Parameter Identification: Selected Topics

F. Kappel

Institute for Mathematics and Scientific Computing

University of Graz

Parameter Identification

More complex models $\Longrightarrow$ Large number of parameters

Clinical applications $\Longrightarrow$ Necessity to "individualize" models

Problems:

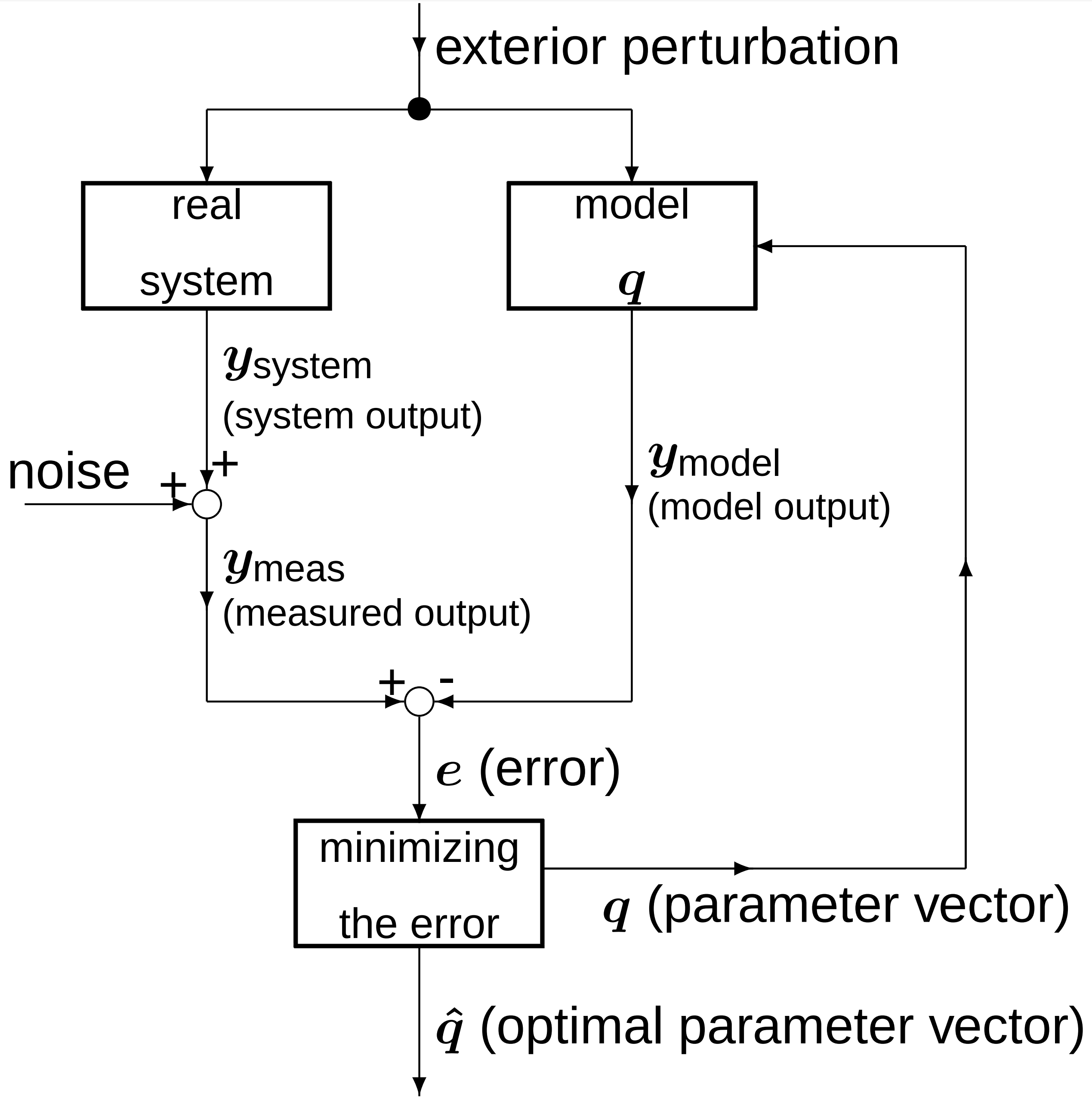
a) Identifiability

b) Unique identifiability and computation of parameter estimates (minimization problems in high-dimensional parameter spaces)

c) Comparatively few measurements, outliers (placement of measurements $\longleftarrow$ generalized sensitivities)

d) Finding the "important" parameters (sensitivity analysis)

Parameter Identification



Parameter Identification

Output-least-squares formulation:

$$y = y(t, \theta) \in^{k} \text{ (model output)}$$

$$\xi_i^{(\varkappa)} \sim y_\varkappa(t_i^{(\varkappa)}, \theta), \quad \kappa = 1, \ldots, k,\ i = 1, \ldots, M_\varkappa$$

(measurements)

Cost functional:

$$J(q) = \sum_{\varkappa=1}^{k} \sum_{i=1}^{M_\varkappa} \alpha_{\varkappa,i} \big(y_i^{(\varkappa)}(t_i^{(\varkappa)}, q) - \xi_i^{(\varkappa)}\big)^2,$$

Parameter Identification

Output-least-squares formulation:

Find $\hat{q} \in \mathcal{Q}_{\text{adm}}$ *such that*

(OLS) $$J(\hat{q}) = \min_{q \in \mathcal{Q}_{\text{adm}}} J(q),$$

$\mathcal{Q}_{\text{adm}}$ ... admissible domain for the parameters

In case of outliers replace J by

$$\tilde{J}(q) = \sum_{\varkappa=1}^{k} \sum_{i=1}^{M_\varkappa} \alpha_{\varkappa,i} \left| y_i^{(\varkappa)}(t_i^{(\varkappa)}, q) - \xi_i^{(\varkappa)} \right|.$$

