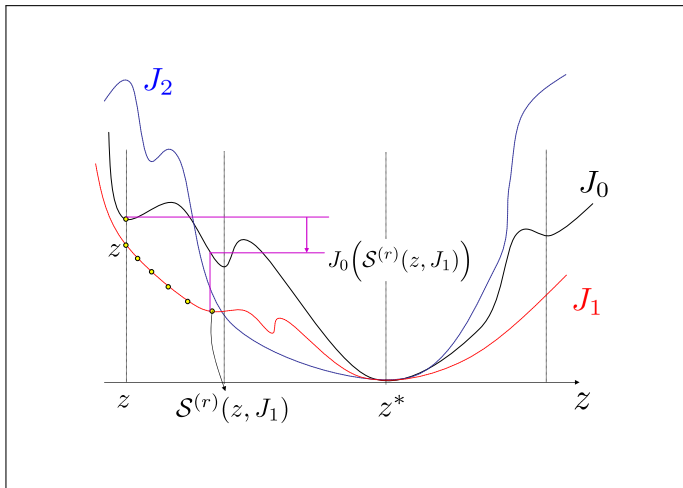


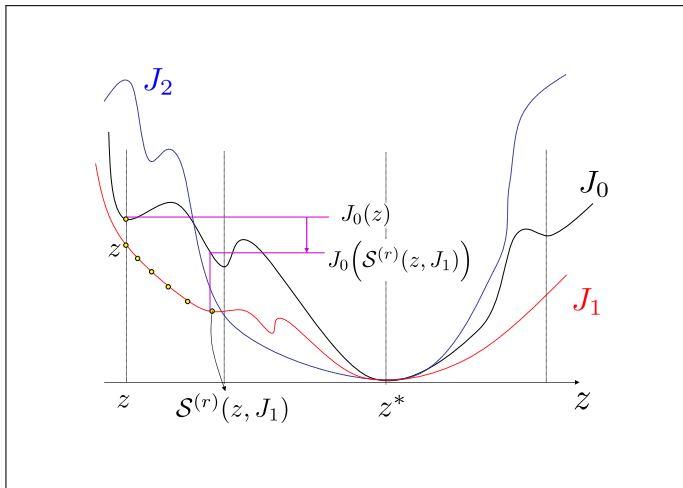


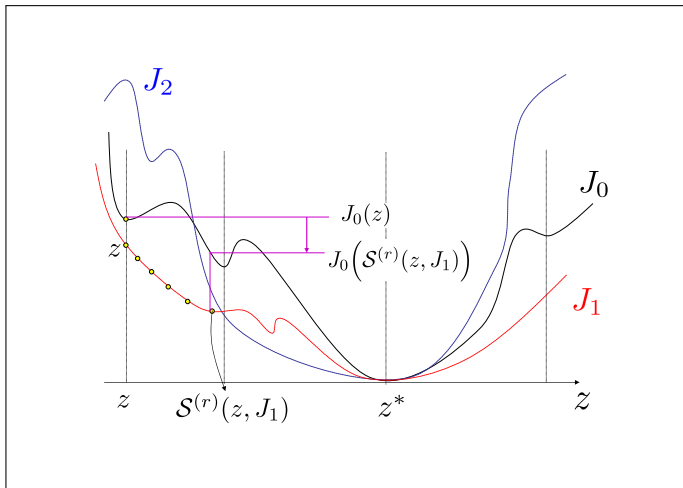




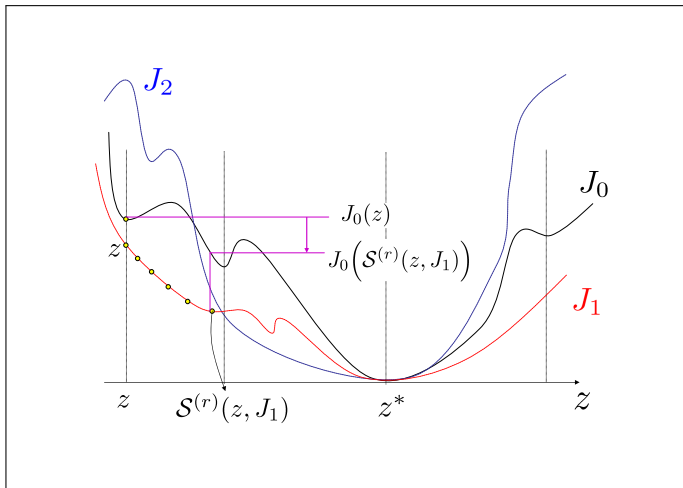








$$\Delta(z) := \min_{i \in \{0, \dots, N\}} \left[J_0(S^{(r^*)}(z, J_i)) - \gamma J_0(z) \right] \leq 0$$



