

The Sensitivity Analysis and Parameter Estimation of Mathematical Models Described by Differential Equations

Hossein ZivariPiran

hzp@cs.toronto.edu

Department of Computer Science
University of Toronto

(part of my PhD thesis under the supervision of professor Wayne Enright)

Outline

- Modeling with Differential Equations (IVPs, DDEs)
- Sensitivity Analysis of Models Described by DEs
- Parameter Estimation of Models Described by DEs

Modeling with DE - Formulation

- An Initial Value Problem (IVP) for Ordinary Differential Equations (ODEs)

$$y'(t) = f(t, y(t))$$

$$y(t_0) = y_0$$

- Retarded Delay Differential Equations (RDDEs)

$$y'(t) = f(t, y(t), y(t - \sigma_1), \dots, y(t - \sigma_\nu)) \text{ for } t_0 \leq t \leq t_F$$

$$y(t) = \phi(t), \text{ for } t \leq t_0$$

$\sigma_i = \sigma_i(t, y(t)) \geq 0$ *delay* (constant / time dependent / state dependent)

$\phi(t)$ *history* function (constant / time dependent)

- Neutral Delay Differential Equations (NDDEs)

$$y'(t) = f(t, y(t), y(t - \sigma_1), \dots, y(t - \sigma_\nu),$$

$$y'(t - \sigma_{\nu+1}), \dots, y'(t - \sigma_{\nu+\omega})) \text{ for } t_0 \leq t \leq t_F$$

$$y(t) = \phi(t), \quad y'(t) = \phi'(t), \text{ for } t \leq t_0,$$

Modeling with DE - Some Areas of Application

Area	Example
Ecology	predator-prey
Epidemiology	spread of infections
Immunology	immune response models
HIV infection	
Physiology	human respiration system
Neural Networks	
Cell Kinetics	
Chemical Kinetics	The Oregonator
Physics	Ring Cavity Lasers , two-body problem of electrodynamics

Modeling with DE - An Example

- A neutral delay logistic Gause-type predator-prey system [Kuang 1991]

$$y_1'(t) = y_1(t)(1 - y_1(t - \tau) - \rho y_1'(t - \tau)) - \frac{y_2(t)y_1(t)^2}{y_1(t)^2 + 1}$$

$$y_2'(t) = y_2(t) \left(\frac{y_1(t)^2}{y_1(t)^2 + 1} - \alpha \right)$$

where $\alpha = 1/10$, $\rho = 29/10$ and $\tau = 21/50$, for t in $[0, 30]$. The history functions are

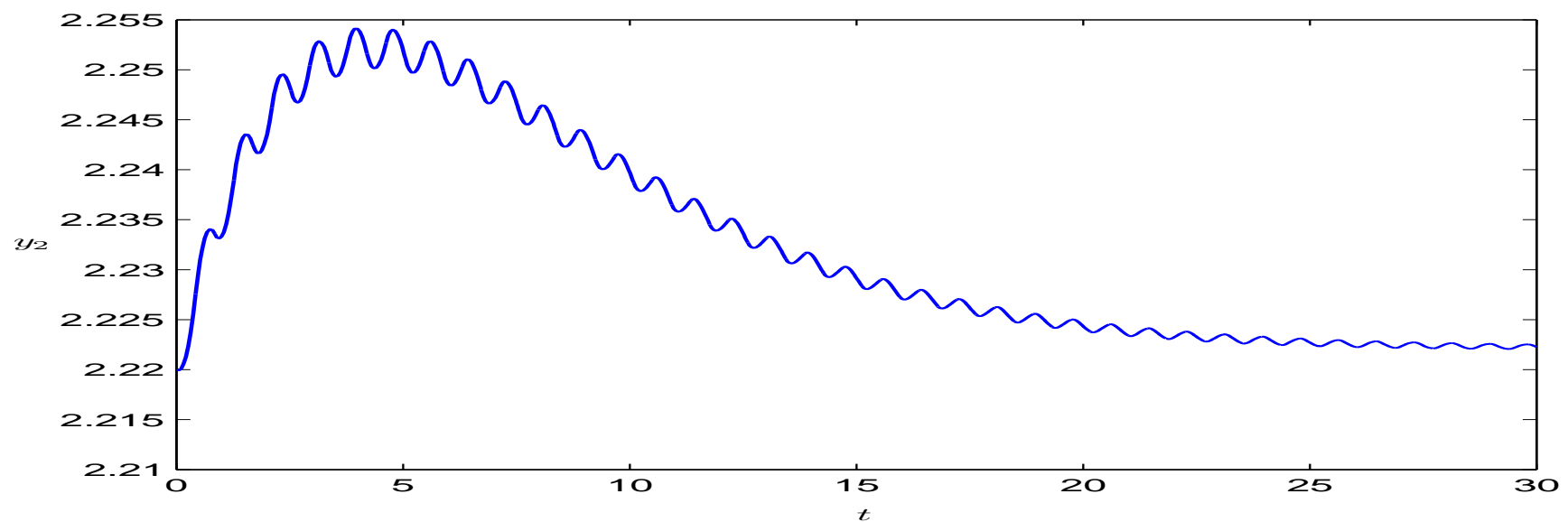
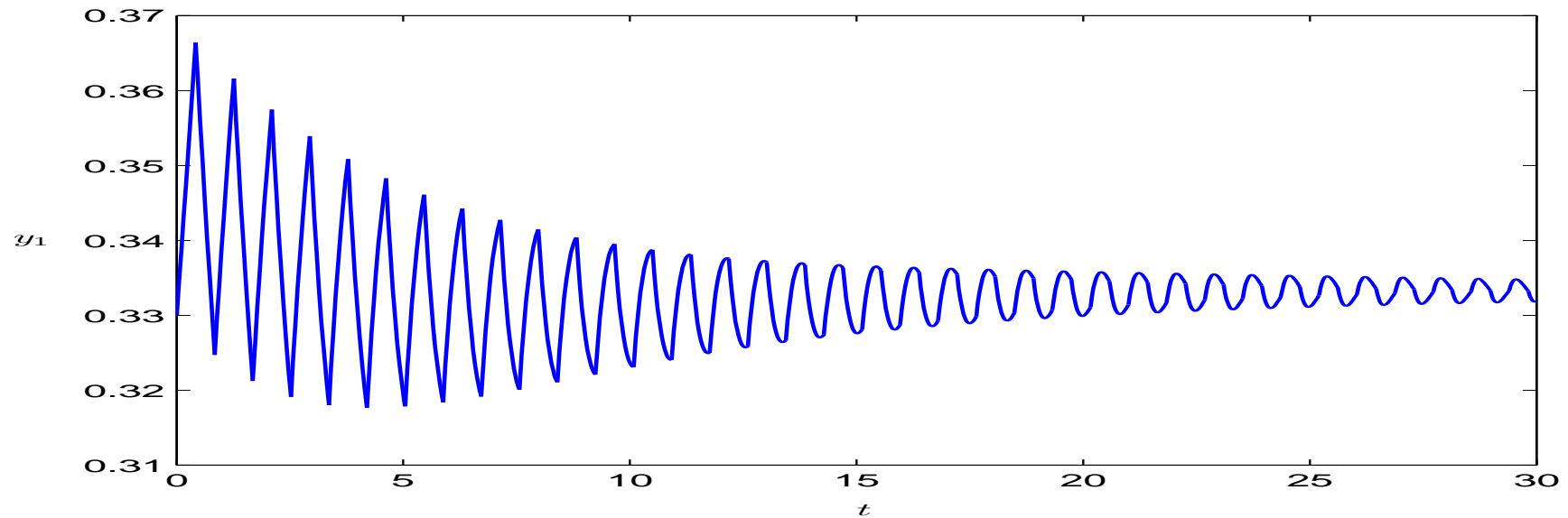
$$\phi_1(t) = \frac{33}{100} - \frac{1}{10}t$$

$$\phi_2(t) = \frac{111}{50} + \frac{1}{10}t$$

for $t \leq 0$.

Modeling with DE - An Example

■ The predator-prey model



Modeling with DE - Numerical Simulation

- Classical Theory of Step by Step Integration for IVPs.

- ◆ Runge-Kutta (RK).
- ◆ Linear Multistep (LM).

+

- Continuous Solution using Polynomial Approximation.

- ◆ Continuous Runge-Kutta (CRK).
- ◆ Linear Multistep methods have natural approximating polynomials.

⇓

- **DDEs** : Combining an “interpolation” method (for evaluating delayed solution values) with an ODE integration method (for solving the resulting “ODE”).

Modeling with DE - Special Difficulties with DDEs

■ Derivative Discontinuities

In general

$$\phi'(t_0) \neq f(t_0, \phi(t_0), \phi(t_0 - \sigma_1), \dots, \phi(t_0 - \sigma_\nu))$$

Due to the existence of *delays*, discontinuities **propagate** along the integration interval.

Solution is **smoothed** for RDDEs but in general not for NDDEs.

The RK and LM methods fail in presence of discontinuities.

Treatment : Tracking Discontinuities and forcing them to be mesh points.