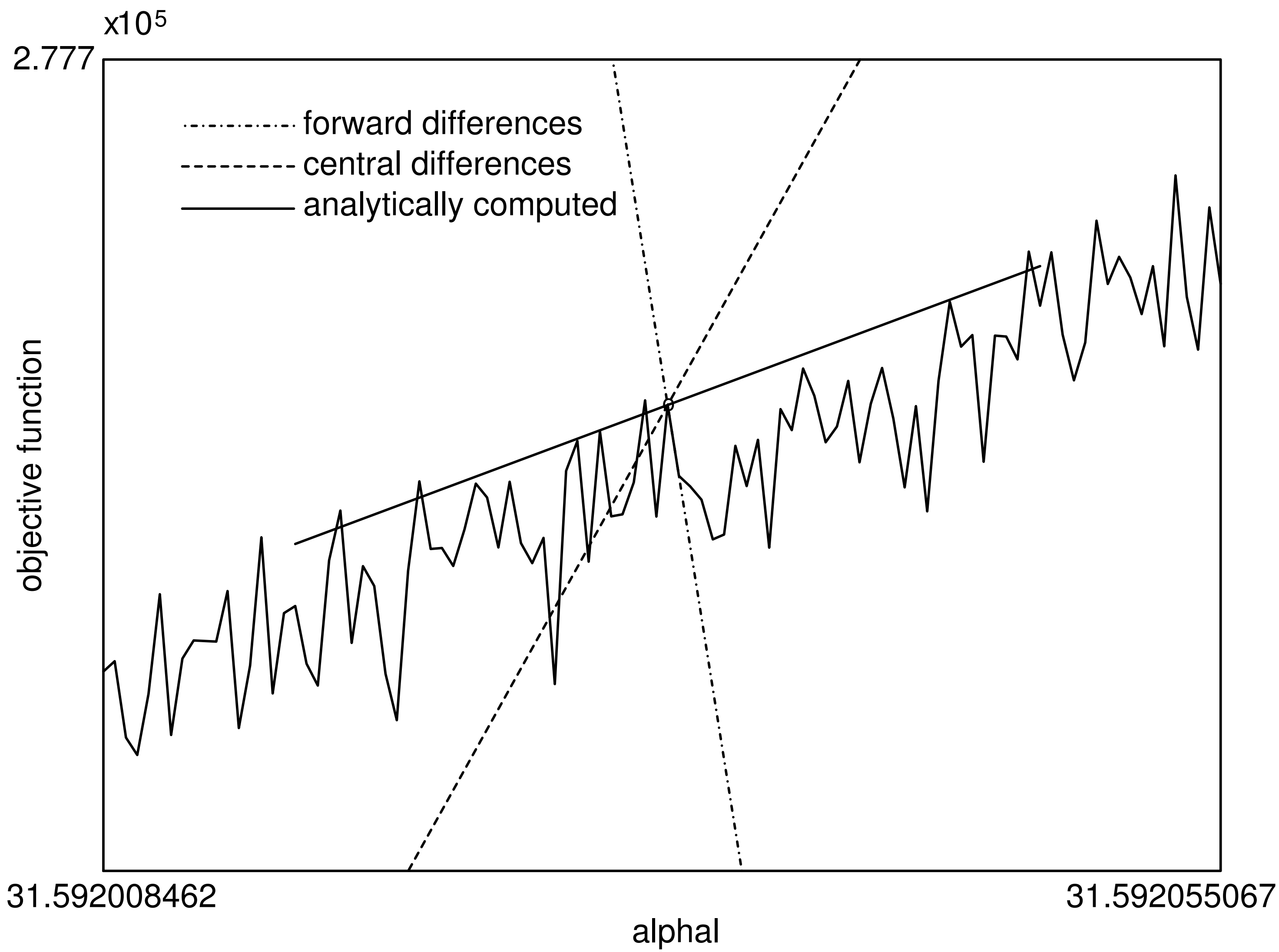
Parameter Identification

Gradients by differences vs. gradients analytically

a) Finite differences (forward, backward or central)
Advantage: No additional efforts for implementation ($p + 1$ forward integrations)
Disadvantage: Can give completely wrong approximations for gradients

b) Analytically computed gradients
Advantage: Gives good approximations for gradients
Disadvantage: Needs computation of the sensitivity equations ($\Longrightarrow$ automated differentiation)

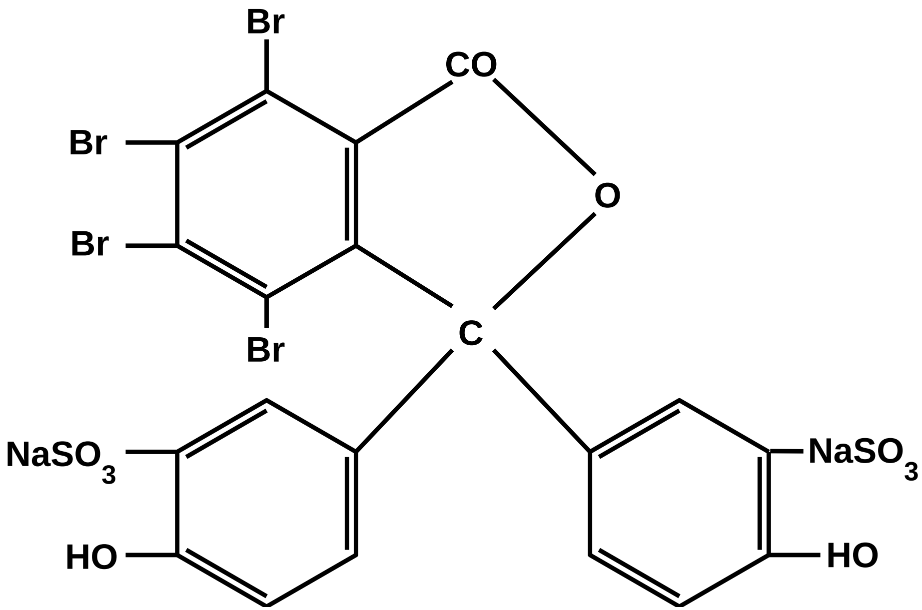
Parameter Identification



Graph of $\alpha_\ell \longrightarrow J(q^0)$ drawn with the step size used for difference approximations of derivatives.

Bromsulphalein test

Test for hepatic functions; the decay rate of bromsulphalein in the liver characterizes certain hepatic functions.

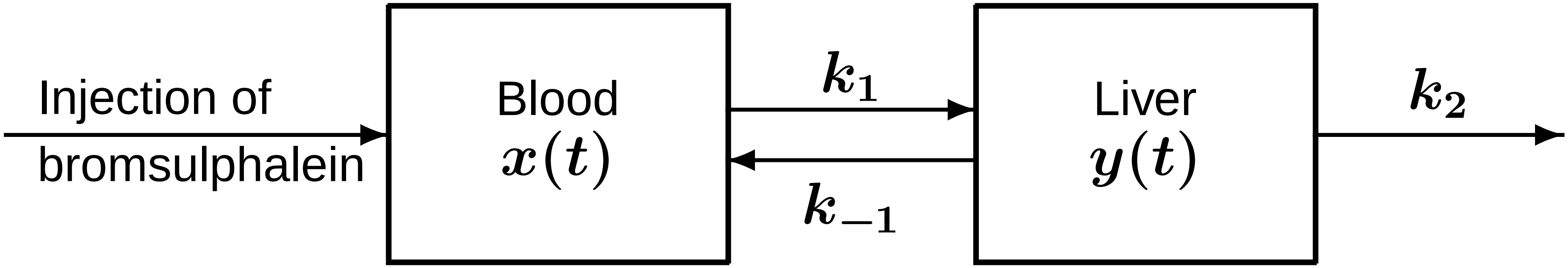


The bromsulphalein retention test:

a) Intravenous application of a dosage d of bromsulphalein at time $t = 0$.

b) Take blood samples approximately every 10 min. and determine the concentration of bromsulphalein in blood.

