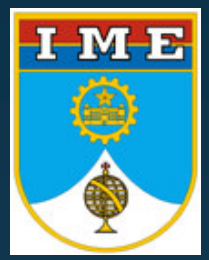
TÉCNICAS DE IDENTIFICAÇÃO DE PARÂMETROS E FUNÇÕES

H. R. B. Orlande
Department of Mechanical Engineering
Federal University of Rio de Janeiro
helcio@serv.com.ufrj.br

M. J. Colaço
Department of Mechanical and Materials Engineering
Military Institute of Engineering
colaco@ime.eb.br



5/3/2007

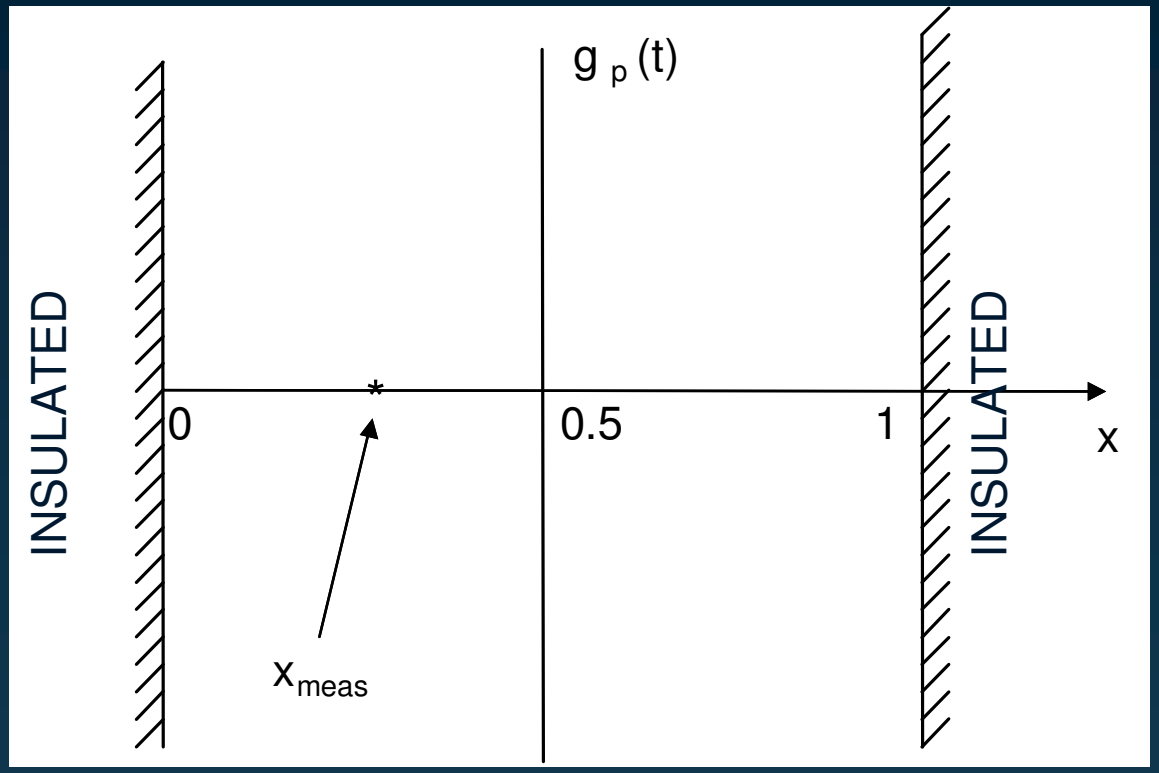
Inverse Problems

- **Introduction**
  - Basic Concepts
  - Classification
  - Overview of Solution Techniques
- **Parameter Estimation**
- **Function Estimation**
- **Examples**

Introduction

1. Inverse Heat Transfer Problem Concept
2. Classification of Inverse Heat Transfer Problems
3. Difficulties in the Solution of Inverse Heat Transfer Problems
4. An Overview of Solution Techniques for Inverse Heat Transfer Problems

Physical Problem



Mathematical Formulation

$$\frac{\partial^2 T(x,t)}{\partial x^2} + g_p(t)\,\delta(x-0.5) = \frac{\partial T(x,t)}{\partial t} \qquad \text{in } 0 < x < 1, \text{ for } t > 0$$

$$\frac{\partial T(0,t)}{\partial x} = 0 \qquad \text{at } x = 0, \text{ for } t > 0$$

$$\frac{\partial T(1,t)}{\partial x} = 0 \qquad \text{at } x = 1, \text{ for } t > 0$$

$$T(x, 0) = 0 \qquad \text{for } t = 0 \text{ , in } 0 < x < 1$$

Inverse Heat Transfer Problem Concept

DIRECT PROBLEM	INVERSE PROBLEM
Known: • Strength of the source term: $g_p(t)$	Known: • *Temperature measurements* taken at a location x_{meas}
Determine: • Temperature field $T(x,t)$	Estimate: • Strength of the source term: $g_p(t)$

Classification Of Inverse Heat Transfer Problems

With respect to the heat transfer mode:

- IHTP of conduction
- IHTP of convection (forced or natural)
- IHTP of surface radiation
- IHTP of radiation in participating medium
- IHTP of simultaneous conjugate modes
- IHTP of phase change

Classification Of Inverse Heat Transfer Problems

With respect to the unknown quantity:

- IHTP of boundary conditions
- IHTP of thermophysical properties
- IHTP of initial condition
- IHTP of source term
- IHTP of geometric characteristics of a heated body

Direct Problem Versus Inverse Problem

New Research Paradigm (Prof. J. BECK)

The results obtained from numerical simulations and from experiments are not simply compared *a posteriori*, but a close synergism exists between experimental and theoretical researchers during the course of the study, in order to obtain the maximum of information regarding the physical problem under picture.

Difficulties In The Solution Of Inverse Heat Transfer Problems

Inverse Problems are *ILL-POSED*.

A solution for a *well-posed* problem (Hadamard) must satisfy the conditions of:

- existence
- uniqueness
- stability with respect to the input data

An Overview Of Solution Techniques For Inverse Heat Transfer Problems

The solution of inverse problems requires its reformulation as well-posed problems.

- Integral equation approach.
- Series solution approach.
- Polynomial approach.
- Hyperbolization of the heat conduction equation.
- Space marching techniques together with filtering of the noisy input data
- Iterative filtering techniques.
- Steady-state techniques.
- Beck's sequential function specification method.
- Levenberg-Marquardt's method for the minimization of the least-squares norm.
- Tikhonov's regularization approach.
- Iterative regularization methods for parameter and function estimations.
- Genetic algorithms.

An Overview Of Solution Techniques For Inverse Heat Transfer Problems

The *time domain* over which measurements are used in the inverse analysis may be used to classify the methods of solution. Consider the estimation of the strength of the heat-source term $g_p(t)$ in the time domain. Three different possible time domains for the measurements used in the estimation of $g_p(t)$ up to time t_f include:

a. up to time $t_i < t_f$.
b. up to time $t_i < t_f$ plus few time steps. → **Sequential**

c. the whole time domain $0 < t < t_f$. → **Whole-domain**

Criteria for the Evaluation of IHTP Solution Procedures

- The predicted quantity should be accurate if the measured data are of high accuracy.
- The method should be stable with respect to measurement errors.
- The method should have a statistical basis and permit various statistical assumptions for the measurement errors.
- The method should not require the input data to be *a priori* smoothed.
- The method should be stable for small time steps or intervals. This permits a better resolution of the time variation of the unknown quantity than is permitted by large time steps.
- Temperature measurements from one or more sensors should be permitted.
- The method should not require continuous first derivatives of unknown functions. Furthermore, the method should be able to recover functions containing jump discontinuities.

Criteria for the Evaluation of IHTP Solution Procedures

- Knowledge of the precise starting time of the application of an unknown surface heat flux or source term should not be required.
- The method should not be restricted to any fixed number of measurements.
- The method should be able to treat complex physical situations, including, among others, composite solids, moving boundaries, temperature dependent properties, convective and radiative heat transfer, combined modes of heat transfer, multi-dimensional problems and irregular geometries.
- The method should be easy for computer programming.
- The computer cost should be moderate.
- The user should not have to be highly skilled in mathematics in order to use the method.
- The method should permit extension to more than one unknown.

Minimization Techniques

Minimize: $S=(\mathbf{Y}-\mathbf{T})^T(\mathbf{Y}-\mathbf{T})$

When the transient readings Y_i taken at times t_i, i=1,...I of a *single sensor* are used in the inverse analysis:

$$(\mathbf{Y}-\mathbf{T})^T=(Y_1-T_1\ ,\ Y_2-T_2\ ,\ldots,\ Y_I-T_I)$$

$$S=(\mathbf{Y}-\mathbf{T})^T\ (\mathbf{Y}-\mathbf{T})=\sum_{i=1}^{I}(Y_i-T_i)^2$$

Remark: Statistical hypotheses.