Modeling and Sensitivity Analysis - Definitions

■ Parameterized Models

◆ A parameterized IVP

$$\begin{aligned}y'(t; \mathbf{p}) &= f(t, y(t; \mathbf{p}); \mathbf{p}) \\ y(t_0) &= y_0(\mathbf{p})\end{aligned}$$

For example $\mathbf{p} = [\mu]$ in the Van der Pol oscillator

$$\begin{aligned}y_1'(t) &= y_2(t) \\ y_2'(t) &= \mu(1 - y_1^2)y_2 - y_1\end{aligned}$$

◆ A simple parameterized DDE

$$\begin{aligned}y'(t; \mathbf{p}) &= f(t, y(t; \mathbf{p}), y(t - \sigma(t; \mathbf{p})); \mathbf{p}) \text{ for } t_0(\mathbf{p}) \leq t \\ y(t; \mathbf{p}) &= \phi(t; \mathbf{p}), \text{ for } t \leq t_0(\mathbf{p})\end{aligned}$$

Modeling and Sensitivity Analysis - Definitions

■ Forward Sensitivity Analysis

- ◆ The (first order) solution sensitivity with respect to the model parameter p_i is defined as the vector

$$s_i(t; \mathbf{p}) = \left\{ \frac{\partial}{\partial p_i} \right\} y(t; \mathbf{p}), \quad (i = 1, \dots, \mathcal{L})$$

- ◆ The second order solution sensitivity with respect to the model parameters p_i and p_j is defined as the vector

$$r_{ij}(t; \mathbf{p}) = \left\{ \frac{\partial}{\partial p_j} \right\} s_i(t; \mathbf{p}) = \left\{ \frac{\partial^2}{\partial p_j \partial p_i} \right\} y(t; \mathbf{p}), \quad (i, j = 1, \dots, \mathcal{L})$$

Modeling and Sensitivity Analysis - Importance

Sensitivity information can be used to:

- Estimate which parameters are most influential in affecting the behavior of the simulation. Such information is crucial for
 - ◆ Experimental Design
 - ◆ Data Assimilation
 - ◆ Reduction of complex nonlinear models
- Study of Dynamical Systems : Periodic orbits, the Lyapunov exponents, chaos indicators, and bifurcation analysis are fundamental objects for the complete study of a dynamical system, and they require computation of the sensitivities with respect to the initial conditions of the problem.
- Evaluate optimization gradients and Jacobians in the setting of
 - ◆ Dynamic Optimization
 - ◆ Parameter Estimation

Numerical Sensitivity Analysis of IVPs - Forward

- Finite Difference Approach

$$\left\{ \frac{\partial}{\partial p_i} \right\} y(t; \mathbf{p}) \approx \frac{y(t; \mathbf{p} + \mathbf{e}_i \Delta p_i) - y(t; \mathbf{p})}{\Delta p_i}$$

Due to the rounding errors, the approximation is only $\mathcal{O}(\sqrt{Tol})$ with the best choice for Δp_i .

- Internal Differentiation

$$y'(t; \mathbf{p}) = f(t, y(t; \mathbf{p}); \mathbf{p}), \quad y(t_0) = y_0(\mathbf{p})$$

⇓

Differentiation + Chain Rule + Clairaut's Theorem

⇓

$$s'_i = \frac{\partial f}{\partial y} s_i + \frac{\partial f}{\partial p_i}, \quad s_i(t_0) = \frac{\partial y_0(\mathbf{p})}{\partial p_i}, \quad (i = 1, \dots, \mathcal{L})$$

- Taylor Series method using extended rules of AD [Barrio 2006].
⇒ second(or higher)-order sensitivities.

Numerical Sensitivity Analysis of IVPs - An Example

■ The Van der Pol oscillator

$$\begin{aligned}y_1'(t) &= y_2(t) \\ y_2'(t) &= \mu(1 - y_1^2)y_2 - y_1\end{aligned}$$

⇓

$$\begin{pmatrix} s'_{1(1)}(t) \\ s'_{1(2)}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2\mu y_1 y_2 - 1 & \mu(1 - y_1^2) \end{pmatrix} \begin{pmatrix} s_{1(1)} \\ s_{1(2)} \end{pmatrix} + \begin{pmatrix} 0 \\ (1 - y_1^2)y_2 \end{pmatrix}$$

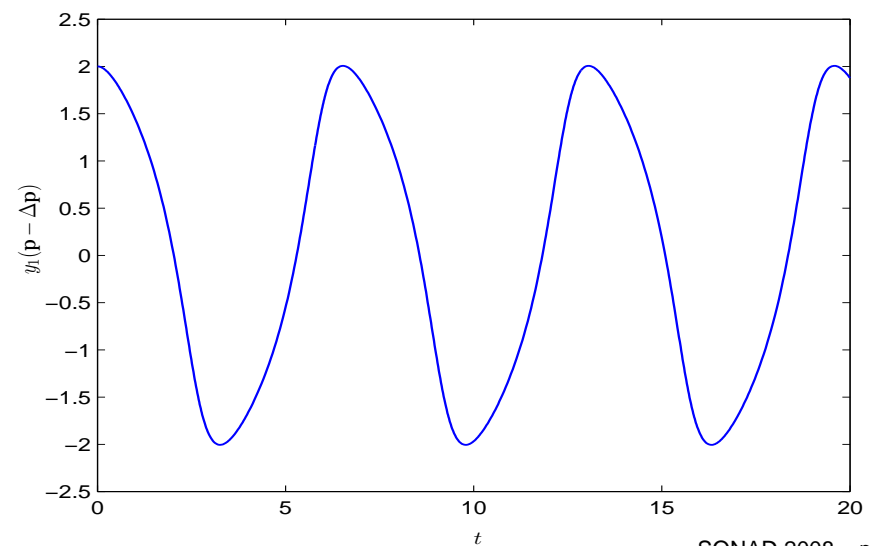
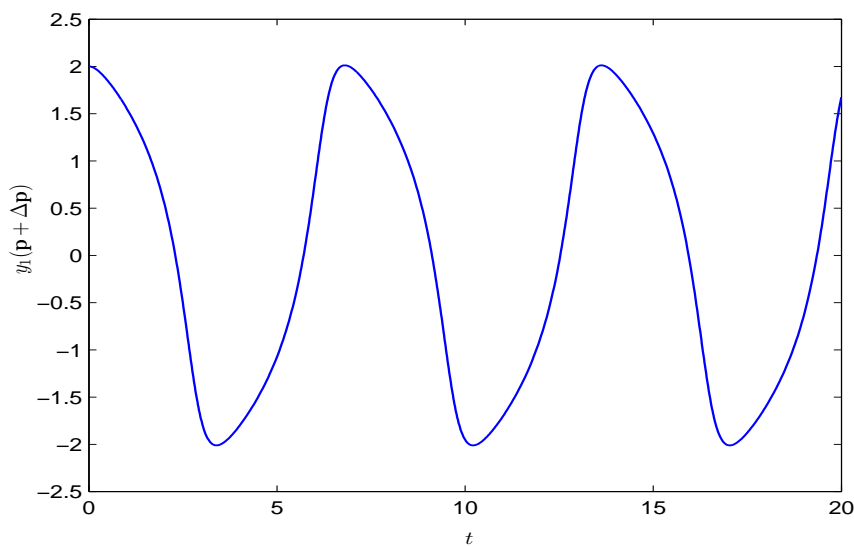
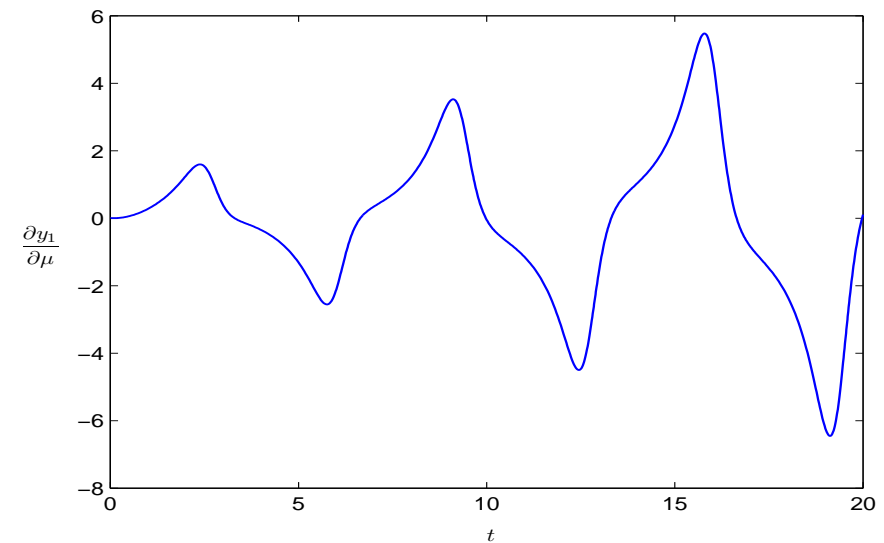
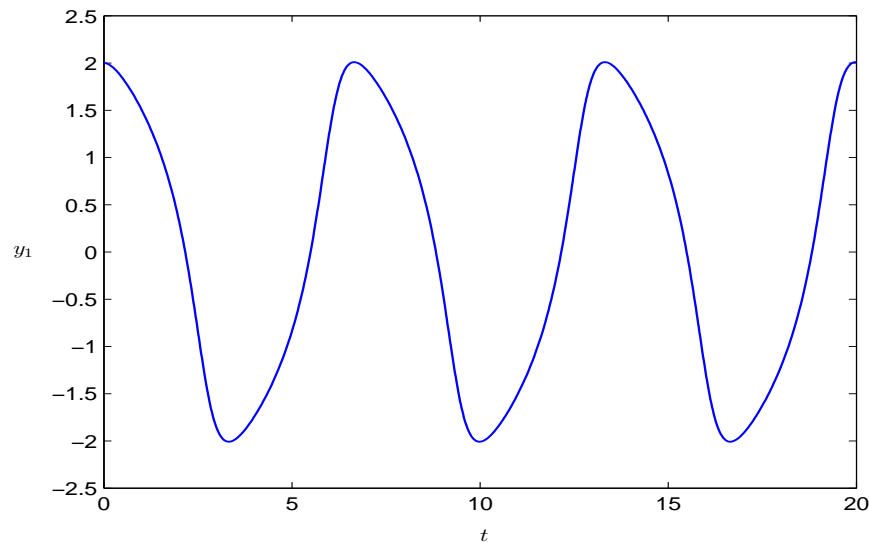
⇓

$$s'_{1(1)}(t) = s_{1(2)}$$

$$s'_{1(2)}(t) = (-2\mu y_1 y_2 - 1)s_{1(1)} + \mu(1 - y_1^2)s_{1(2)} + (1 - y_1^2)y_2$$