Compartment model:



$x(t)$ ... mass of BS in blood at time t

$y(t)$ ... mass of BS in liver at time t

k_1, k_{-1}, k_2 ... transition rates

Mass-balance:

$$\dot{x}(t) = -k_1 x(t) + k_{-1} y(t), \quad (1)$$

$$\dot{y}(t) = k_1 x(t) - (k_{-1} + k_2) y(t) \quad (2)$$

Initial conditions: $x(0) = d$, $y(0) = 0$.

$\Longrightarrow$

$$x(t) = c_1 \frac{\lambda_1 + k_{-1} + k_2}{k_1} e^{\lambda_1 t} + c_2 \frac{\lambda_2 + k_{-1} + k_2}{k_1} e^{\lambda_2 t},$$
$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

$$\lambda_{1,2} = \frac{1}{2}\left(k_1 + k_{-1} + k_2 \pm \sqrt{(k_1 + k_{-1} + k_2)^2 - 4k_1 k_2} \right)$$

$$c_1 = \frac{k_1 d}{\lambda_1 - \lambda_2}, \quad c_2 = -\frac{k_1 d}{\lambda_1 - \lambda_2}.$$

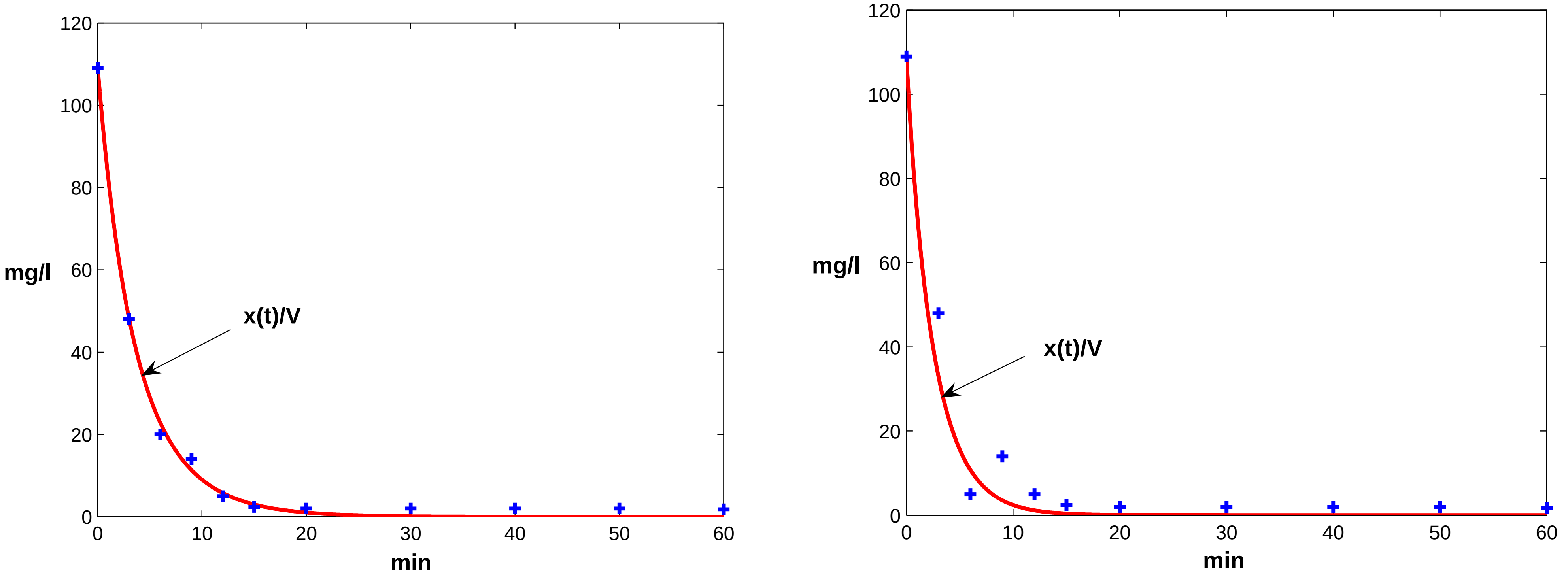
$\Longrightarrow x(t) = x(t; k_1, k_{-1}, k_2), \qquad y(t) = y(t; k_1, k_{-1}, k_2)$

Task: Identification of k_2 (and k_1, k_{-1})

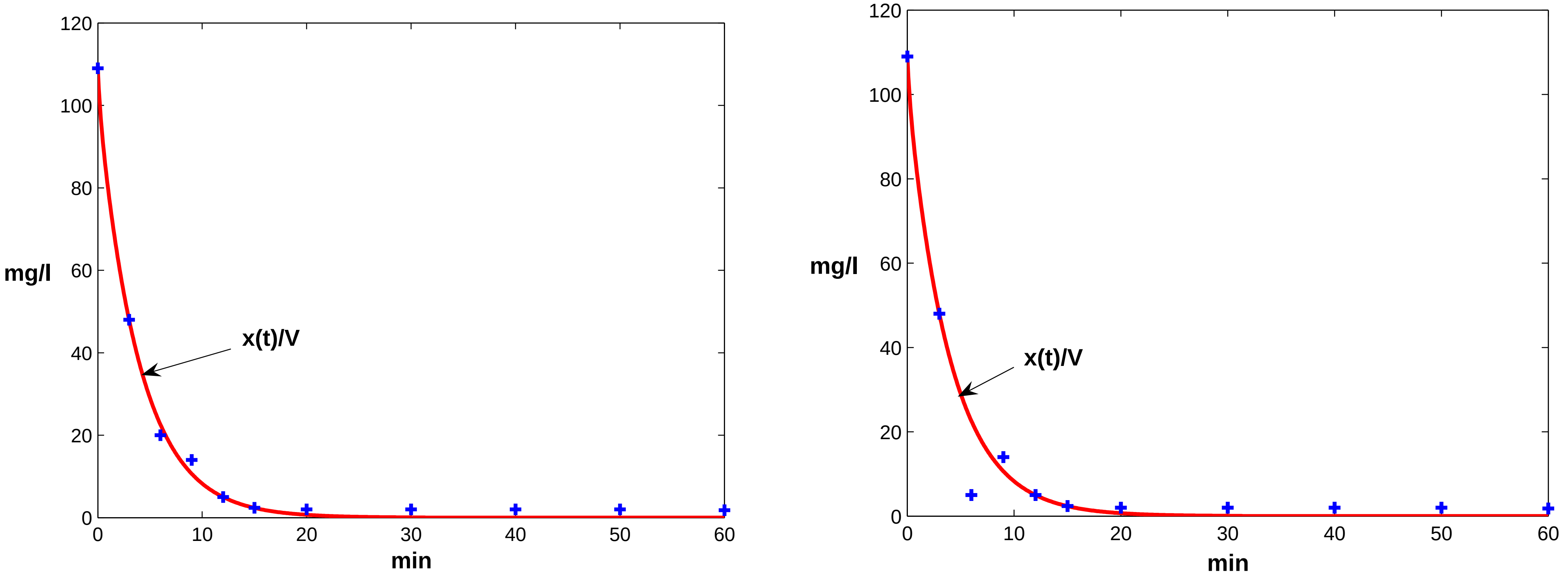
Data: ξ_i, $i = 1, \ldots, N$,. . . concentration of bromsulphalein in blood at times $t_1 < \cdots < t_N$.