Parameter Estimation

The unknown function $g_p(t)$ is approximated as:

$$g_p(t) = \sum_{j=1}^{N} P_j C_j(t)$$

where: $C_j(t)$ are *known basis functions*

N is the *number of basis functions* used in the approximation (known for the analysis)

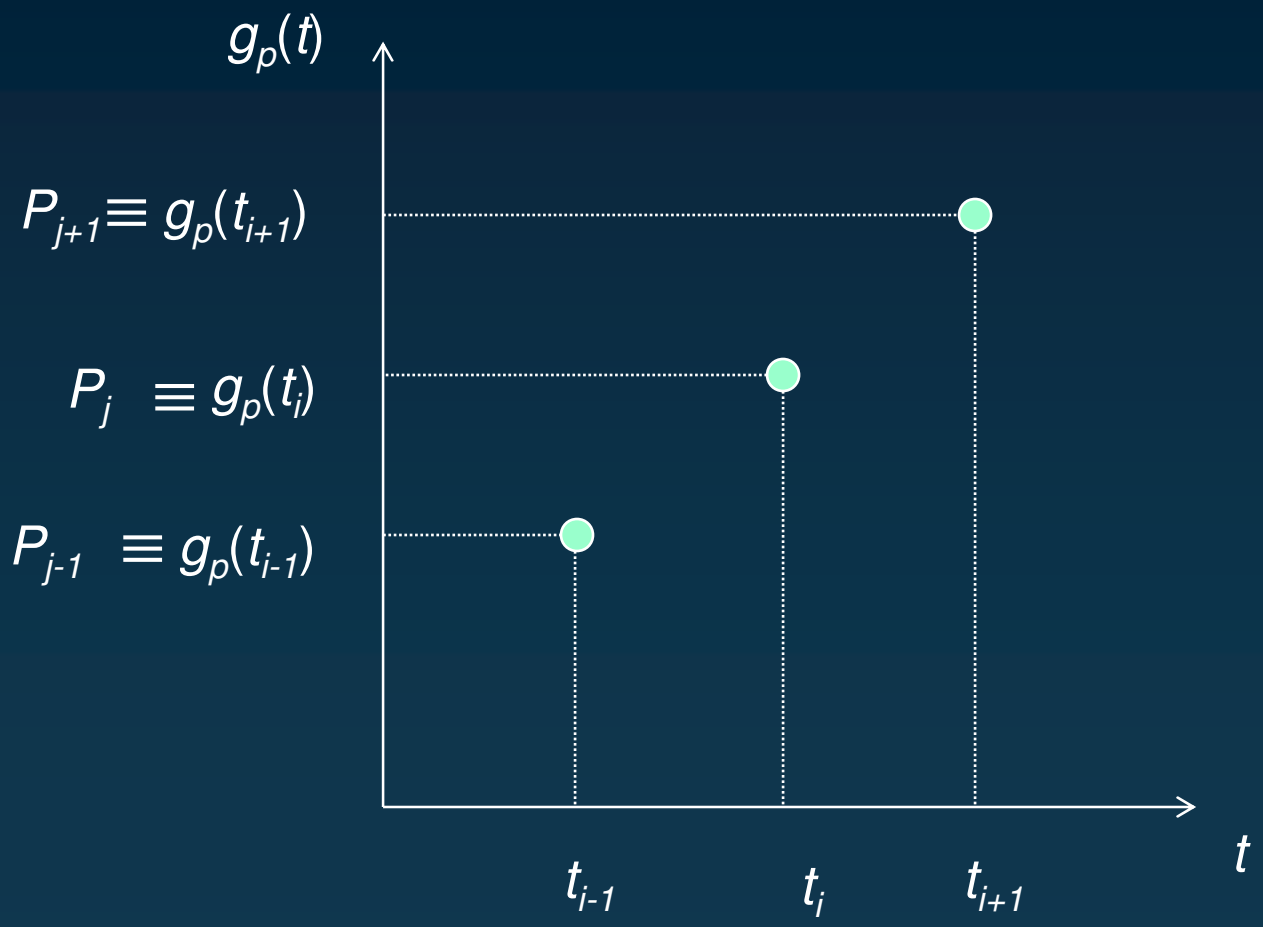
P_j are the *unknown* parameters

Function Estimation

- No assumption is made regarding the functional form of the unknown.
- Minimization is performed in an infinite dimensional space of functions, or minimization is performed in a finite dimensional space where N is large, e.g., $C_j(t) = \delta(t_i)$, $i = 1,...,I$, $N = I$.

I = Number of measurements
N = Number of unknown parameters

Function Estimation



An Overview Of Solution Techniques For Inverse Heat Transfer Problems

Remark: If the inverse heat transfer problem involves the estimation of only few unknown parameters from transient temperature measurements, the use of the ordinary least squares norm can be stable. However, if the inverse problem involves the estimation of a large number of parameters, such as the recovery of the unknown transient strength of the heat source term $g_p(t_i)$ at times t_i, $i=1,\ldots,I$, excursion and oscillation of the solution may occur. In this case, ***regularization*** (or stabilization) techniques are required.

An Overview Of Solution Techniques For Inverse Heat Transfer Problems

Tikhonov's Whole-Domain Regularization

Zeroth order:

$$S[g_p(t)] = \sum_{i=1}^{I} (Y_i - T_i)^2 + \alpha_0 \sum_{i=1}^{I} [g_p(t_i)]^2$$

First order:

$$S[g_p(t)] = \sum_{i=1}^{I} (Y_i - T_i)^2 + \alpha_1 \sum_{i=1}^{I-1} [g_p(t_{i+1}) - g_p(t_i)]^2$$

- α_0 and α_1 are the regularization parameters.

An Overview Of Solution Techniques For Inverse Heat Transfer Problems

Beck's Sequential Function Specification Technique

$$S[g_p(t_i)] = \sum_{s=i}^{i+r-1} (Y_s - T_s)^2$$

where *r* is the number of future measurements

- *Regularization* is obtained from the least-squares averaging capabilities and from the measurements taken at future time steps.

An Overview Of Solution Techniques For Inverse Heat Transfer Problems

Alifanov's Iterative Regularization

- Regularization is obtained from the stopping criterion utilized for the iterative procedure.

Parameter Estimation

- **Deterministic Methods**
  - Newton-Gauss' Method
  - Levenberg-Marquardt's Method
  - Steepest Descent Method
  - Conjugate Gradient Method
  - Newton's Method
  - Quasi-Newton Methods
- **Evolutionary and Stochastic Methods**
  - Genetic Algorithms
  - Differential Evolution
  - Particle Swarm
  - Simulated Annealing
- **Hybrid Methods**
- **Sequential Parameter Estimation Technique**

Newton-Gauss' Method

Maximum Likelihood Objective Function

$$S_{ML}(\mathbf{P}) = \left[\mathbf{Y} - \mathbf{T}(\mathbf{P})\right]^T \mathbf{W} \left[\mathbf{Y} - \mathbf{T}(\mathbf{P})\right]$$

Hypotheses:

- The errors are additive, with zero mean and normally distributed.
- The statistical parameters describing the errors are known.
- There are no errors in the independent variables.
- There is no prior information about **P**.

Newton-Gauss' Method

For uncorrelated measurements:

$$\mathbf{W} = \begin{bmatrix} 1/\sigma_1^2 & & & 0 \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ 0 & & & 1/\sigma_I^2 \end{bmatrix}$$

For the minimization of $S_{ML}(\mathbf{P})$:

$$\frac{\partial S_{ML}(\mathbf{P})}{\partial P_1} = \frac{\partial S_{ML}(\mathbf{P})}{\partial P_2} = \cdots = \frac{\partial S_{ML}(\mathbf{P})}{\partial P_N} = 0$$

$$\nabla S(\mathbf{P}) = -2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y} - \mathbf{T}(\mathbf{P})] = 0$$

where **J** is the *Sensitivity Matrix*.

Sensitivity Matrix And Sensitivity Coefficients

$$\mathbf{J} = \left[\frac{\partial \mathbf{T}^T}{\partial \mathbf{P}}\right]^T = \begin{bmatrix} \frac{\partial T_1}{\partial P_1} & \frac{\partial T_1}{\partial P_2} & \frac{\partial T_1}{\partial P_3} & \cdots & \frac{\partial T_1}{\partial P_N} \\ \frac{\partial T_2}{\partial P_1} & \frac{\partial T_2}{\partial P_2} & \frac{\partial T_2}{\partial P_3} & \cdots & \frac{\partial T_2}{\partial P_N} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial T_I}{\partial P_1} & \frac{\partial T_I}{\partial P_2} & \frac{\partial T_I}{\partial P_3} & \cdots & \frac{\partial T_I}{\partial P_N} \end{bmatrix}$$

$$J_{ij} = \frac{\partial T_i}{\partial P_j}$$

How To Compute The Sensitivity Coefficients ?

(i) Direct analytic solution

(ii) Boundary value problem

(iii) Finite-difference approximation

(i) Direct Analytic Solution For Determining Sensitivity Coefficients

The analytical solution for our test-problem at the measurement position is given by:

$$T(x_{meas},t)=\int_{t'=0}^{t} g_p(t')dt'+2\sum_{m=1}^{\infty} e^{-\beta_m^2 t}\cos(\beta_m x_{meas})\cos(0.5\beta_m)\int_{t'=0}^{t} e^{\beta_m^2 t'} g_p(t')dt'$$

By substituting the strength of the source term $g_p(t)$ given by

$$g_p(t)=\sum_{j=1}^{N} P_j C_j(t)$$

into the equation above and differentiating the resulting expression with respect to P_j, we find the expression for the sensitivity coefficient for the parameter P_j as

$$J_j \equiv \frac{\partial T}{\partial P_j}=\int_{t'=0}^{t} C_j(t')dt'+2\sum_{m=1}^{\infty} e^{-\beta_m^2 t}\cos(\beta_m x_{meas})\cos(0.5\beta_m)\int_{t'=0}^{t} e^{\beta_m^2 t'} C_j(t')dt'$$

(ii) The Boundary Value Problem Approach For Determining The Sensitivity Coefficients

By differentiating the direct problem with respect to the parameter P_j, after substituting $g_p(t)$ in its parameterized form, we obtain the sensitivity problem governing the sensitivity coefficients $J_j(x,t)$ as

$$\frac{\partial^2 J_j(x,t)}{\partial x^2} + C_j(t)\,\delta(x-0.5) = \frac{\partial J_j(x,t)}{\partial t} \qquad \text{in } 0 < x < 1, \quad \text{for } t > 0$$

$$\frac{\partial J_j}{\partial x} = 0 \qquad \text{at } x = 0, \quad \text{for } t > 0$$

$$\frac{\partial J_j}{\partial x} = 0 \qquad \text{at } x = 1, \quad \text{for } t > 0$$

$$J_j = 0 \qquad \text{for } t = 0\,, \quad \text{in } 0 < x < 1$$

(iii) Finite Difference Approximation For Determining Sensitivity Coefficients

$$J_{ij} \cong \frac{T_i(P_1, P_2, \ldots, P_j + \varepsilon P_j, \ldots, P_N) - T_i(P_1, P_2, \ldots, P_j, \ldots, P_N)}{\varepsilon P_j}$$

or

$$J_{ij} \cong \frac{T_i(P_1, P_2, \ldots, P_j + \varepsilon P_j, \ldots, P_N) - T_i(P_1, P_2, \ldots, P_j - \varepsilon P_j, \ldots, P_N)}{2\varepsilon P_j}$$

Linear Problems

$$-2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P})]=0$$

Linear Problems: **J** does not depend on **P** ⇒ $\mathbf{T}(\mathbf{P}) = \mathbf{J}\,\mathbf{P}$

$$\mathbf{P} = [\mathbf{J}^T\mathbf{W}\,\mathbf{J}]^{-1}[\mathbf{J}^T\mathbf{W}\,\mathbf{Y}]$$