Numerical Sensitivity Analysis of IVPs - An Example

- The Van der Pol oscillator ($\Delta p = 0.2$). The decrease of y_1 at $t = 20$ can be explained as second order sensitivities being dominant ($\frac{\partial y_1}{\partial \mu}(t = 20) \gtrsim 0$, $\frac{\partial^2 y_1}{\partial \mu^2}(t = 20) < 0$).



DDEs and Sensitivity Analysis - Governing Equations

$$\begin{aligned}y'(t; \mathbf{p}) &= f(t, y(t; \mathbf{p}), y(\alpha(t, y; \mathbf{p}); \mathbf{p}), y'(\alpha(t, y; \mathbf{p}); \mathbf{p}); \mathbf{p}) \\y(t; \mathbf{p}) &= \phi(t; \mathbf{p}), \text{ for } t \leq t_0(\mathbf{p})\end{aligned}$$



Differentiation + Chain Rule + Clairaut's Theorem



$$\begin{aligned}s'_i(t) &= \frac{\partial f}{\partial y} s_i(t) + \frac{\partial f}{\partial y(\alpha_k)} \left(y'(\alpha_k) \left(\frac{\partial \alpha}{\partial y} s_i(t) + \frac{\partial \alpha}{\partial \mathbf{p}} \right) + s_i(\alpha) \right) \\&+ \frac{\partial f}{\partial y'(\alpha)} \left(y''(\alpha) \left(\frac{\partial \alpha}{\partial y} s_i(t) + \frac{\partial \alpha}{\partial \mathbf{p}} \right) + s'_i(\alpha) \right) \\&+ \frac{\partial f}{\partial \mathbf{p}}\end{aligned}$$

DDEs and Sensitivity Analysis - Proper Handling of Jumps

■ Hybrid ODE systems [Tolsma & Barton]

- ◆ Continuous transition at $t = \lambda$,

$$y(\lambda^+) = y(\lambda^-)$$

⇓

$$\frac{\partial y}{\partial p_l}(\lambda^+) = \frac{\partial y}{\partial p_l}(\lambda^-) + (y'(\lambda^-) - y'(\lambda^+)) \frac{\partial \lambda}{\partial p_l}$$

- ◆ Triggered by,

$$g(t, y, y'; \mathbf{p}) = 0$$

⇓

$$\frac{\partial g}{\partial y'} \left(\frac{\partial}{\partial t} \left(\frac{\partial y}{\partial p_l} \right) + y'' \frac{\partial \lambda}{\partial p_l} \right) + \frac{\partial g}{\partial y} \left(\frac{\partial y}{\partial p_l} + y' \frac{\partial \lambda}{\partial p_l} \right) + \frac{\partial g}{\partial p_l} + \frac{\partial g}{\partial t} \frac{\partial \lambda}{\partial p_l} = 0$$

DDEs and Sensitivity Analysis - Proper Handling of Jumps

■ DDEs

- ◆ Discontinuity points of $y'(t; \mathbf{p})$,

$$\Lambda(\mathbf{p}) \equiv \{\lambda_1(\mathbf{p}), \lambda_2(\mathbf{p}), \dots\}.$$

- ◆ Identified by,

$$\alpha(\lambda_{r+1}(\mathbf{p}), y; \mathbf{p}) = \lambda_r(\mathbf{p}) \quad r = 1, 2, \dots$$

- ◆ Can be viewed as $\lambda_{r+1}(\mathbf{p})$ being a solution of,

$$\hat{g}(t, y; \mathbf{p}) = \alpha(t, y; \mathbf{p}) - \lambda_r(\mathbf{p}) = 0.$$

- ◆ Using the result for hybrid ODEs,

$$\frac{\partial \lambda_{r+1}(\mathbf{p})}{\partial p_l} = - \frac{\frac{\partial \alpha}{\partial y} \frac{\partial y}{\partial p_l} + \frac{\partial \alpha}{\partial p_l} - \frac{\partial \lambda_r(\mathbf{p})}{\partial p_l}}{\frac{\partial \alpha}{\partial y} y' + \frac{\partial \alpha}{\partial t}} \quad \text{for } r \geq 1$$
$$\frac{\partial \lambda_1(\mathbf{p})}{\partial p_l} = \frac{\partial t_0(\mathbf{p})}{\partial p_l}$$

DDEs and Sensitivity Analysis - Proper Handling of Jumps

■ Integration (first order sensitivities)

1. Initialize ($\lambda_1 = t_0(\mathbf{p})$).

2. $r \leftarrow 1$.

3. Integrate the equations up to a C^1 -discontinuity point (λ_{r+1}).

4. Update the state variables (sensitivities) using

$$\frac{\partial y}{\partial p_l}(\lambda_{r+1}^+) = \frac{\partial y}{\partial p_l}(\lambda_{r+1}^-) + (y'(\lambda_{r+1}^-) - y'(\lambda_{r+1}^+)) \frac{\partial \lambda_{r+1}(\mathbf{p})}{\partial p_l}$$

5. $r \leftarrow r + 1$ and restart (3).

DDEs and Sensitivity Analysis - Examples

■ The predator-prey model

$$y_1'(t) = y_1(t)(1 - y_1(t - \tau) - \rho y_1'(t - \tau)) - \frac{y_2(t)y_1(t)^2}{y_1(t)^2 + 1}$$
$$y_2'(t) = y_2(t) \left(\frac{y_1(t)^2}{y_1(t)^2 + 1} - \alpha \right)$$

where $\alpha = 1/10$, $\rho = 29/10$ and $\tau = 21/50$, for t in $[0, 30]$. The history functions are

$$\phi_1(t) = \frac{33}{100} - \frac{1}{10}t$$
$$\phi_2(t) = \frac{111}{50} + \frac{1}{10}t$$

for $t \leq 0$.