$$\xi_i = \frac{x(t_i)}{V},$$

V . . . blood volume



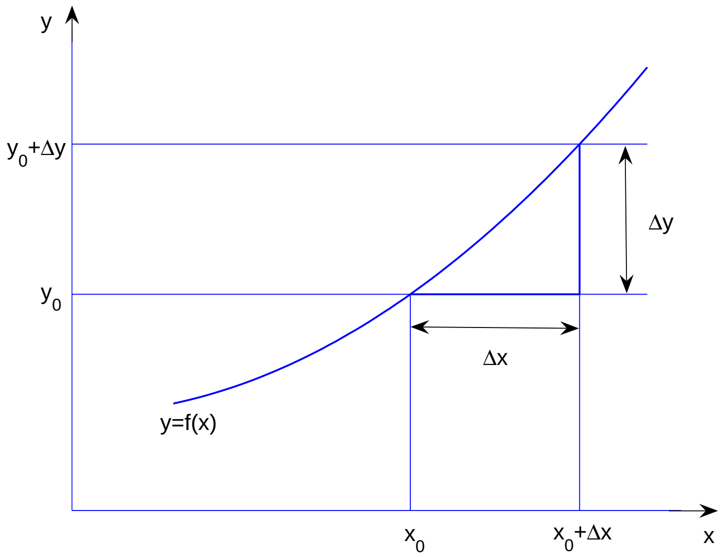
Bromsulphalein concentration (solid line) versus data (crosses), cost functional J (without outlier left panel, with outlier right panel).



Bromsulphalein concentration (solid line) versus data (crosses), cost functional $\tilde{J}$ (without outlier left panel, with outlier right panel).

Costs	Outlier	$\hat{V}$	$\hat{k}_1$	$\hat{k}_{-1}$	$\hat{k}_2$
J	no	5.5373	0.2706	0.0393	0.3182
$\tilde{J}$	no	5.5046	0.5021	1.8297	2.0852
J	yes	5.4897	0.4427	0.1273	1.0444
$\tilde{J}$	yes	5.5046	0.4944	1.7460	2.0560

Classical sensitivities



Relative errors:

$\Delta x / x_0$ and $\Delta y / y_0$

Sensitivity of y with respect to x at x_0:

$$\sigma_{y,x}(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y / y_0}{\Delta x / x_0} = \frac{x_0}{y_0} y'(x_0).$$

Classical sensitivities

Sensitivity:
Measure for the dependence of outputs or states on parameters.

Note that

$\boldsymbol{\sigma_{y,x}(x_0)}$ is dimensionless and thus independent of units.

Generalized sensitivities

Generalized sensitivity:
Sensitivity of the parameter estimates with respect to variations in the measurements.

Generalized sensitivities

Single output system: $y(t) = f(t, \theta)$, $0 \le t \le T$,

$\theta = (\theta_1, \dots, \theta_p)^{\mathsf{T}}$ . . . model parameters

ξ_k . . . measurements for $y(t_k), 0 \le t_1 < \cdots < t_M \le T$

$$\xi_k = z(t_k) + e_k, \quad k = 1, \dots, M,$$

$z(t)$ . . . 'true' output of the system

e_k . . . measurement noise for ξ_k

Generalized sensitivities

Assumptions on e_k:

(i) e_k has zero mean, $k = 1, \ldots, M$.

(ii) The e_k's are identically distributed.

(iii) The variance σ_k^2 of e_k is not dependent on θ.

Generalized sensitivities

Basic assumption:

$$\exists \theta_0 : \quad z(t_k) = f(t_k, \theta_0), \; k = 1, \ldots, M.$$

Output least squares formulation ($\xi = (\xi_1, \ldots, \xi_M)$):

$$\hat{\theta}_0 = \mathrm{argmin}_\theta \, J(\xi, \theta),$$

$$J(\xi, \theta) = \sum_{k=1}^{M} \frac{1}{\sigma_k^2} \big(\xi_k - f(t_k, \theta)\big)^2$$

Generalized sensitivities

Assumptions:

a) Unique local identifiability at $\boldsymbol{\theta_0}$.

$$\forall \boldsymbol{\theta_1} \in \boldsymbol{\mathcal{U}}(\boldsymbol{\theta_0}) \exists ! \hat{\boldsymbol{\theta}}_1 : \hat{\boldsymbol{\theta}}_1 = \mathbf{argmin}_{\boldsymbol{\theta}} \boldsymbol{J}(\ldots \boldsymbol{f}(\boldsymbol{t_k}, \boldsymbol{\theta_1}) + \boldsymbol{e_k} \ldots, \boldsymbol{\theta})$$

Note: We consider $\boldsymbol{\xi} \sim \boldsymbol{\xi}(\boldsymbol{\theta_0})$, because $\boldsymbol{\xi} = \boldsymbol{F}(\boldsymbol{\theta_0}) + \boldsymbol{e}$

b) Estimates are unbiased, i.e., $\boldsymbol{E}(\hat{\boldsymbol{\theta}}_0) = \boldsymbol{\theta_0}$.

$\Longrightarrow$

$$\boldsymbol{\nabla_{\theta} J}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}}_0) = \mathbf{0}, \quad \boldsymbol{\nabla_{\theta\,\theta} J}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}}_0) > \mathbf{0}.$$

Generalized sensitivities

Quantity of interest: $$\frac{\partial\hat{\theta}}{\partial\theta}$$

$$\frac{\partial\hat{\theta}(\theta)}{\partial\theta} = -\left(\nabla_{\theta\theta} J(\xi(\theta),\hat{\theta}(\theta))\right)^{-1} \nabla_{\xi\theta} J(\xi(\theta),\hat{\theta}(\theta)) \frac{\partial F(\theta)}{\partial\theta}$$

$$F(\theta) = \big(f(t_1,\theta),\ldots,f(t_M,\theta)\big)$$

Generalized sensitivities

$\Longrightarrow$ (taking expected values)

$$\frac{\partial\hat{\theta}(\theta)}{\partial\theta} = \sum_{k=1}^{M} \frac{1}{\sigma_k^2}\left(\mathcal{M}^{-1}\left(\nabla_\theta f(t_k,\theta)\right)^{\mathsf{T}}\right)\nabla_\theta f(t_k,\theta)$$
$$= \mathcal{M}^{-1}\mathcal{M} \equiv I.$$

where

$$\mathcal{M} = \sum_{j=1}^{M} \frac{1}{\sigma_j^2}\left(\nabla_\theta f(t_j,\theta)\right)^{\mathsf{T}}\nabla_\theta f(t_j,\theta)$$

Estimate $\hat{\theta}_j$ for $(\theta_0)_j$ is independent from the estimate $\hat{\theta}_k$ for $(\theta_0)_k$, $k \neq j$

Generalized sensitivities

Modification: All deviations $\xi_k - f(t_k, \theta)$, $k = 1, \ldots, M$, still enter the cost functional, but we only have access to the measurements $\xi_1, \ldots, \xi_{k_0}$, $1 \leq k_0 \leq M$.