$$(\exists r^*) (\forall z) \quad \Delta(z) := \min_{i \in \{0, \dots, N\}} \left[J_0(\mathcal{S}^{(r^*)}(z, J_i)) - \gamma J_0(z) \right] \leq 0$$

Next

- ▶ Bio-reactor State estimation example
- ▶ Parameter estimation example / comparison:
 - ▶ Multiple starting points
 - ▶ Redundancy-Based singularity avoidance

Example 1: Recombinant Escherichia Coli

$$\begin{aligned}\dot{X} &= \mu(S)X - k_d \exp\left(-\frac{k_p}{P}\right)X \\ \dot{S} &= -y_s \mu X - k_m S \\ \dot{P} &= y_p \mu(S) \frac{I}{I + k_I} X - k_d \exp\left(-\frac{k_p}{P}\right)P\end{aligned}$$

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

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

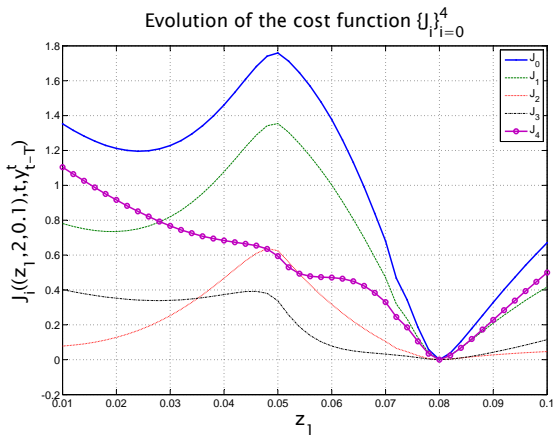


Escherichia coli under 15000
magnification factor

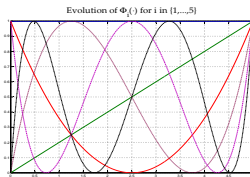
Output measurement:

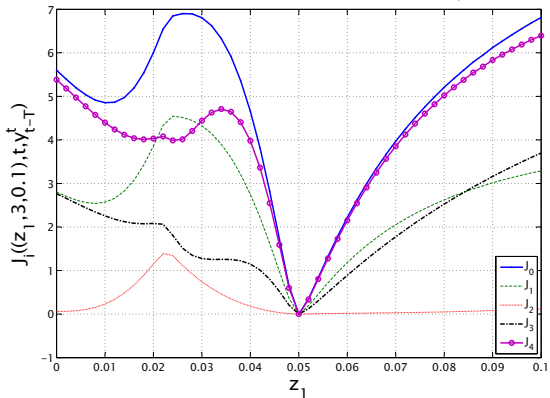
Light produced by the bioluminescence:

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

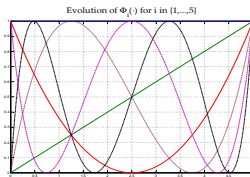


- ▶ $T = 10$
- ▶ $x(t - T) = (0.08, 2, 0.1)$
- ▶ $z_2 = x_2(t - T),$
 $z_3 = x_3(t - T)$



Evolution of the cost functions $\{J_i\}_{i=0}^4$ 

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



Algorithm 1: Multiple starting points

0. Initialization $p^{(0,0)}$ initial guess,
 $m \leftarrow 0$; $i \leftarrow 1$
while ($J_0(p^{(m,i)}) > 0$) **do**
 $failure \leftarrow true$
 while ($failure$) **do**
 $\xi^{(m,i)} \leftarrow \mathcal{S}^{r_1^*}(p^{(m,i)}, J_0)$
 $failure$
 $\leftarrow (J_0(\xi^{(m,i)}) > (1 - \varepsilon)J_0(p^{(m,i)}))$
 If $failure$ **then**
 $(m, i) \leftarrow (0, i + 1)$, $p^{(m,i)} \leftarrow random$
 Else $m \leftarrow m + 1$, $p^{(m,i)} \leftarrow \xi^{(m,i)}$
 End while
End while