Parameters are

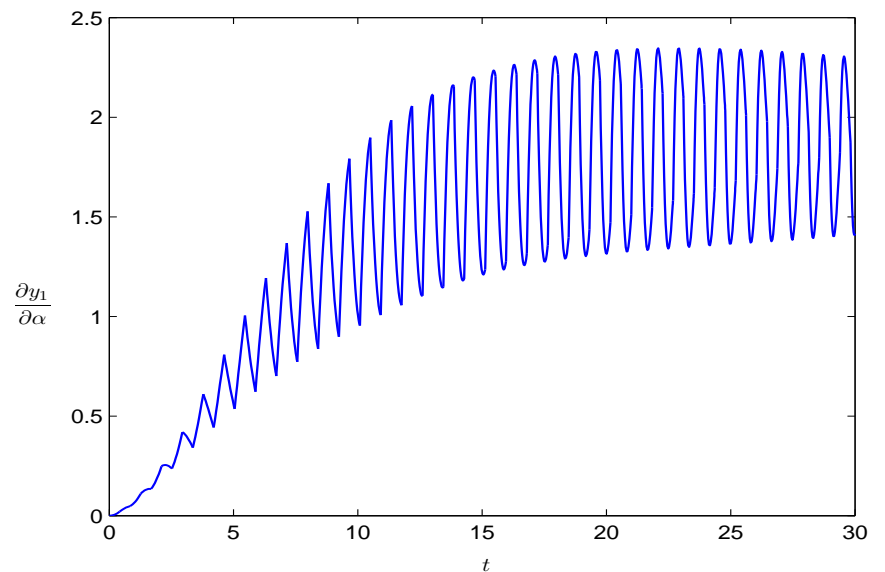
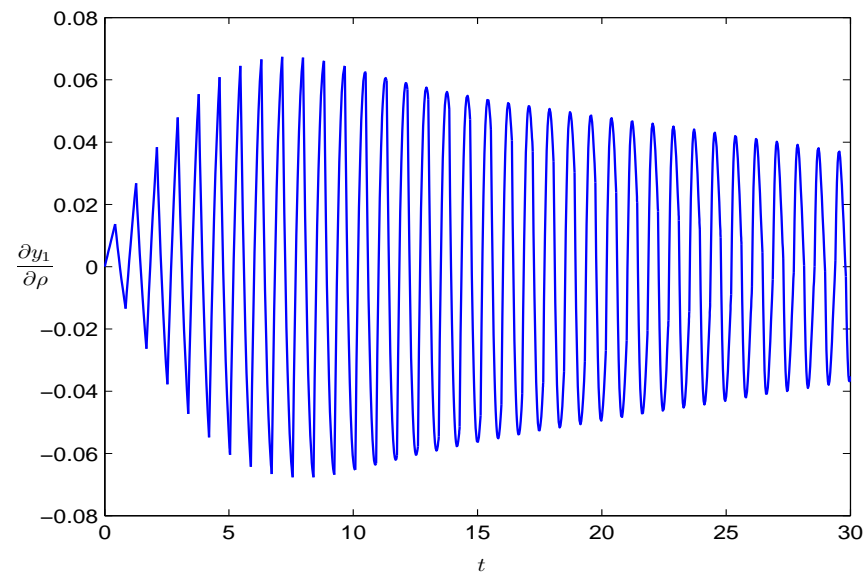
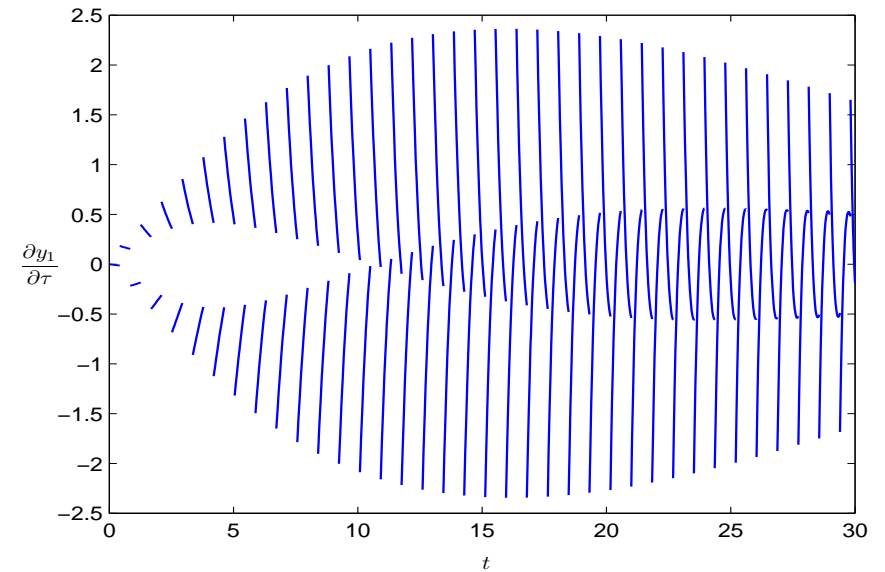
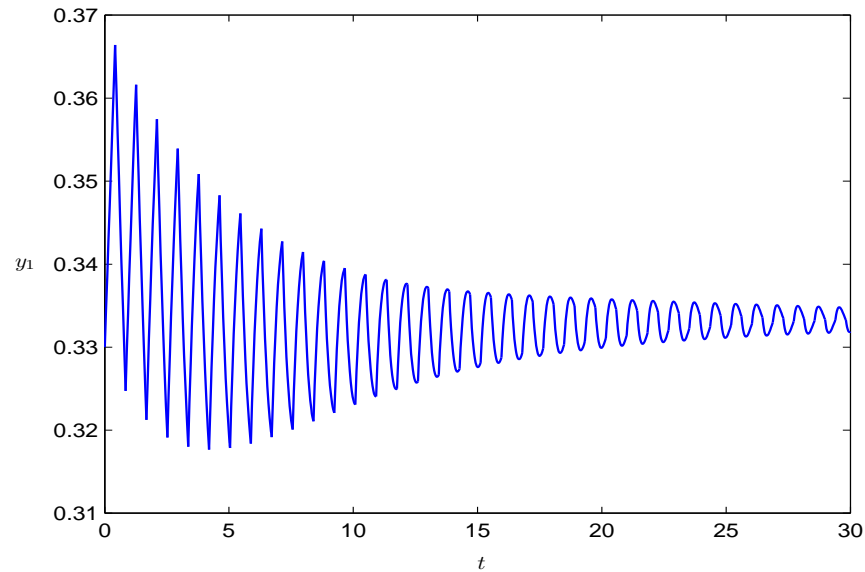
$$\mathbf{p} = [\tau, \rho, \alpha, a, b, c, d]$$

where

$$\phi_1(t) = a + b t$$
$$\phi_2(t) = c + d t$$

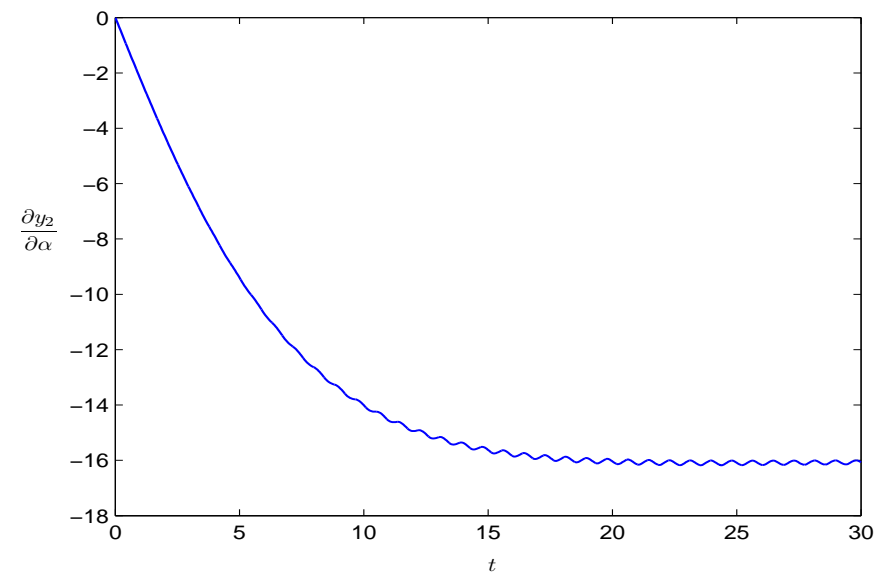
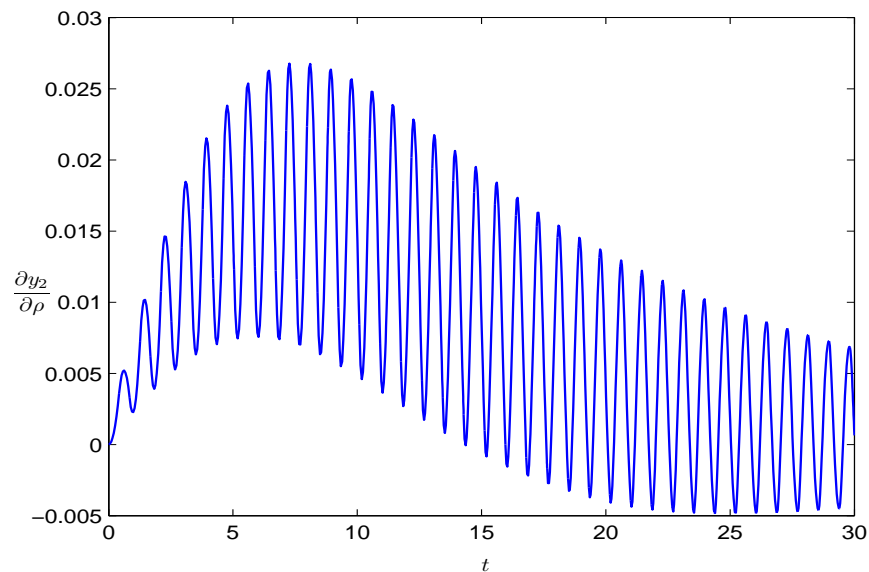
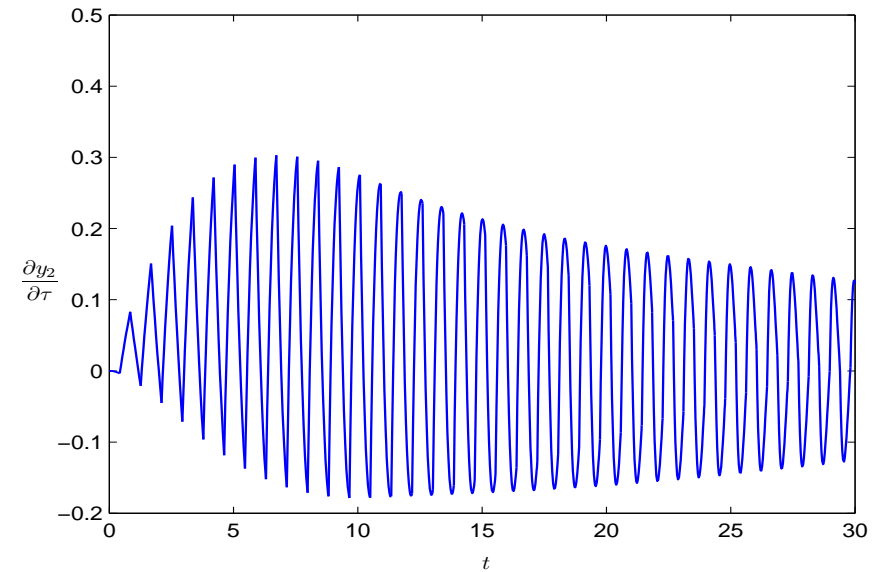
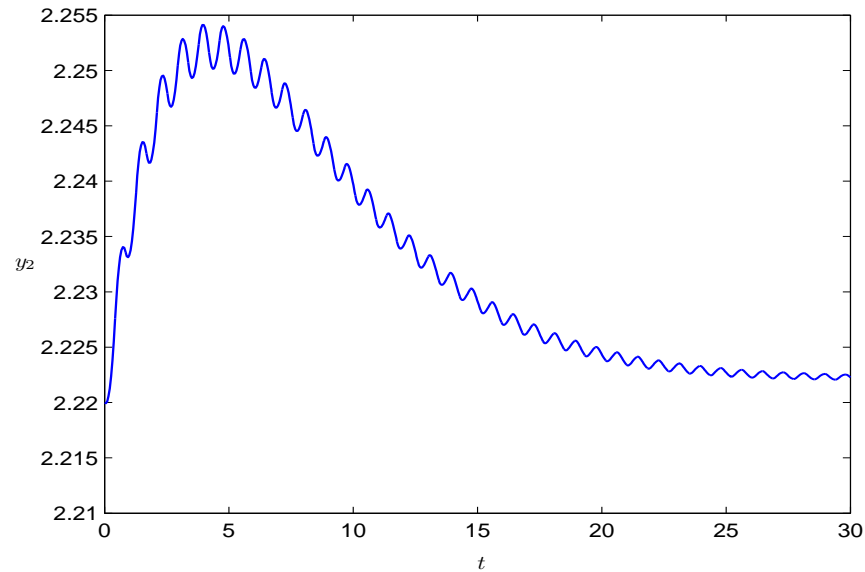
DDEs and Sensitivity Analysis - Examples