$\Longrightarrow$

$$\frac{\partial \hat{\theta}}{\partial \theta} = \mathcal{M}^{-1} \sum_{k=1}^{k_0} \frac{1}{\sigma_k^2} \nabla_\theta f(t_k, \theta)\big)^{\mathsf{T}} \nabla_\theta f(t_k, \theta).$$

Generalized sensitivities

Generalized sensitivity for parameter θ_i:

$$g_i(t_{k_0}) = \sum_{k=1}^{k_0} \frac{1}{\sigma_k^2} \Big(\mathcal{M}^{-1} \big(\nabla_\theta f(t_k, \theta)\big)^{\mathsf{T}} \Big)_i \big(\nabla_\theta f(t_k, \theta)\big)_i .$$

Note that

$$\text{rank}\,\big(\nabla_\theta f(t_j, \theta)\big)^{\mathsf{T}} \nabla_\theta f(t_j, \theta) = 1, \quad \text{but} \quad \mathcal{M} \in \mathbb{R}^{p \times p}$$

Incremental gen. sensitivities

Incremental generalized sensitivity functions:

$$g_{\text{inc},i}(t_{k_0}) = g_i(t_{k_0}) - g_i(t_{k_0-1}), \quad k_0 = 1, \ldots, M.$$

Explicit representation:

$$g_{\text{inc},i}(t_{k_0}) = \frac{1}{\sigma_{k_0}^2}\Big(\mathcal{M}^{-1}(\nabla_\theta f(t_{k_0},\theta))^{\mathsf{T}}\Big)_i (\nabla_\theta f(t_{k_0},\theta))_i,$$
$$k_0 = 1, \ldots, M,\ i = 1, \ldots, p.$$

Interpretation

Fisher information matrix:

$$\mathcal{J}(\theta) = \sum_{k=1}^{M} \frac{1}{\sigma_k^2} \left(\nabla_\theta f(t_k, \theta) \right)^{\mathsf{T}} \nabla_\theta f(t_k, \theta) \eta_k$$

η_k ...weight for measurement ξ_k at time t_k ($= 1$ in our case, i.e., $\mathcal{J}(\theta) = \mathcal{M}$)

Information index: $\ln\left(\det \mathcal{J}(\theta)\right)$

$\frac{\partial}{\partial \eta_k} \ln(\det \mathcal{J}(\theta))$ ...information on the parameters θ provided by the measurement ξ_k.

Interpretation

$$\boxed{\frac{\partial}{\partial \eta_k} \ln(\det \mathcal{J}(\theta)) = \sum_{i=1}^{p} g_{i,\text{inc}}(t_k) \geq 0}$$

$\Longrightarrow \qquad k \to \sum_{i=1} g_i(t_k)$ is increasing

Identification procedure is efficient, i.e.,

$\text{Cov}\,\hat{\theta} = \mathcal{J}(\hat{\theta})^{-1} = \mathcal{M}^{-1}$ (Cramer-Rao bound)

Interpretation

a) Information provided by the measurements for given parameters uncorrelated$\Longrightarrow$generalized sensitivity functions for these parameters monotonically increasing

b) In case of a) measurements in that time interval, where the generalized sensitivity function of a parameter has most of its increase from 0 to 1, are the measurements which carry most of the information on that parameter.

c) Is the information on given parameters rather strongly correlated$\Longrightarrow$oscillations of the generalized sensitivity functions

CVS

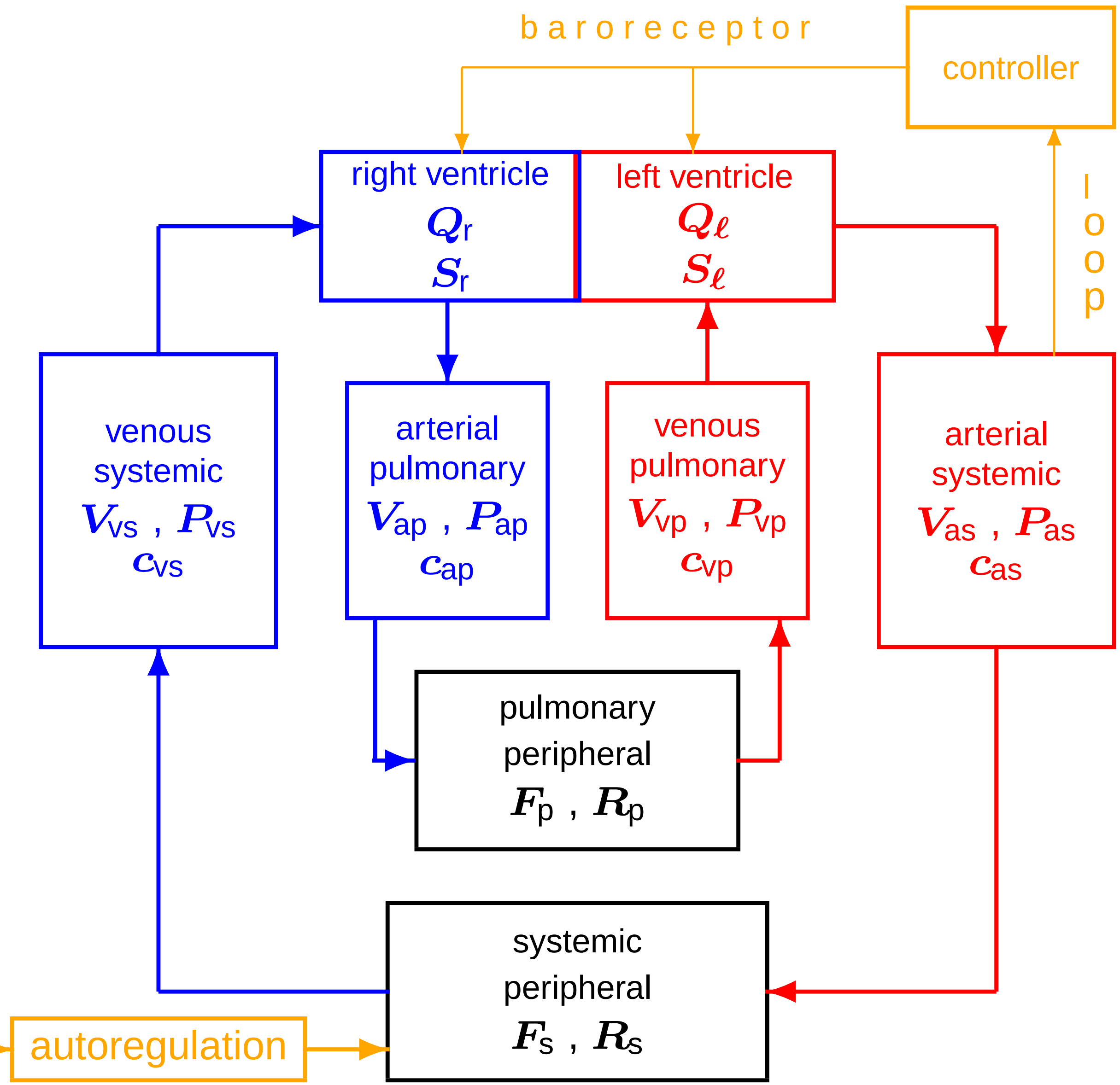
$$F_{\mathrm{s}} = \frac{P_{\mathrm{as}} - P_{\mathrm{vs}}}{R_{\mathrm{s}}}$$