Algorithm A₂: Redundancy based approach

0. Initialization $p^{(0)}$ initial guess,
 $m \leftarrow 0$
while ($J_0(p^{(m)}) > 0$) **do**
 $i \leftarrow 0$; $failure \leftarrow true$
 while ($failure$) **do**
 $\xi^{(m,i)} \leftarrow \mathcal{S}^{r^*}(p^{(m)}, J_i)$
 $failure$
 $\leftarrow (J_0(\xi^{(m,i)}) > (1 - \varepsilon)J_0(p^{(m)}))$
 If $failure$ **then**
 $i \leftarrow (i + 1) \cdot mod \cdot N$
 Else $m \leftarrow m + 1$, $p^{(m)} \leftarrow \xi^{(m,i)}$
 End while
End while

Example 2: Parameter estimation

$$\dot{x}_1 = -p_1 x_2$$

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

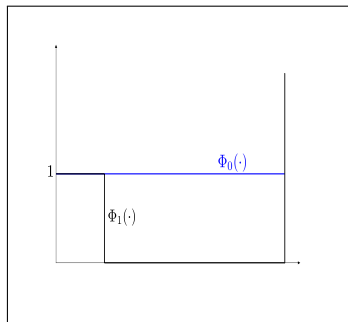
$$y = x_1 + x_2 + \nu$$

Example 2: Parameter estimation

$$\dot{x}_1 = -\rho_1 x_2$$

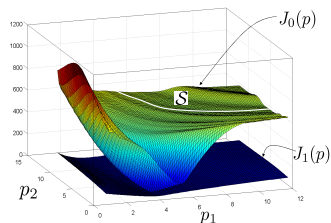
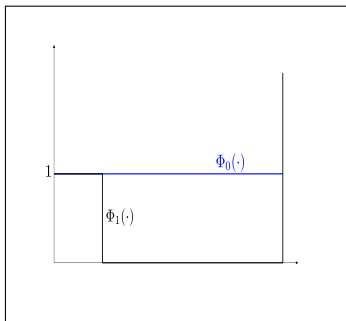
$$\dot{x}_2 = (1 + \rho_2)x_1 + (1 - x_1^2)x_2$$

$$y = x_1 + x_2 + \nu$$



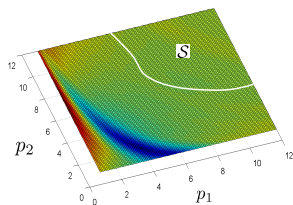
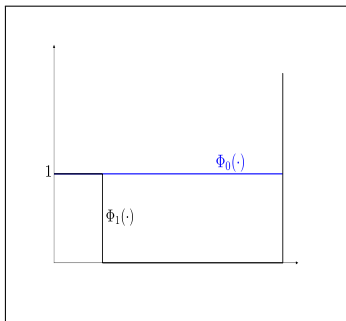
Example 2: Parameter estimation

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



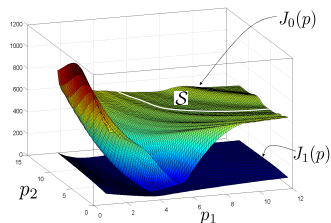
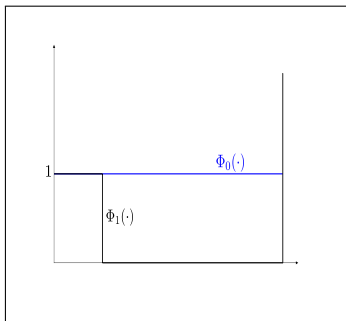
Example 2: Parameter estimation

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



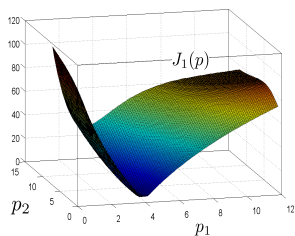
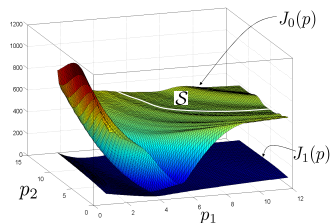
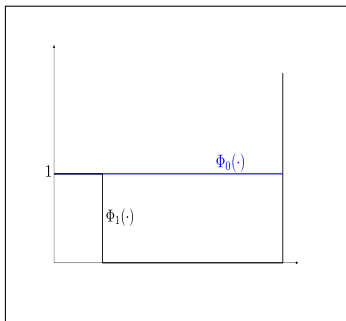
Example 2: Parameter estimation