Newton-Gauss' Method

$$-2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P})]=0$$

Nonlinear Problems: $\mathbf{J}\equiv\mathbf{J}(\mathbf{P})$ ⇒ $\mathbf{T}(\mathbf{P})=\mathbf{T}(\mathbf{P}^k)+\mathbf{J}^k(\mathbf{P}-\mathbf{P}^k)$

$$\mathbf{P}^{k+1}=\mathbf{P}^k+[\mathbf{J}^T\mathbf{W}\mathbf{J}]^{-1}\{\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P}^k)]\}$$

Identifiability Condition

$$\left|\mathbf{J}^T\mathbf{W}\mathbf{J}\right| \neq 0 \qquad \text{or} \qquad \left|\mathbf{J}^T\mathbf{J}\right| \neq 0$$

ILL-CONDITIONED PROBLEMS: $\left|\mathbf{J}^T\mathbf{J}\right| \approx 0$

- Small magnitudes for the sensitivity coefficients
- Columns of the sensitivity matrix are linearly dependent

The Levenberg-Marquardt Method

$$\mathbf{P}^{k+1} = \mathbf{P}^{k} + [\mathbf{J}^{T}\mathbf{W}\mathbf{J} + \lambda^{k}\Omega^{k}]^{-1}\mathbf{J}^{T}\mathbf{W}[\mathbf{Y} - \mathbf{T}(\mathbf{P}^{k})]$$

where λ^k is the *damping parameter* and Ω^k is a *diagonal matrix.*

- The Levenberg-Marquardt Method is related to ***Tikhonov's regularization*** approach.
- Compromise between steepest-descent method and Gauss' method.
- Simple, powerful and straightforward iterative procedure.
- Capable of treating complex physical situations.
- Easy to program.
- Stable and converges fast.

The levenberg-Marquardt method

Stopping Criteria

$$S(\mathbf{P}^{k+1}) < \varepsilon_1$$

$$\left\| (\mathbf{J}^k)^T \left[\mathbf{Y} - \mathbf{T}(\mathbf{P}^k) \right] \right\| < \varepsilon_2$$

$$\left\| \mathbf{P}^{k+1} - \mathbf{P}^k \right\| < \varepsilon_3$$

where ε_1, ε_2 and ε_3 are user prescribed tolerances and || . || is the vector Euclidean norm.

The Levenberg-Marquardt Method

Suppose that temperature measurements $\mathbf{Y}=(Y_1, Y_2, ..., Y_I)$ are given at times t_i, $i=1,...,I$. Also, suppose an initial guess $\mathbf{P}^0$ is available for the vector of unknown parameters $\mathbf{P}$. Choose a value for μ^0, say, $\mu^0 = 0.001$ and set $k=0$. Then,

1. Solve the direct heat transfer problem with the available estimate $\mathbf{P}^k$ in order to obtain the temperature vector $\mathbf{T}(\mathbf{P}^k)=(T_1, T_2, ..., T_I)$.
2. Compute $S(\mathbf{P}^k)$.
3. Compute the sensitivity matrix $\mathbf{J}^k$ and then the matrix $\mathbf{W}^k$.
4. Compute the new estimate $\mathbf{P}^{k+1}$
5. Solve the direct problem with the new estimate $\mathbf{P}^{k+1}$ in order to find $\mathbf{T}(\mathbf{P}^{k+1})$. Then compute $S(\mathbf{P}^{k+1})$
6. If $S(\mathbf{P}^{k+1}) > S(\mathbf{P}^k)$, replace μ^k by $10\mu^k$ and return to step 4.
7. If $S(\mathbf{P}^{k+1}) < S(\mathbf{P}^k)$, accept the new estimate $\mathbf{P}^{k+1}$ and replace μ^k by $0.1\mu^k$.
8. Check the stopping criteria.Stop the iterative procedure if any of them is satisfied; otherwise, replace k by $k+1$ and return to step 3.

Deterministic Methods

➔ Steepest-Descent Method:

Iterative process:

$$\mathbf{P}^{k+1} = \mathbf{P}^k + \alpha^k \mathbf{d}^k$$

$$\mathbf{d}^k = -\nabla S\left(\mathbf{P}^k\right)$$

where:

P is the vector of parameters to be optimized

α is the search-step size

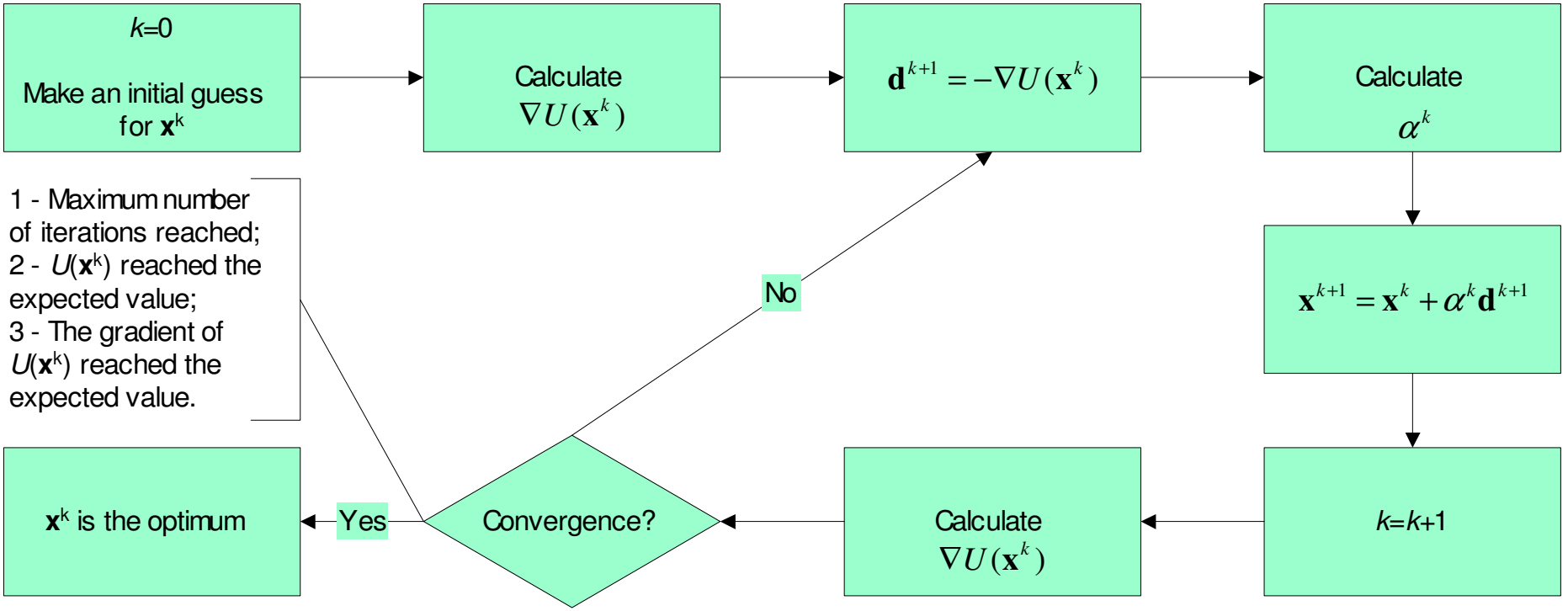
d is the direction of descent

S is the objective function

k is the iteration number

Deterministic Methods

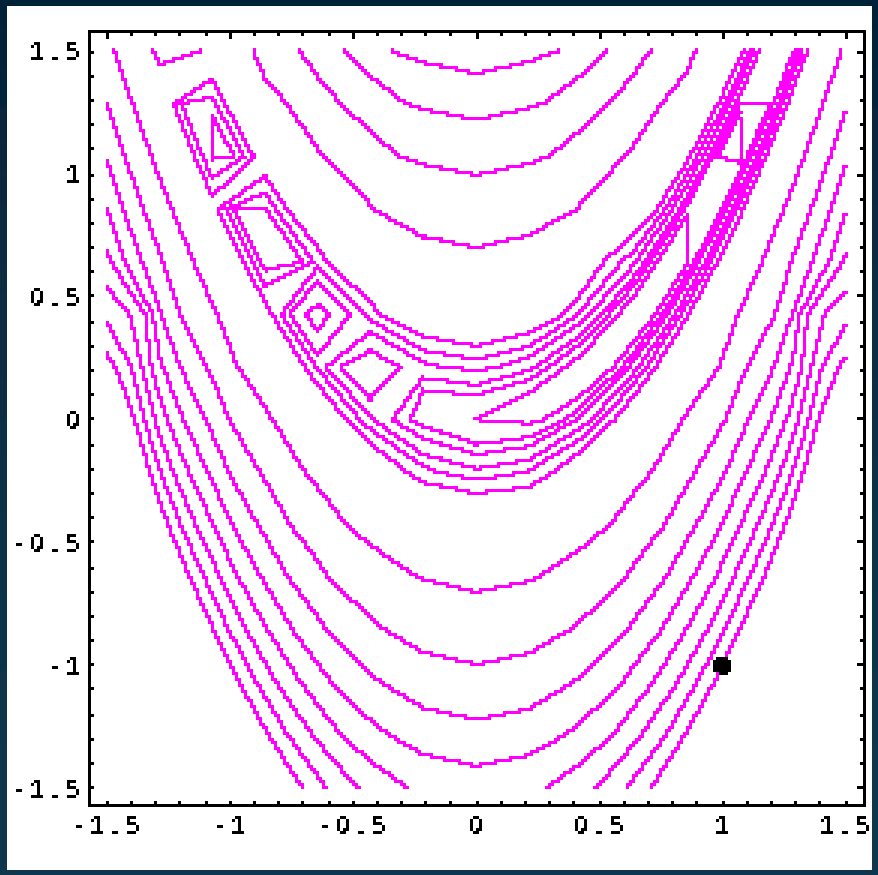
➔ Steepest-Descent Method:



Deterministic Methods

- Steepest-Descent Method:

SLOW!!!!!!