$$F_{\mathrm{p}} = \frac{P_{\mathrm{ap}} - P_{\mathrm{vp}}}{R_{\mathrm{p}}}$$

$$Q_{\ell} = H V_{\mathrm{str},\ell}$$

$$Q_{\mathrm{r}} = H V_{\mathrm{str,r}}$$

$$V = cP$$



CVS

$$c_{\mathsf{as}}\dot{P}_{\mathsf{as}} = Q_\ell - F_{\mathsf{s}},$$

$$c_{\mathsf{vs}}\dot{P}_{\mathsf{vs}} = F_{\mathsf{s}} - Q_{\mathsf{r}},$$

$$c_{\mathsf{ap}}\dot{P}_{\mathsf{ap}} = Q_{\mathsf{r}} - F_{\mathsf{p}},$$

$$\boxed{\dot{x}(t) = \mathcal{F}(x(t), q) + Bu(t)}$$

$$c_{\mathsf{vp}}\dot{P}_{\mathsf{vp}} = F_{\mathsf{p}} - Q_\ell,$$

$$\ddot{S}_\ell + \gamma_\ell \dot{S}_\ell + \alpha_\ell S_\ell = \beta_\ell H,$$

$$\ddot{S}_{\mathsf{r}} r + \gamma_{\mathsf{r}} \dot{S}_{\mathsf{r}} + \alpha_{\mathsf{r}} S_{\mathsf{r}} = \beta_{\mathsf{r}} H,$$

$$\dot{R}_{\mathsf{s}} = \frac{1}{K}\Big(A_{\mathsf{pesk}}\big(F_{\mathsf{s}} C_{\mathsf{a,O_2}} - M\big) - (P_{\mathsf{as}} - P_{\mathsf{vs}})\Big),$$

$$\dot{H} = u(t)$$

CVS

Choose $u(t)$ such that

$$J(u(\cdot), x^{\text{rest}}) = \int_0^\infty (q_{\text{as}}^2 (P_{\text{as}}(t) - P_{\text{as}}^{\text{exer}})^2 + u(t)^2)\, dt \rightarrow \min$$