- The predator-prey model (structure-related parameters, y_1)



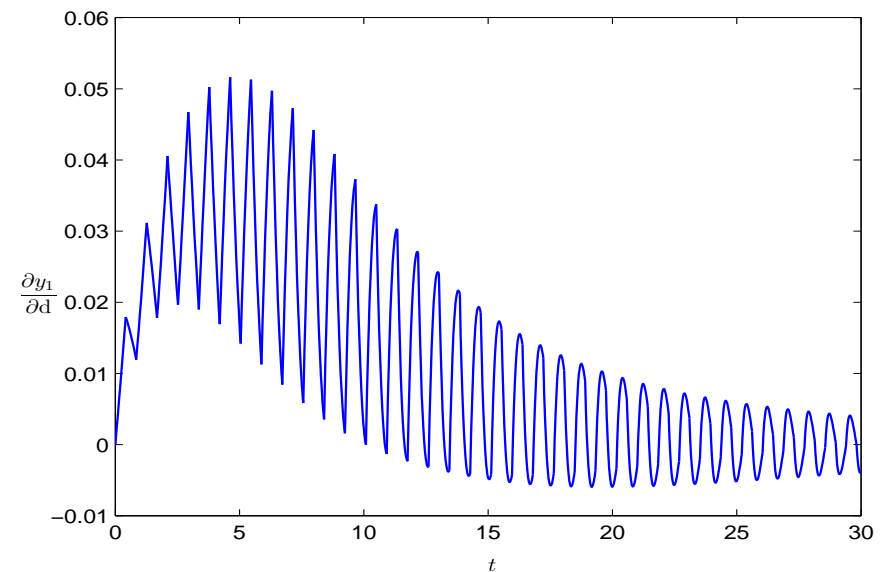
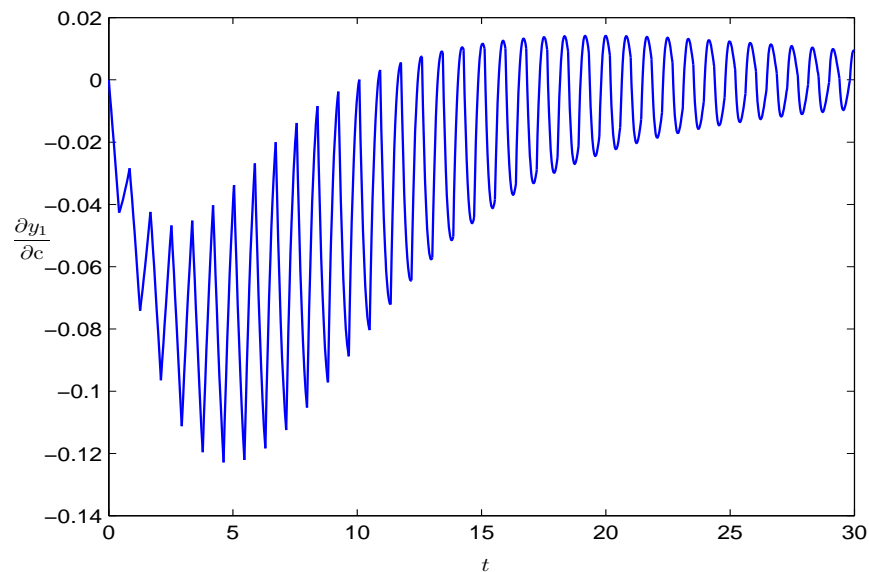
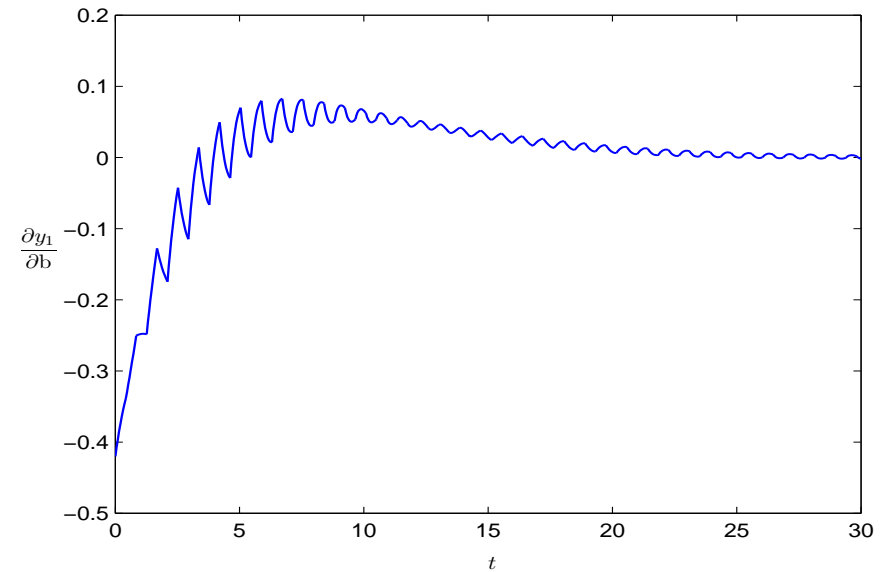
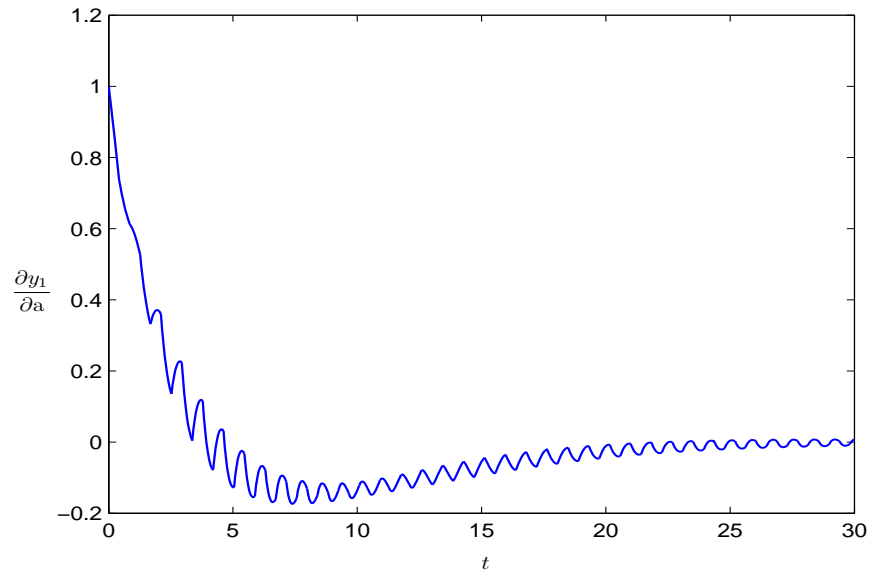
DDEs and Sensitivity Analysis - Examples

■ The predator-prey model (structure-related parameters, y_2)