Deterministic Methods

➔ Conjugate Gradient Method:

Iterative process:

$$\mathbf{P}^{k+1} = \mathbf{P}^{k} + \alpha^{k}\mathbf{d}^{k}$$

$$\mathbf{d}^{k} = -\nabla S\left(\mathbf{P}^{k}\right) + \gamma^{k}\mathbf{d}^{k-1}$$

where:

P is the vector of parameters to be optimized

α is the search-step size

d is the direction of descent

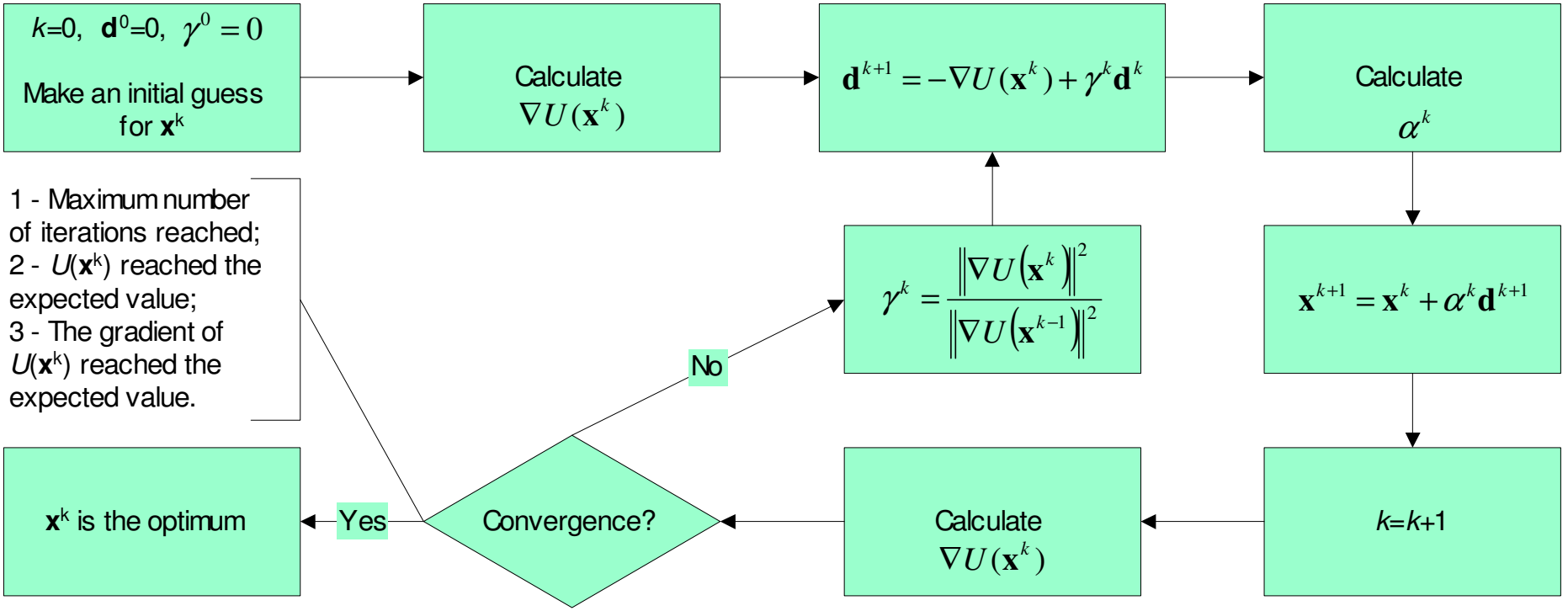
S is the objective function

k is the iteration number

γ is the conjugation coefficient

Deterministic Methods

➔ Conjugate Gradient Method:

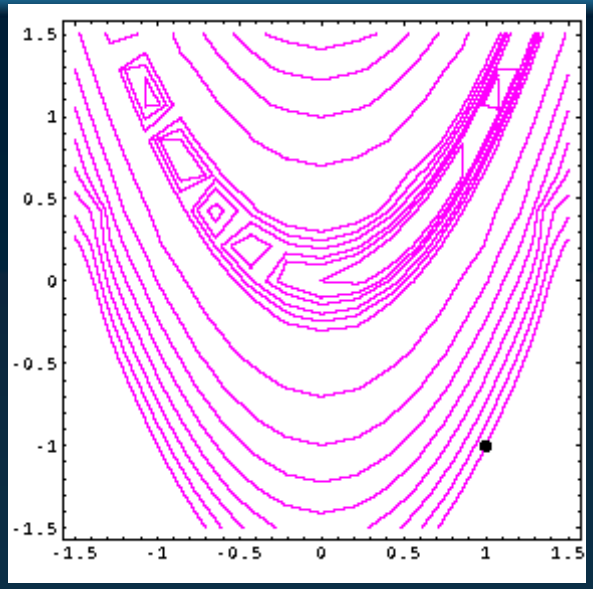


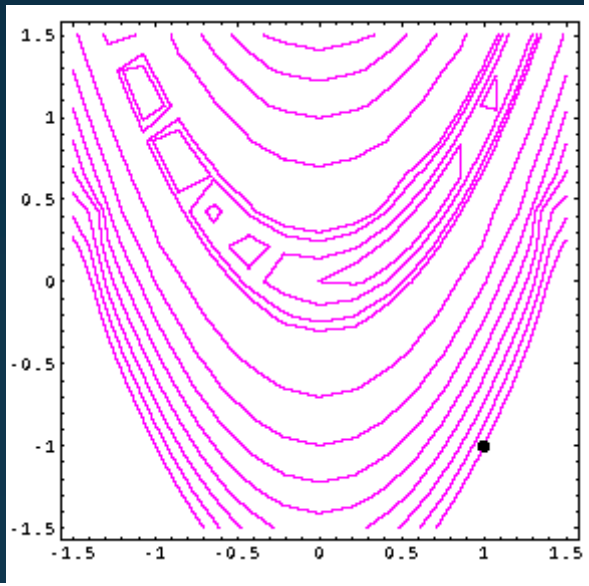
Deterministic Methods

- Conjugate Gradient Method:

Faster than the

Steepest-Descent!





Deterministic Methods

➔ Newton's Method:

Iterative process:

$$\mathbf{P}^{k+1} = \mathbf{P}^k + \alpha^k \mathbf{d}^k$$

$$\mathbf{d}^k = -\mathbf{H}^k \nabla S\left(\mathbf{P}^k\right)$$

where:

x is the vector of parameters to be optimized

α is the search-step size

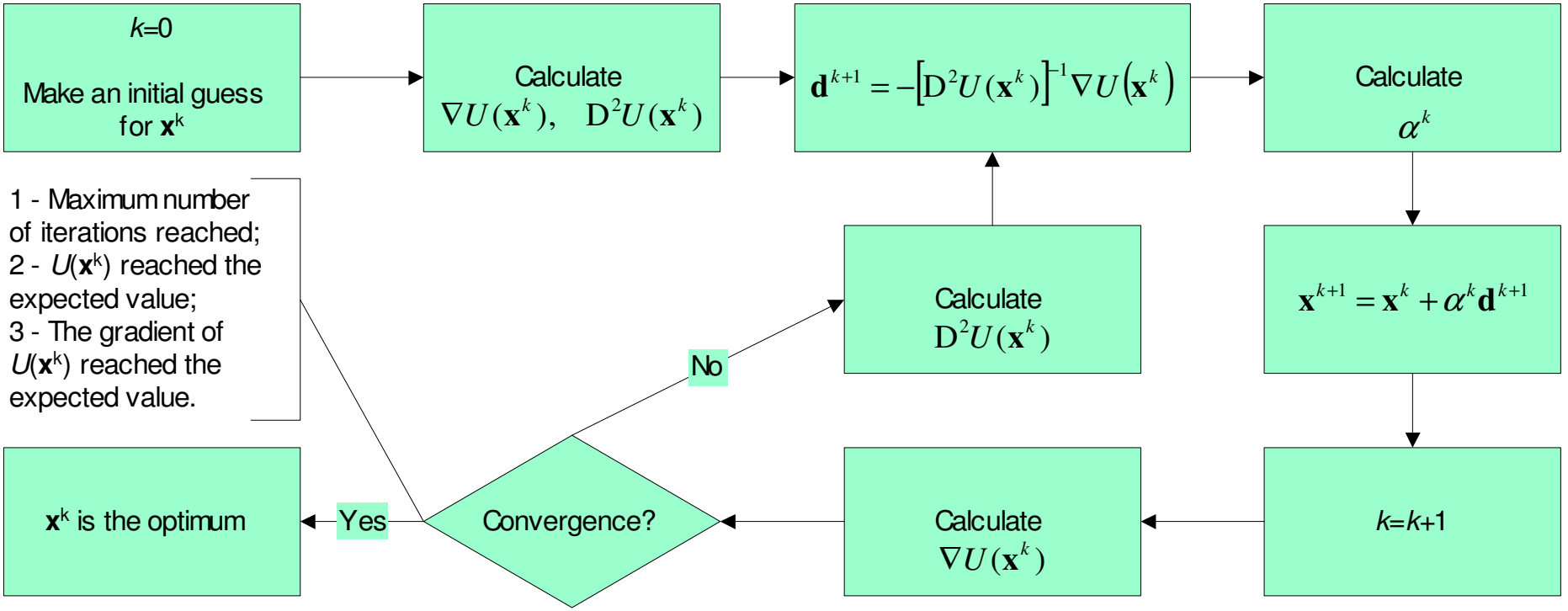
H is the matrix of 2nd order derivatives – Expensive in terms of computational cost!