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



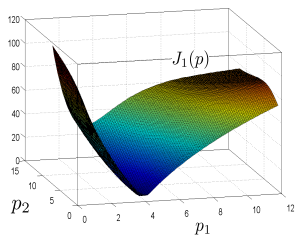
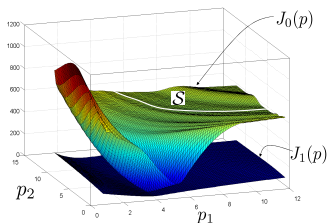
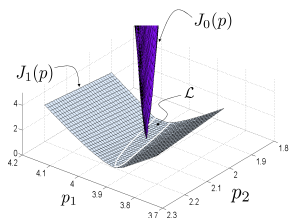
Example 2: Parameter estimation

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



Example 2: Parameter estimation

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

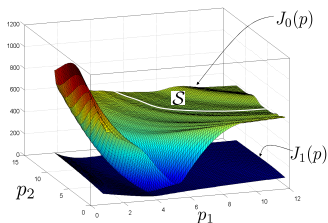
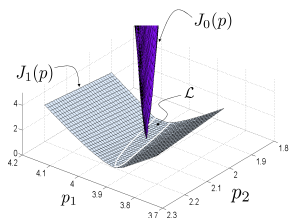


Example 2: Parameter estimation

$$\dot{x}_1 = -p_1 x_2$$

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

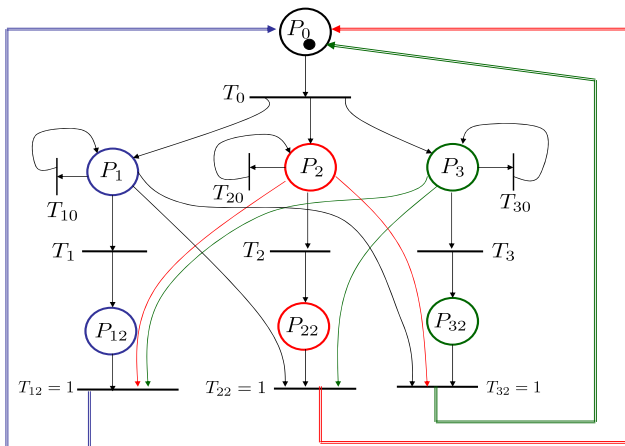
$$y = x_1 + x_2 + \nu$$



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

Conclusion

- ▶ Dynamic inversion leads to particular optimization problems
- ▶ Singularity crossing algorithm based on redundancy
- ▶ \neq multiple starting points
- ▶ Scheme that is solver-independent
- ▶ Additional *layer* to any existing *global* scheme



$$x^+ = f(x, u, p)$$

$$y = h(x, p) + w$$

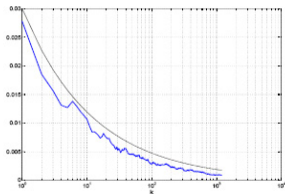
- $u_k = \mathcal{U}(k, q)$ with $q \in \mathbb{Q}$
- $y_k^0(q, p)$ prediction under $w \equiv 0$
- y_k^m effective measurement
- p^r true value of p
- $\epsilon_k(q, p^r, p) = y_k^0(q, p) - y_k^0(q, p^r)$

$$d^+(w, w^*) := \max\{0, |w| - w^*\}$$

(component wise)

Assumption on measurement errors w

$$\frac{\sum_{i=0}^{k^*-1} w(k+i)}{k^*} \in \phi_w(k^* - 1) \cdot [-\bar{w}, \bar{w}]$$



The l.h.s of the equation beside for for a MATLAB's $0.01 \times \text{RANDN}$ function (maximum over 1000 trials) and its approximated upper bound $\bar{w} \cdot \phi_w(k) = 0.03/(k^{0.4})$