$$x(0) = x^{\text{rest}}$$

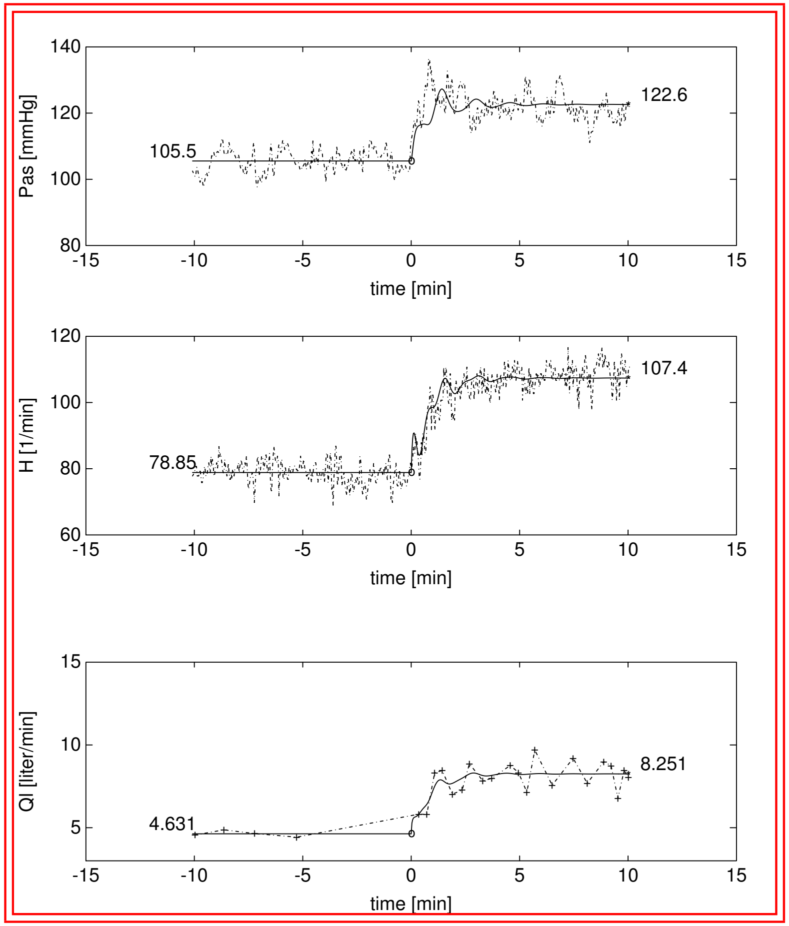
$\Longrightarrow$

$$u(t) = K(x(t) - x^{\text{exer}}), \quad K = -B^{\top} X,$$

$$XA + A^{\top} X - XBB^{\top} X + C^{\top} C = 0,$$

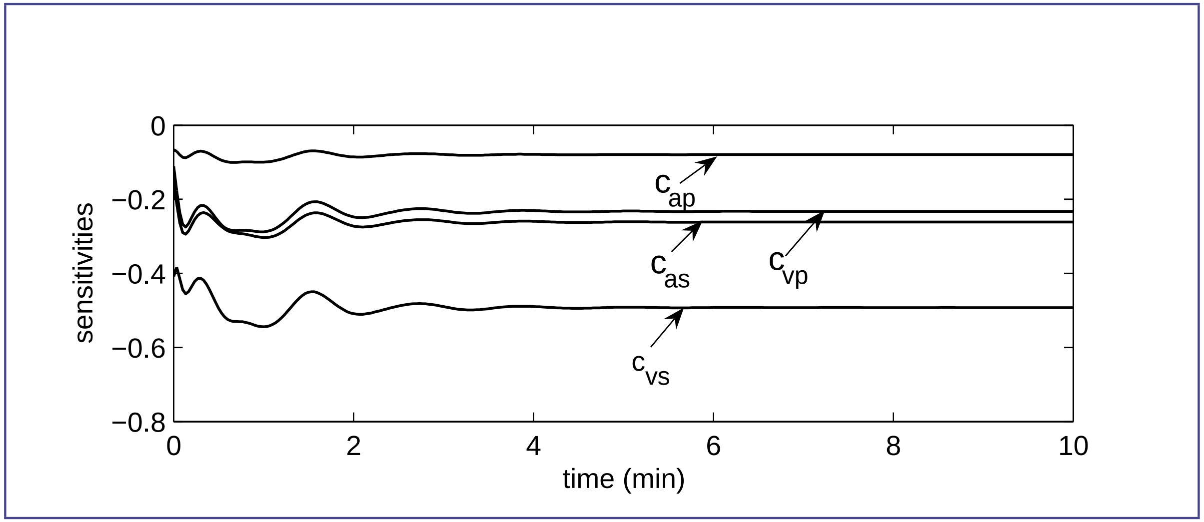
where $A = (\partial \mathcal{F} / \partial)(x^{\text{exer}}, q)$, $B = \text{col}(0, \ldots, 1)$,
$C = (q_{\text{as}}, 0, \ldots, 0)$.

CVS

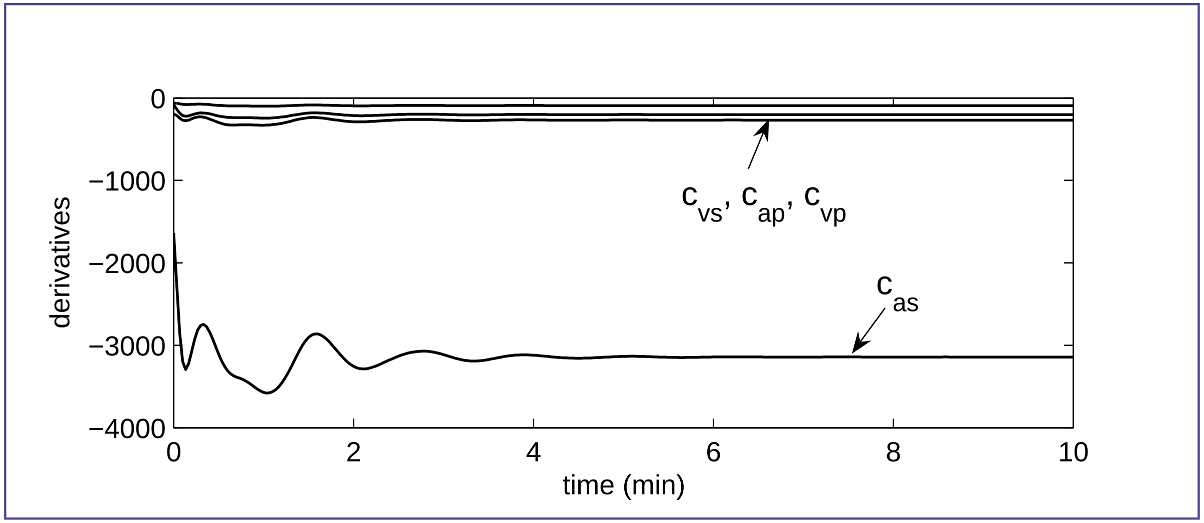


Measurements (dash-dots) for P_{as}, H, Q_ℓ and model output (solid).