S is the objective function

k is the iteration number

Deterministic Methods

- Newton Method:

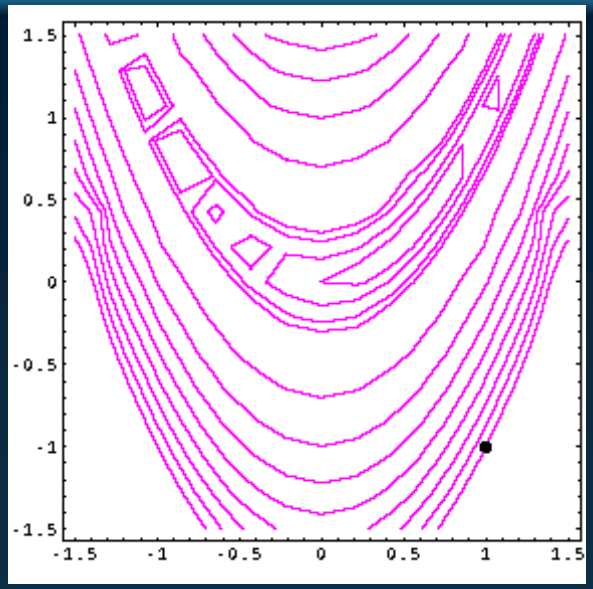


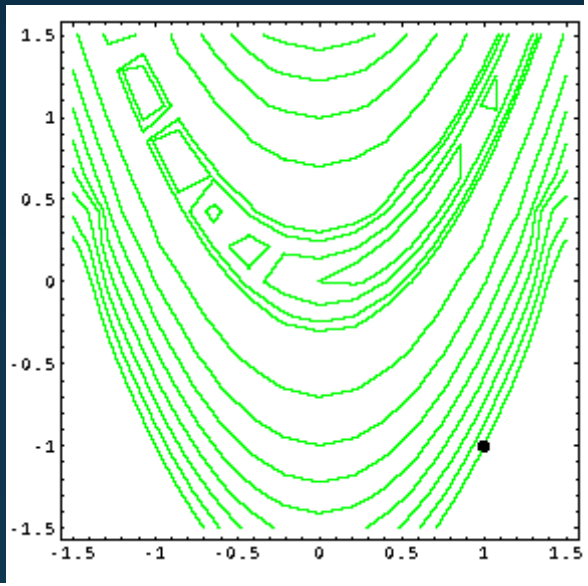
Deterministic Methods

- Newton's Method:

Faster than the

Conjugate Gradient!





Deterministic Methods

- BFGS (Broyden-Fletcher-Goldfarb-Shanno) Method:
  - Quasi-Newton method, similar to the DFP method, but less dependent on the search-step size choice.
  - Uses an iterative approximation for the Hessian

$$\mathbf{H}^{k} = \mathbf{H}^{k-1} + \mathbf{M}^{k-1} + \mathbf{N}^{k-1}$$

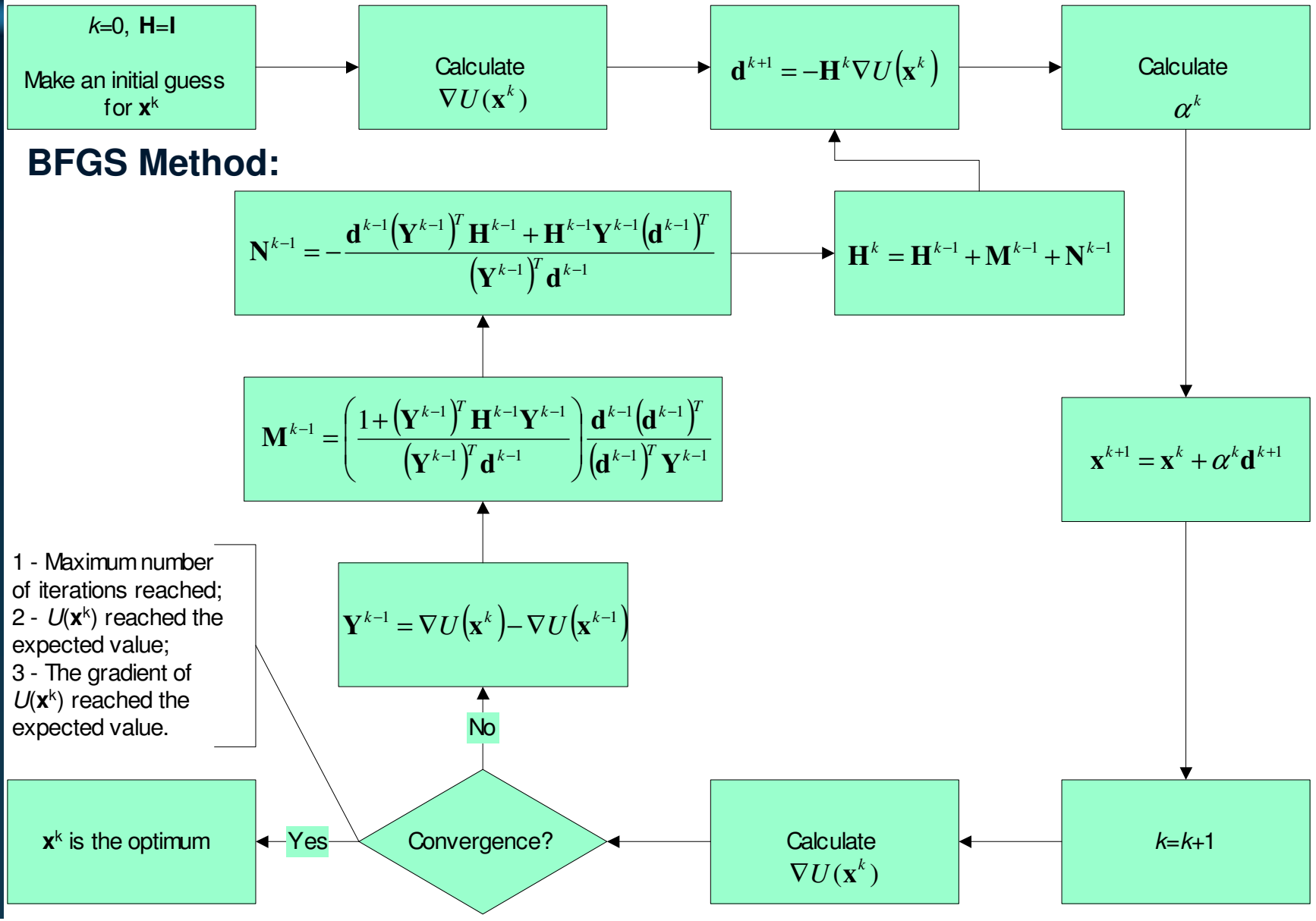
where

$$\mathbf{M}^{k-1} = \left(\frac{1 + \left(\mathbf{Y}^{k-1}\right)^{T} \mathbf{H}^{k-1} \mathbf{Y}^{k-1}}{\left(\mathbf{Y}^{k-1}\right)^{T} \mathbf{d}^{k-1}} \right) \frac{\mathbf{d}^{k-1} \left(\mathbf{d}^{k-1}\right)^{T}}{\left(\mathbf{d}^{k-1}\right)^{T} \mathbf{Y}^{k-1}}$$

$$\mathbf{N}^{k-1} = -\frac{\mathbf{d}^{k-1} \left(\mathbf{Y}^{k-1}\right)^{T} \mathbf{H}^{k-1} + \mathbf{H}^{k-1} \mathbf{Y}^{k-1} \left(\mathbf{d}^{k-1}\right)^{T}}{\left(\mathbf{Y}^{k-1}\right)^{T} \mathbf{d}^{k-1}}$$

$$\mathbf{Y}^{k-1} = \nabla S\left(\mathbf{P}^{k}\right) - \nabla S\left(\mathbf{P}^{k-1}\right)$$

Deterministic Methods



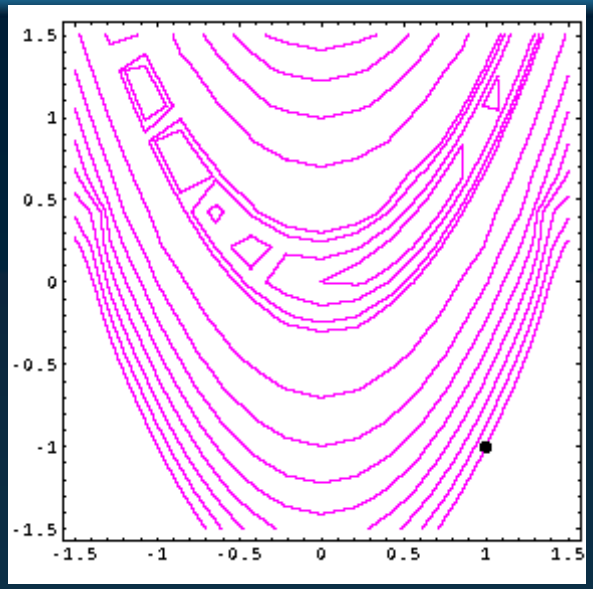
➔ **BFGS Method:**

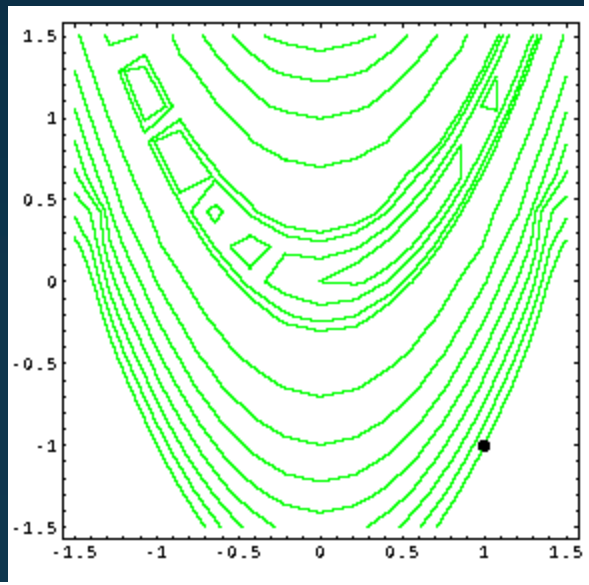
Deterministic Methods

- BFGS Method:

Faster than the

Conjugate Gradient!





Evolutionary and Stochastic Methods

- GA (Genetic Algorithm) and DE (Differential Evolution) Methods:
- Based on Darwin's model for natural selection of species.
  - Members of a certain population matches and have children. Those children are a combination of the parents' chromosomes.
  - The strongest members of the population will have more chances to survive under a certain environment.
  - The combination of the chromosomes is called crossover.
  - Mutations can occur. They can be good or bad mutations.

Evolutionary and Stochastic Methods

- DE Method:
  - Alternative to the Genetic Algorithm method.
  - Proposed in 1995 by Kenneth Price and Rainer Storn from Berkeley.
- The method initializes with a random generated random matrix **P** which contains N vector parameters **P**
- From the initial population matrix, generations are created until the best generation (optimum) is found.

Evolutionary and Stochastic Methods

- DE Method:
- The next generation is created as:

$$\mathbf{P}_i^{k+1} = \delta_1 \mathbf{P}_i^k + \delta_2 \left[\boldsymbol{\alpha} + F\left(\boldsymbol{\beta} - \boldsymbol{\gamma}\right) \right]$$

1st parent **2nd parent Mutation included**

where

α, β and γ are three randomly chosen members of the population matrix.