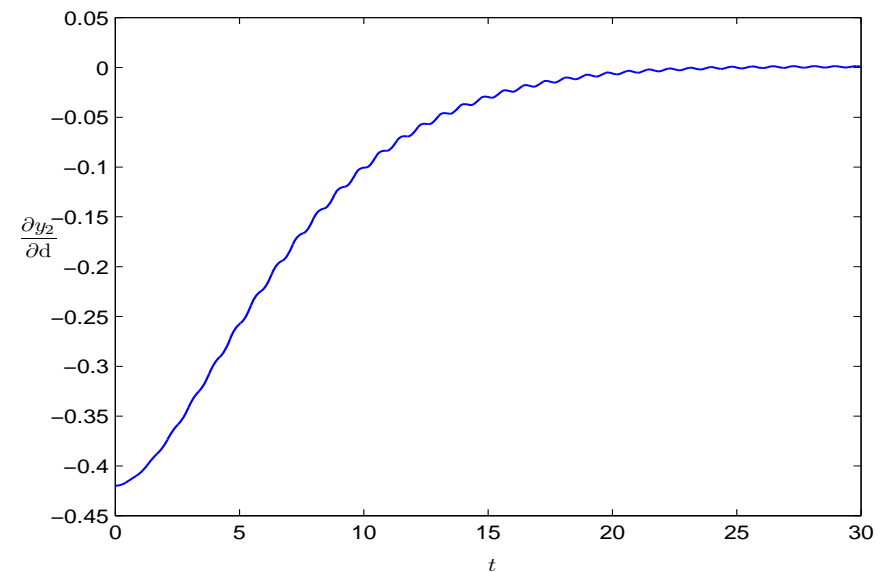
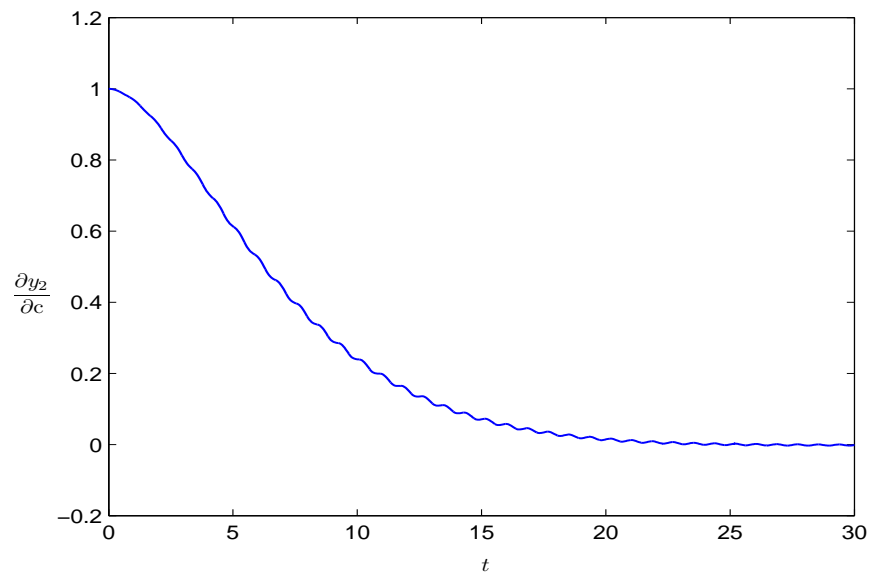
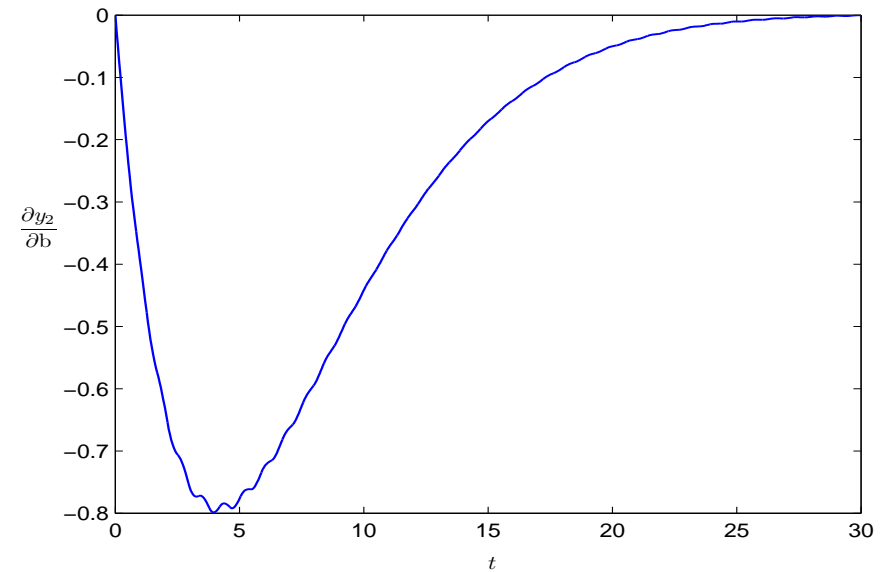
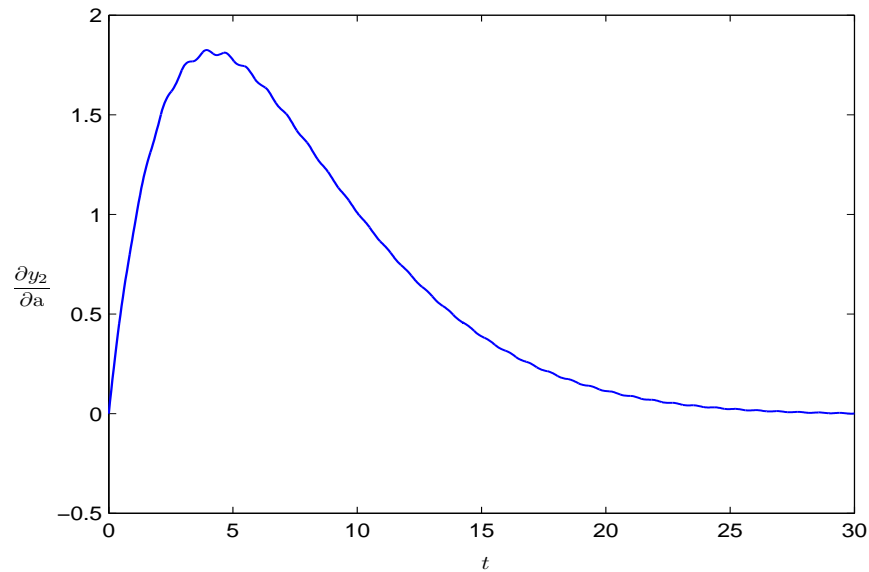
DDEs and Sensitivity Analysis - Examples

■ The predator-prey model (history-related parameters, y_1)



DDEs and Sensitivity Analysis - Examples

■ The predator-prey model (history-related parameters, y_2)



Modeling and Parameter Estimation - Definitions

■ Parameter Estimation Problem

◆ A System of Parameterized IVP

$$\begin{aligned}y'(t; \mathbf{p}) &= f(t, y(t; \mathbf{p}); \mathbf{p}) \\ y(t_0) &= y_0(\mathbf{p})\end{aligned}$$

or DDE

$$\begin{aligned}y'(t; \mathbf{p}) &= f(t, y(t; \mathbf{p}), y(t - \sigma(t; \mathbf{p})); \mathbf{p}) \text{ for } t_0(\mathbf{p}) \leq t \\ y(t; \mathbf{p}) &= \phi(t; \mathbf{p}), \text{ for } t \leq t_0(\mathbf{p})\end{aligned}$$

◆ A Set of Data (Observations/Measurements)

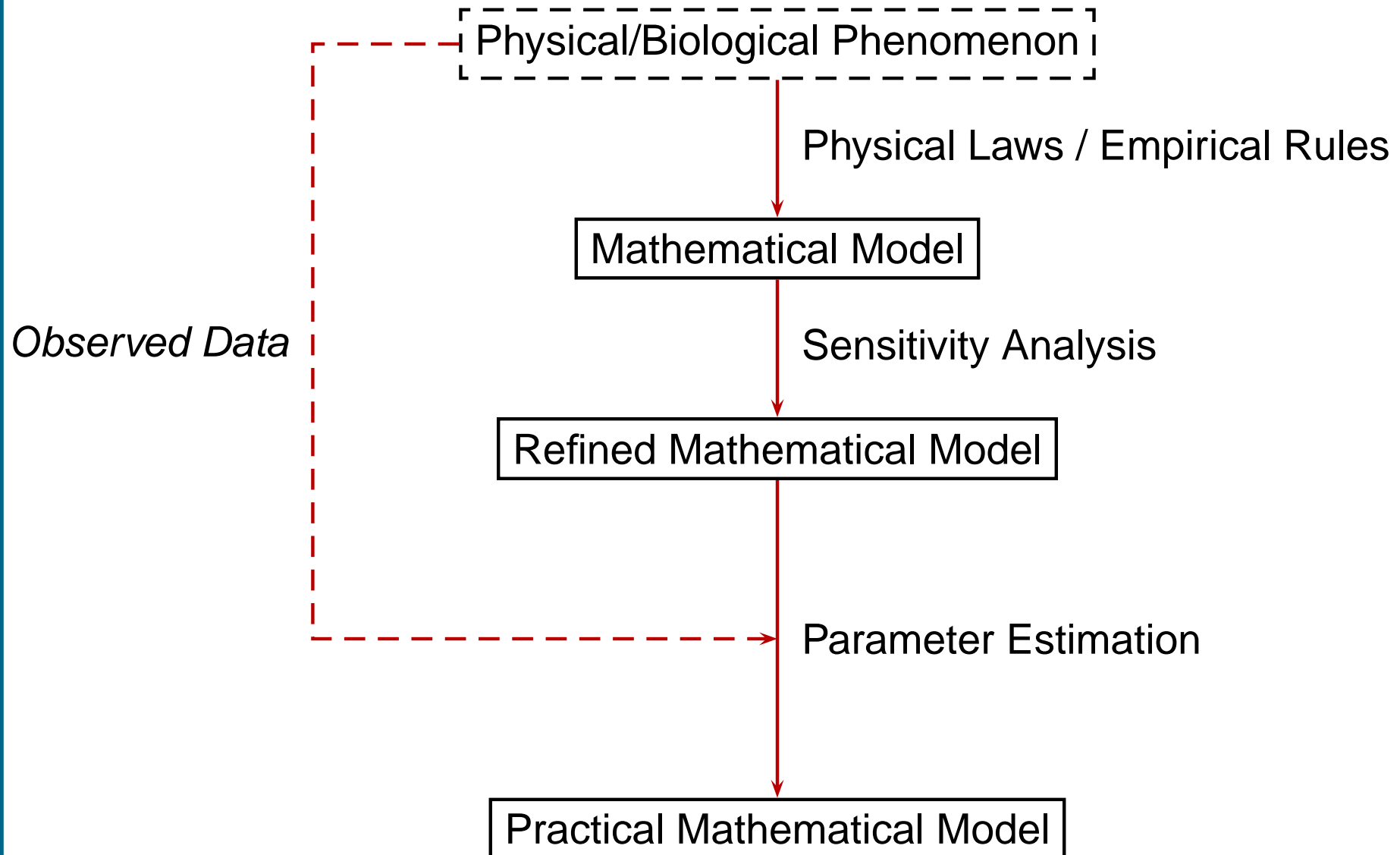
$$\{Y(\gamma_i) \approx y(\gamma_i; \mathbf{p}^*)\}$$