CVS: Classical sensitivities

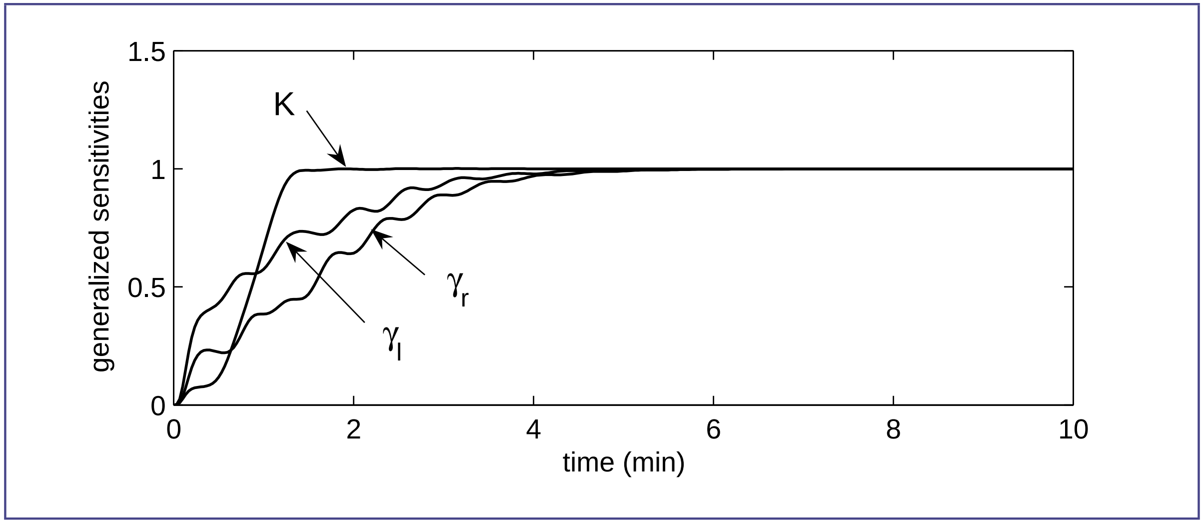


Sensitivities of P_{as}

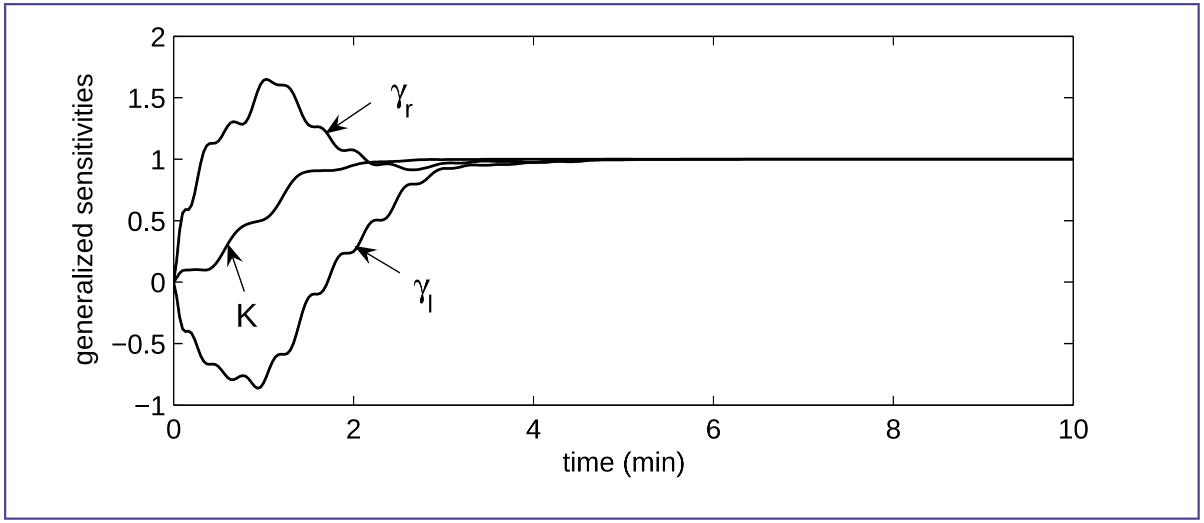


Derivatives of P_{as}

CVS: Generalized sensitivities

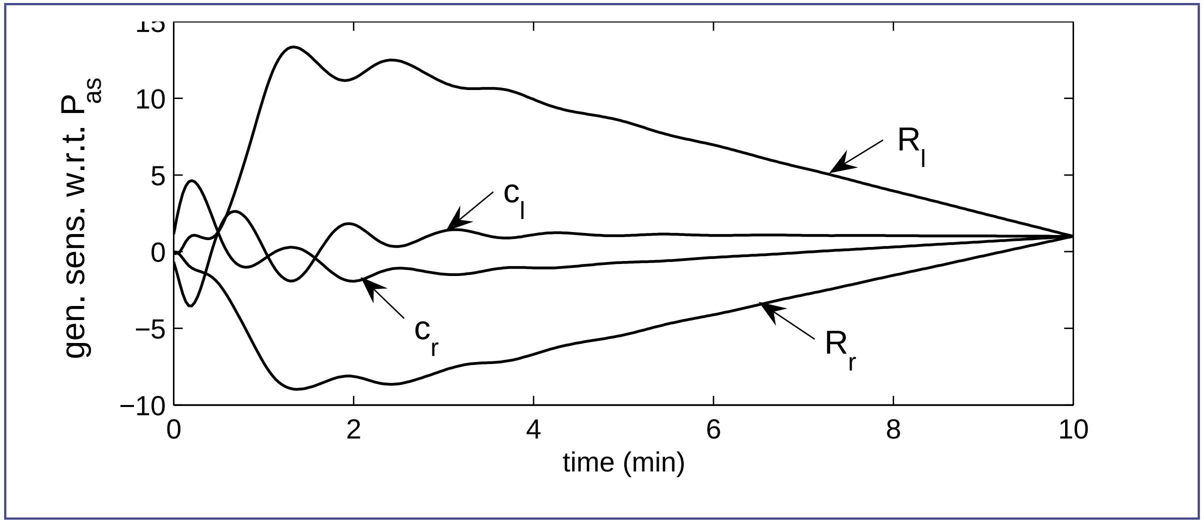


P_{as}

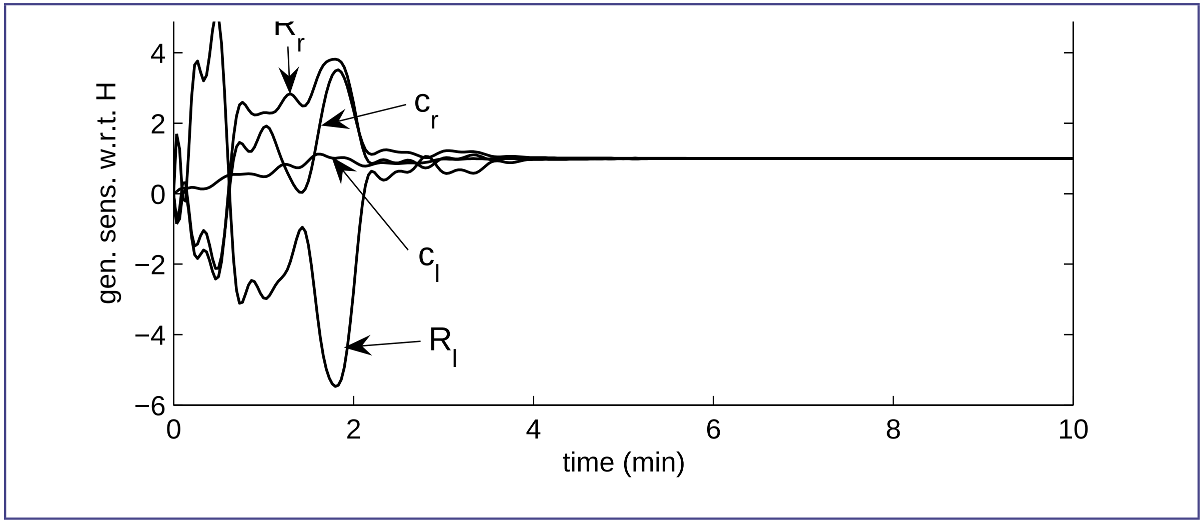


H

CVS: Generalized sensitivities

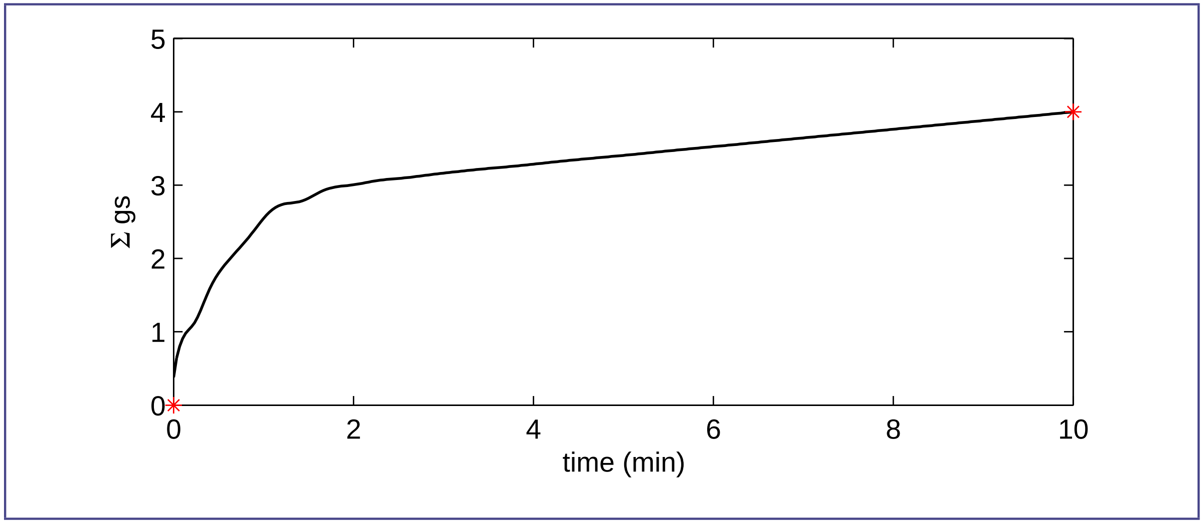


P_{as}



H

CVS: Generalized sensitivities



Sum of generalized sensitivities for P_{as}