F is a weighting function which defines the mutation ($0.5 < F < 1$).

k is the generation counter.

δ_1 and δ_2 are delta Dirac functions that defines the crossover.

If $S(\mathbf{P}^{k+1}) < S(\mathbf{P}^k)$ ⇒ $\mathbf{P}^{k+1}$ replaces $\mathbf{P}^k$ in the population matrix

If $S(\mathbf{P}^{k+1}) > S(\mathbf{P}^k)$ ⇒ $\mathbf{P}^k$ is kept in the population matrix and $\mathbf{P}^{k+1}$ is discarded

Evolutionary and Stochastic Methods

- DE Method:
- The crossover is obtained as:

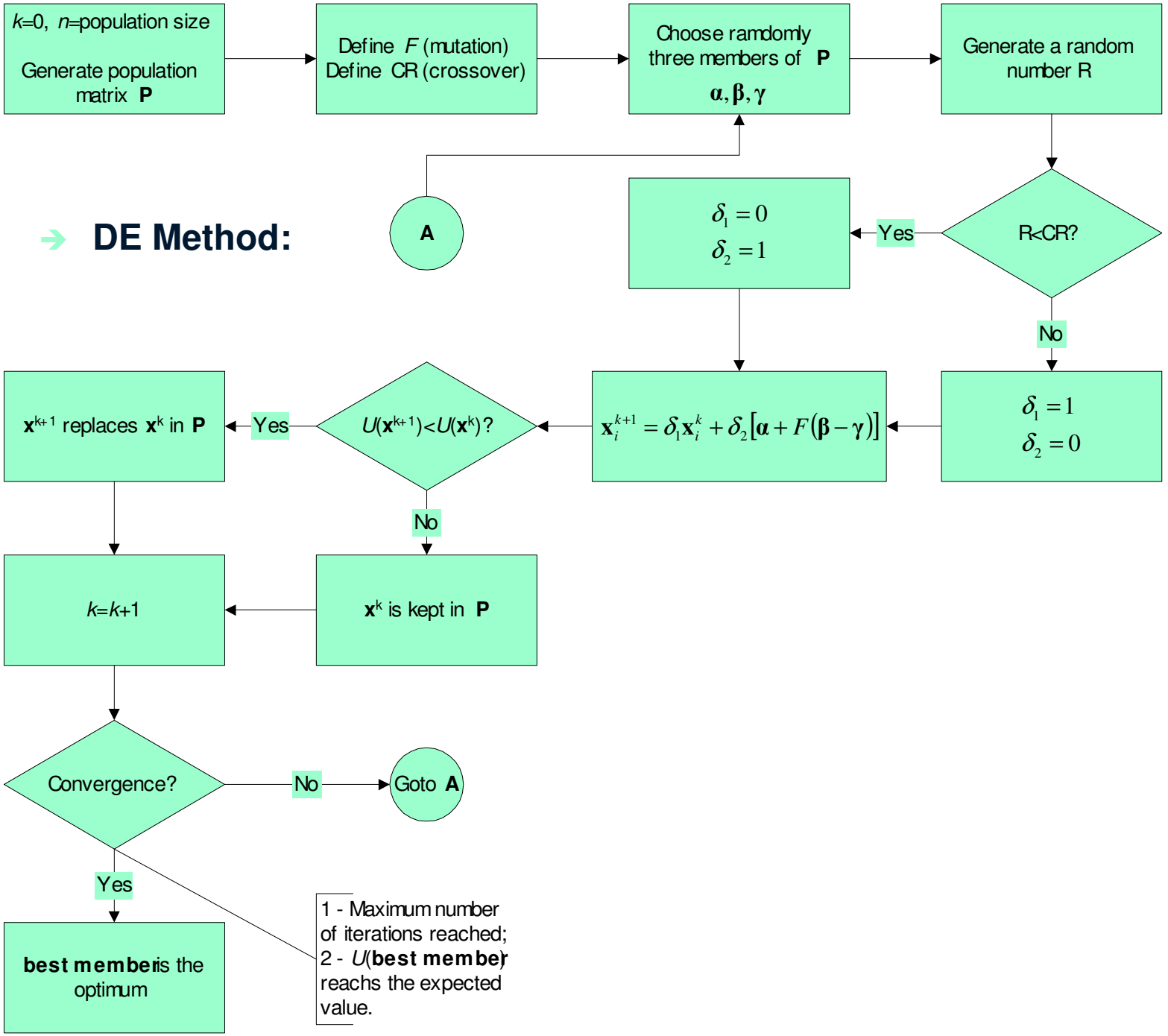
$$\mathbf{P}_i^{k+1} = \delta_1 \mathbf{P}_i^k + \delta_2 \left[\boldsymbol{\alpha} + F\left(\boldsymbol{\beta} - \boldsymbol{\gamma}\right) \right]$$

$\delta_1 =$ → 0, if R < CR; 1, if R > CR

$\delta_2 =$ → 1, if R < CR; 0, if R > CR

- R is a random number with uniform distribution between 0 and 1
- CR is the crossover factor (0.5 < CR < 1)

➔ DE Method:


k=0, n=population size
Generate population matrix P
Define F (mutation)
Define CR (crossover)
Choose ramdomly three members of P
α, β, γ
Generate a random number R
R<CR?
Yes
No
A
δ1 = 0
δ2 = 1
δ1 = 1
δ2 = 0
x_i^{k+1} = δ1 x_i^k + δ2 [α + F(β − γ)]
U(x^{k+1}) < U(x^k)?
Yes
No
x^{k+1} replaces x^k in P
x^k is kept in P
k=k+1
Convergence?
No
Goto A
Yes
best member is the optimum
1 - Maximum number of iterations reached;
2 - U(best member) reachs the expected value.

Evolutionary and Stochastic Methods

- SA (Simulated Annealing) Method:
- Based on thermodynamics and solidification process of liquids and metals.

Slow cooling ⇒ Pure crystal is formed with minimum energy state.

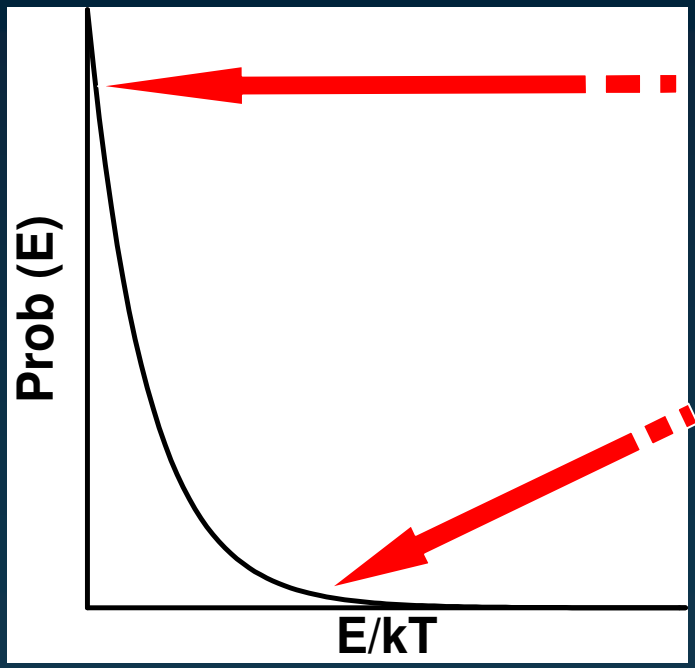
Fast cooling "quenched" ⇒ Polycrystalline or amorphous state is formed with higher energy.

- Gradient methods ⇒ "Fast cooling". They can lead to a local minima.

Evolutionary and Stochastic Methods

- SA Method:
- Boltzmann probability distribution:

$$\text{Prob}(E) \approx \exp\left(-\frac{E}{kT}\right)$$



High temperature

High Probability of high energy state

Low temperature

Small Probability of high energy state

- The method can move uphill as well as downhill depending on the probability of high energy states.

Evolutionary and Stochastic Methods

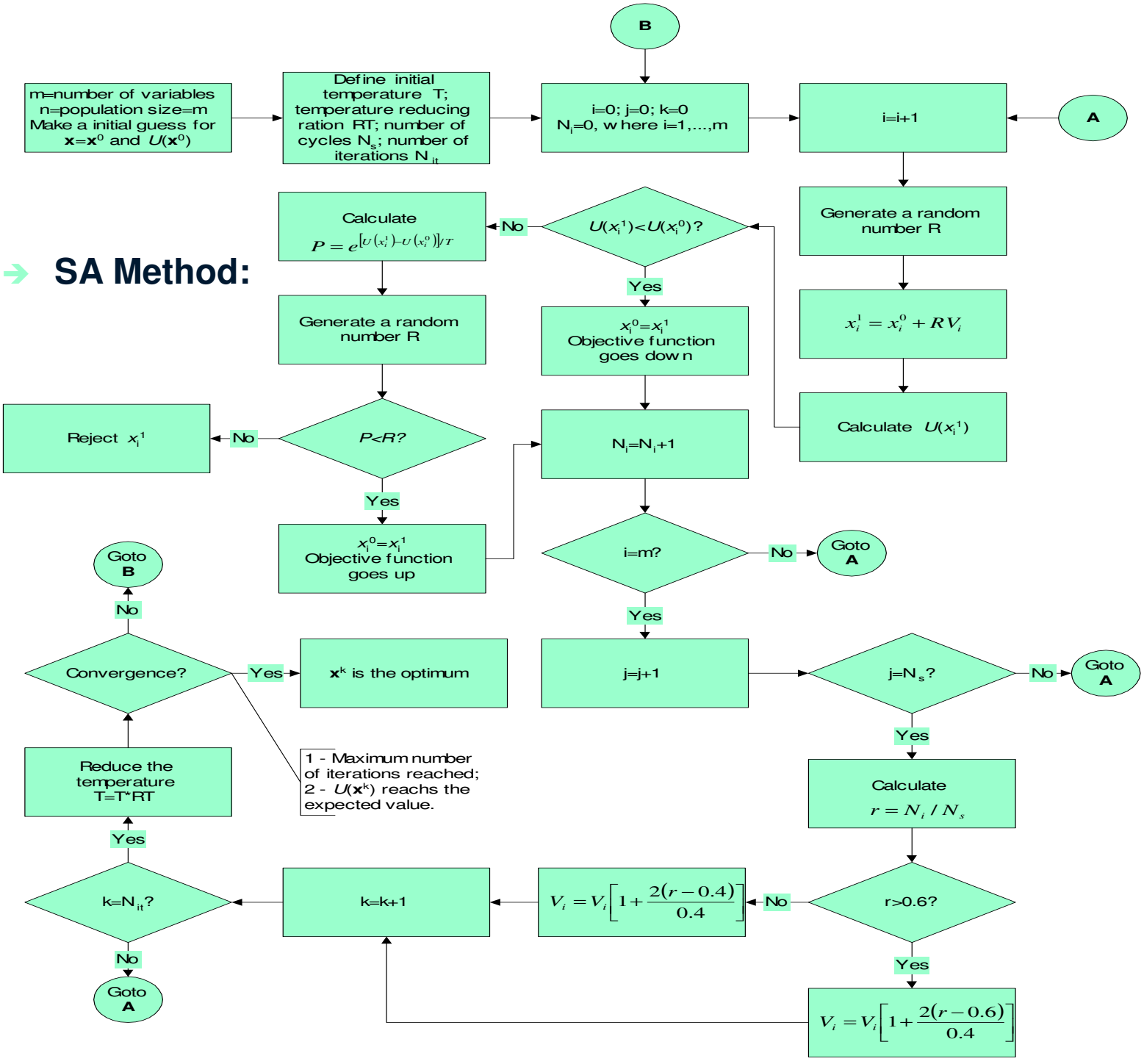
- SA Method:
- Iterative process is presented in the following papers:

  - Corana, A., Marchesi, M., Martini, C. e Ridella, S., “Minimizing Multimodal Functions of Continuous Variables with the ‘Simulated Annealing Algorithm’”, *ACM Transactions on Mathematical Software*, vol. 13, pp. 262-280, 1987.

  - Goffe, W. L., Ferrier, G. D. e Rogers, J., “Global Optimization of Statistical Functions with Simulated Annealing”, *Journal of Econometrics*, vol. 60, pp. 65-99, 1994.

- Excessive number of objective function evaluations!!!

SA Method:


B
m=number of variables
n=population size=m
Make a initial guess for
x=x⁰ and U(x⁰)
Define initial temperature T; temperature reducing ration RT; number of cycles N_s; number of iterations N_it
i=0; j=0; k=0
N_i=0, where i=1,...,m
i=i+1
A
Generate a random number R
x_i¹ = x_i⁰ + R V_i
Calculate U(x_i¹)
U(x_i¹)<U(x_i⁰)?
No
Yes
Calculate P = e^[(U(x_i¹)−U(x_i⁰))/T]
Generate a random number R
P<R?
No
Reject x_i¹
Yes
x_i⁰=x_i¹ Objective function goes down
N_i=N_i+1
x_i⁰=x_i¹ Objective function goes up
i=m?
No
Goto A
Yes
j=j+1
j=N_s?
No
Goto A
Yes
Calculate r = N_i / N_s
r>0.6?
No
V_i = V_i[1 + 2(r−0.4)/0.4]
Yes
V_i = V_i[1 + 2(r−0.6)/0.4]
k=k+1
k=N_it?
Yes
No
Goto A
Reduce the temperature T=T*RT
Convergence?
Yes
x^k is the optimum
No
Goto B
1 - Maximum number of iterations reached;
2 - U(x^k) reachs the expected value.

Evolutionary and Stochastic Methods

- PS (Particle Swarm) method:
  - Created in 1995 by an Electric Engineer (Russel Eberhart) and a Social-Psychologist (James Kennedy) as an alternative to Genetic Algorithm.
  - Based on the social behavior of various species (including humans).
  - Balances the individuality and sociability of individuals in order to find a optimum.

↑ Individuality → ↑ Chances to find alternatives places

↓ Convergence

↑ Sociability → ↑ Learning process among the individuals

↓ Chances to find alternatives places. Individuals can find a local minima

Evolutionary and Stochastic Methods

- PS method:
- Update process

$$\mathbf{P}_i^{k+1} = \mathbf{P}_i^k + \mathbf{v}_i^{k+1}$$

$$\mathbf{v}_i^{k+1} = \alpha \mathbf{v}_i^k + \beta \mathbf{r}_{1i}\left(\mathbf{p}_i - \mathbf{P}_i^k\right) + \beta \mathbf{r}_{2i}\left(\mathbf{p}_g - \mathbf{P}_i^k\right)$$

Individuality **Sociability**

where

$\mathbf{P}_i$ is i-th individual of the vector of parameters

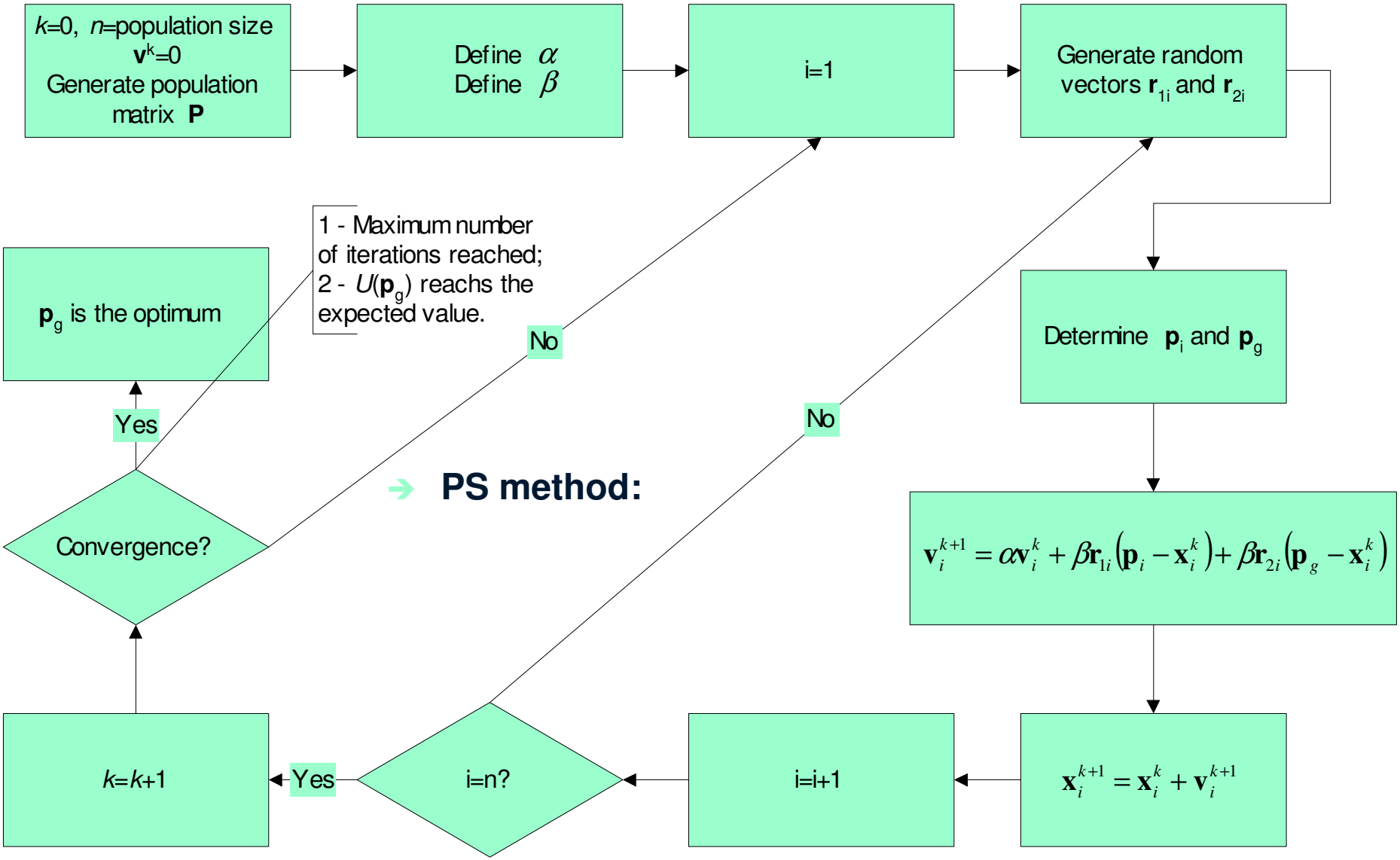
$\mathbf{r}_{1i}$ and $\mathbf{r}_{2i}$ are are random numbers with uniform distribution between 0 and 1