◆ Estimate \mathbf{p}^* by minimizing an objective function.

e.g.

$$W(\mathbf{p}) = \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})]^2.$$

Modeling and Parameter Estimation - Importance



■ Algorithms for Nonlinear Least-Squares

◆ Unconstrained

$$\min_{\mathbf{p}} W(\mathbf{p}) = \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})]^2.$$

↓

Levenberg-Marquardt

Variations of Sequential Quadratic Programming (SQP)

◆ Constrained

$$\min_{\mathbf{p}} W(\mathbf{p}) = \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})]^2,$$

$$c_j(\mathbf{p}) = 0, \quad j \in \mathcal{E},$$

$$c_j(\mathbf{p}) \geq 0, \quad j \in \mathcal{I}.$$

↓

Sequential Quadratic Programming (SQP)

- The *smoothness* of functions involved in the problem, the objective function and constraints, is a *necessary* assumption.

Numerical Parameter Estimation of IVPs

Steps :

1. Choose an initial guess for the parameters
2. Solve model equations
3. Check optimality conditions, (if satisfied \Rightarrow stop).
4. Choose a better value for the parameters and continue with (2)

If the optimization method needs to compute the **gradient** or the **Hessian** of the objective function,

$$\left(\frac{\partial W(\mathbf{p})}{\partial p_l} \right) = -2 \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right)$$

$$\left(\frac{\partial^2 W(\mathbf{p})}{\partial p_l \partial p_m} \right) = 2 \sum_i \left[\left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right) \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_m} \right) - [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial^2 y(\gamma_i; \mathbf{p})}{\partial p_l \partial p_m} \right) \right]$$

the sensitivity equations are usually used to provide the required values. An alternative approach is to use a divided-difference approximation.

DDEs and Parameter Estimation

[Paul , 1997] :

Assume that jumps in the derivative of $y(t; \mathbf{p})$ with respect to t occur at the points

$$\Lambda(\mathbf{p}) \equiv \{\lambda_1(\mathbf{p}), \lambda_2(\mathbf{p}), \dots\}.$$

Such discontinuities, when arising from the initial point $t_0(\mathbf{p})$ (and the initial function $\phi(t; \mathbf{p})$), may propagate into $W(\mathbf{p})$ via the solution values $\{y(\gamma_i; \mathbf{p})\}$.

The first and second order partial derivatives of the objective function are

$$\left(\frac{\partial W(\mathbf{p})}{\partial p_l} \right)_{\pm} = -2 \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right)_{\pm}$$

$$\left(\frac{\partial^2 W(\mathbf{p})}{\partial p_l \partial p_m} \right)_{\pm\pm} = 2 \sum_i \left[\left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right)_{\pm} \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_m} \right)_{\pm} - [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial^2 y(\gamma_i; \mathbf{p})}{\partial p_l \partial p_m} \right)_{\pm\pm} \right]$$

DDEs and Parameter Estimation

- Non-smooth optimization

- ◆ Consider the continuous function

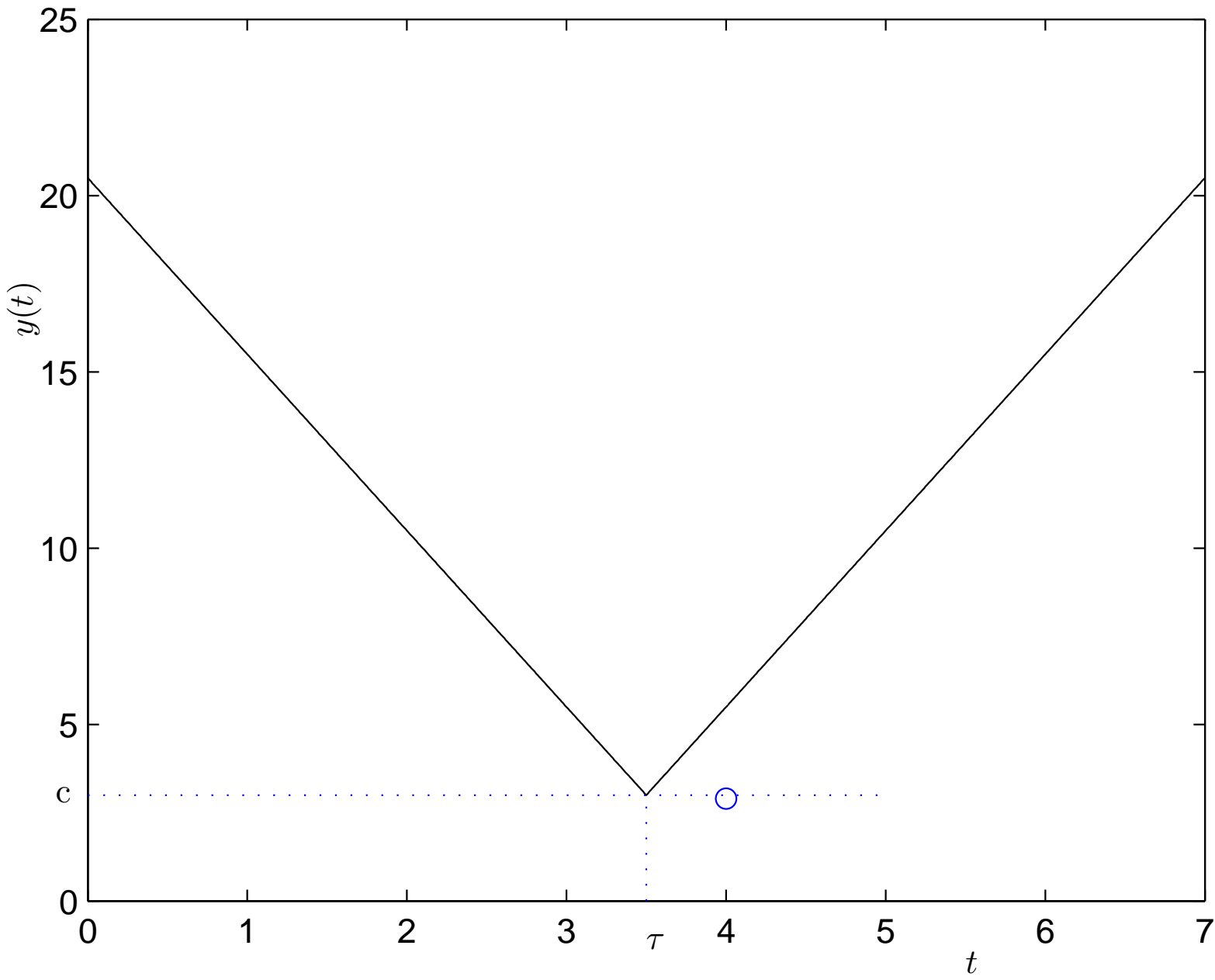
$$y(t) = \begin{cases} -5(t - \tau) + c, & \text{if } t < \tau \\ 5(t - \tau) + c, & \text{if } t \geq \tau \end{cases}$$

- ◆ with discontinuous derivative

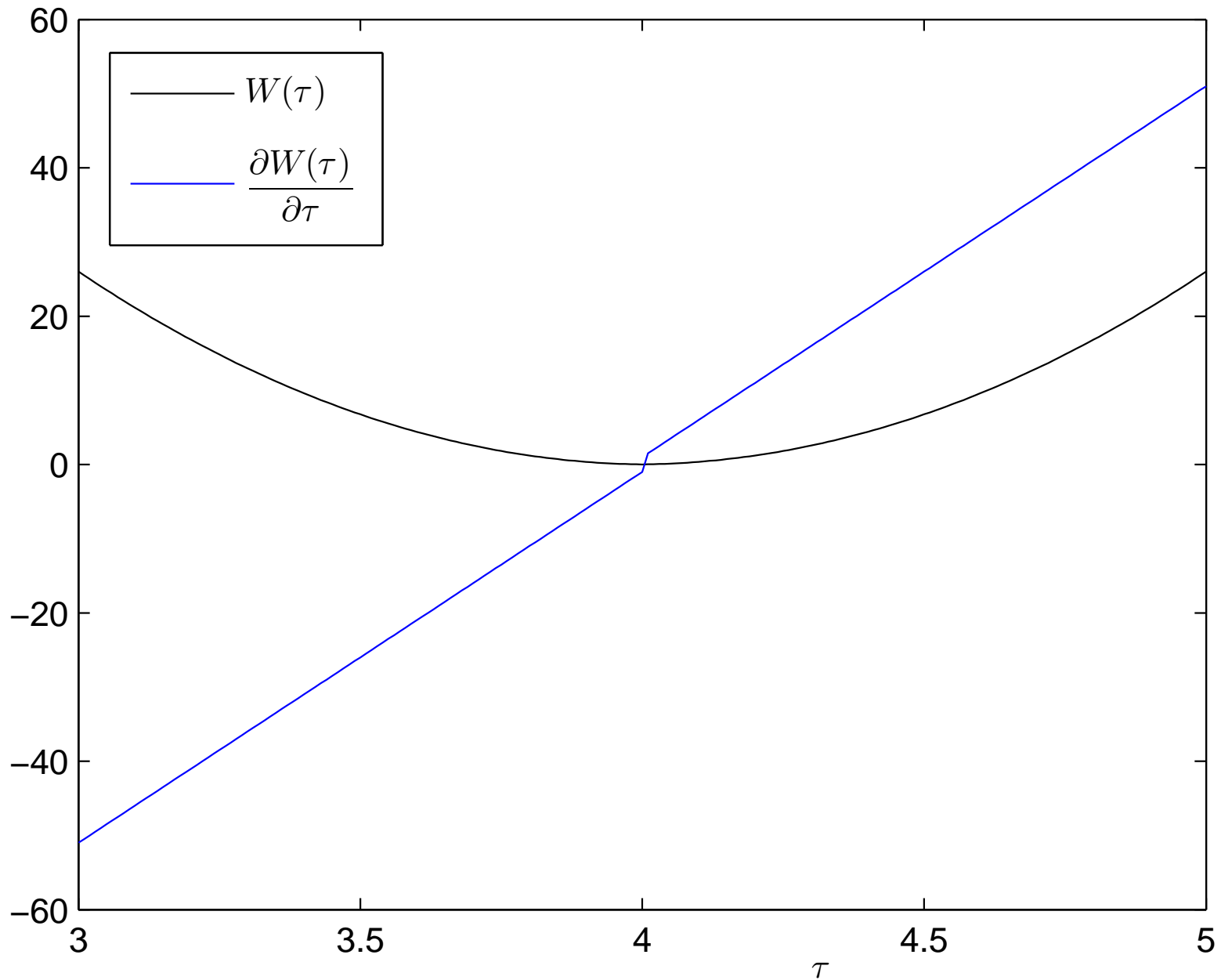
$$y'(t) = \begin{cases} -5, & \text{if } t < \tau \\ 5, & \text{if } t \geq \tau \end{cases}$$

- ◆ and observed value of y at the discontinuity point $\tau^* = 4$.

DDEs and Parameter Estimation



DDEs and Parameter Estimation



DDEs and Parameter Estimation

- Try to find τ^* using MATLAB's unconstrained minimization routine *fminunc*



31 iterations .

- Try to use MATLAB's constrained minimization routine *fmincon* with the added constraint

$$\tau \leq 4$$



2 iterations .

- Try to use MATLAB's constrained minimization routine *fmincon* with the added constraint

$$\tau \geq 4$$



2 iterations .

DDEs and Parameter Estimation - Safe Minimization

■ Appearance of Non-smoothness

- ◆ Partial derivatives (gradient) of the objective function

$$\left(\frac{\partial W(\mathbf{p})}{\partial p_l} \right)_{\pm} = -2 \sum_i [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right)_{\pm}$$

$$\left(\frac{\partial^2 W(\mathbf{p})}{\partial p_l \partial p_m} \right)_{\pm\pm} = 2 \sum_i \left[\left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_l} \right)_{\pm} \left(\frac{\partial y(\gamma_i; \mathbf{p})}{\partial p_m} \right)_{\pm} - [Y(\gamma_i) - y(\gamma_i; \mathbf{p})] \left(\frac{\partial^2 y(\gamma_i; \mathbf{p})}{\partial p_l \partial p_m} \right)_{\pm\pm} \right]$$

- ◆ Recall the jump equation for sensitivities

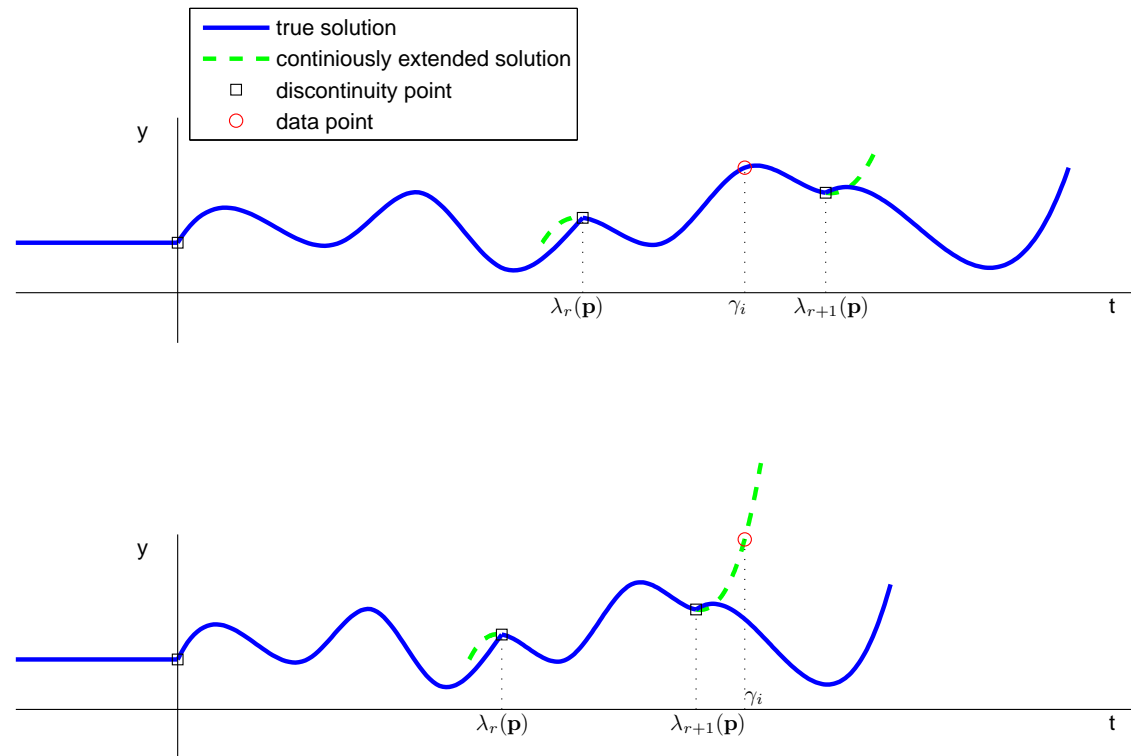
$$\frac{\partial y}{\partial p_l}(\lambda_{r+1}^+) = \frac{\partial y}{\partial p_l}(\lambda_{r+1}^-) + (y'(\lambda_{r+1}^-) - y'(\lambda_{r+1}^+)) \frac{\partial \lambda_{r+1}(\mathbf{p})}{\partial p_l}$$

⇓

- ◆ The General Rule : A jump occurs in $W(\mathbf{p})$ when

a discontinuity point $\lambda_{r+1}(\mathbf{p})$ **passes** a data point γ_i

■ Avoiding The Non-smoothness



Force the ordering by adding $\lambda_r(\mathbf{p}) \leq \gamma_i \leq \lambda_{r+1}(\mathbf{p})$ to the set of constraints.

The partial derivatives (gradient) of the new constraints, $\frac{\partial \lambda_{r+1}(\mathbf{p})}{\partial \mathbf{p}}$, can be computed recursively.

DDEs and Parameter Estimation - A Test Case

- Estimating τ for the predator-prey model

$$y_1'(t) = y_1(t)(1 - y_1(t - \tau) - \rho y_1'(t - \tau)) - \frac{y_2(t)y_1(t)^2}{y_1(t)^2 + 1}$$
$$y_2'(t) = y_2(t) \left(\frac{y_1(t)^2}{y_1(t)^2 + 1} - \alpha \right)$$

- Start with up to 10% random perturbation in original τ , and up to 3% randomly perturbed $y(t; \tau)$ as Data (Y). For γ 's we choose 10 random points, one of which is a discontinuity point.
- Run the parameter estimator 3 times.
- Results

Estimator Choice	FCN	OBJ
Very Simple	72,2697	0.000142917
Using Sensitivities	26,942	0.000142917
Adding Constraints	9,916	0.000142917