$\mathbf{p}_i$ is the best value found for the vector $\mathbf{P}_i$

$\mathbf{p}_g$ is the best value found for the entire population

$0 < \alpha < 1;\ 1 < \beta < 2$

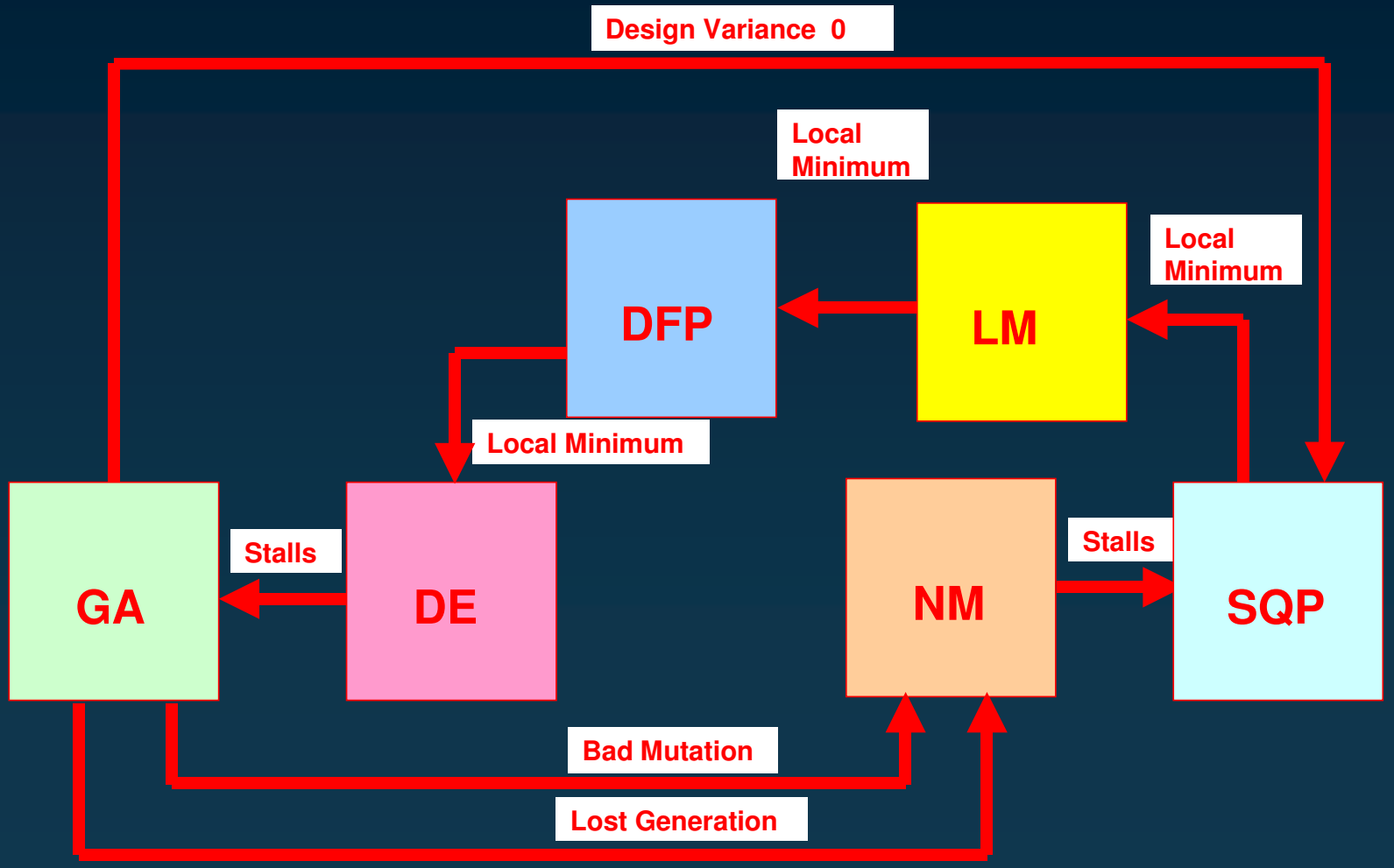
Evolutionary and Stochastic Methods



➔ **PS method:**

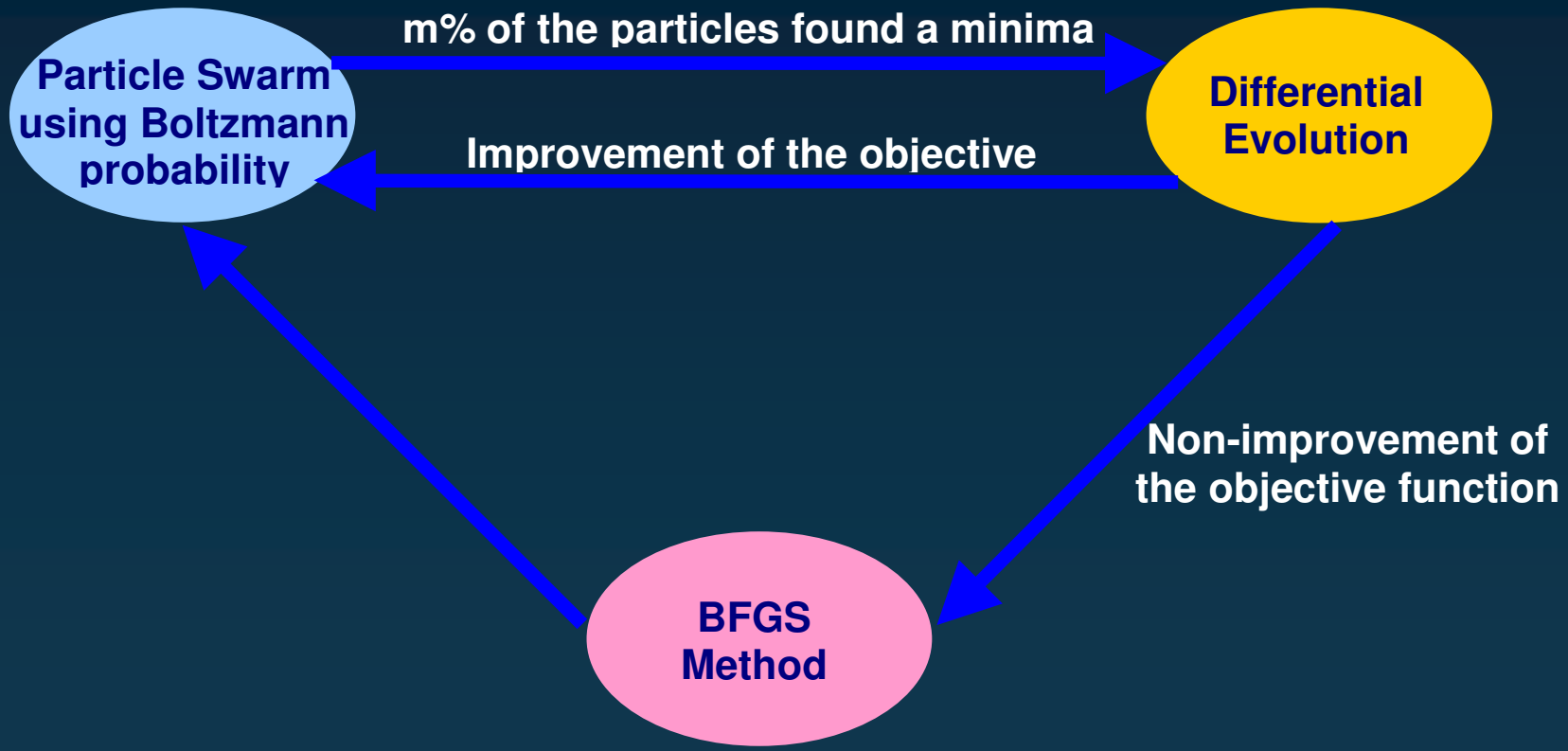
Hybrid Methods

- Hybrid optimizer – version 1 (Martin, Colaço and Dulikravich)



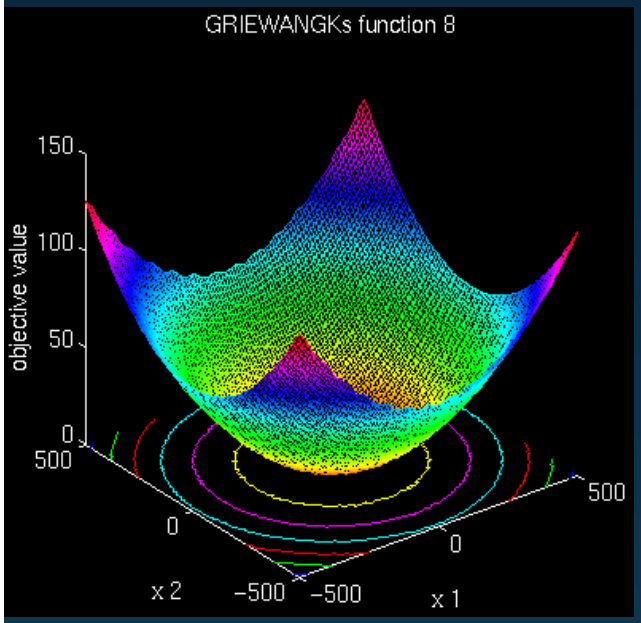
Hybrid Methods

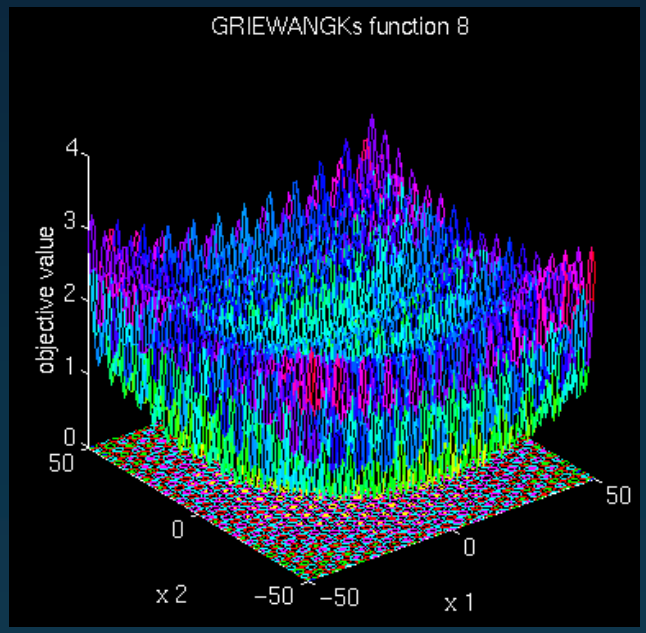
- Hybrid optimizer – version 2 (Colaço and Dulikravich)

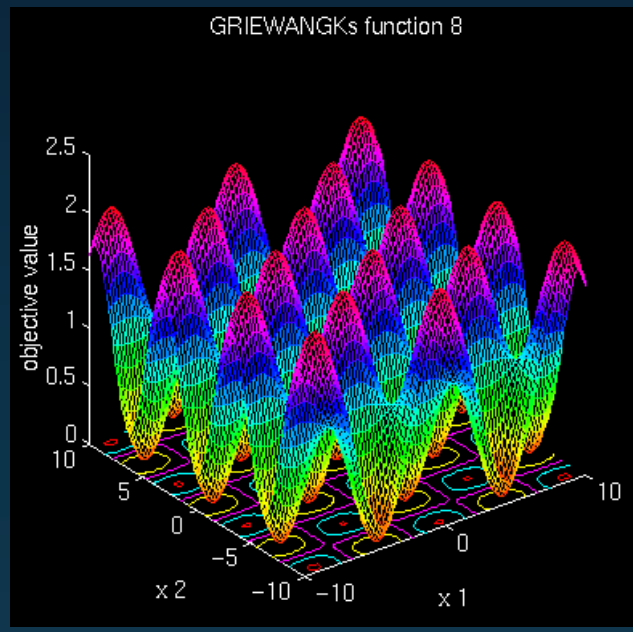


Example 1 - Griewangk's function

- Multiple local minima

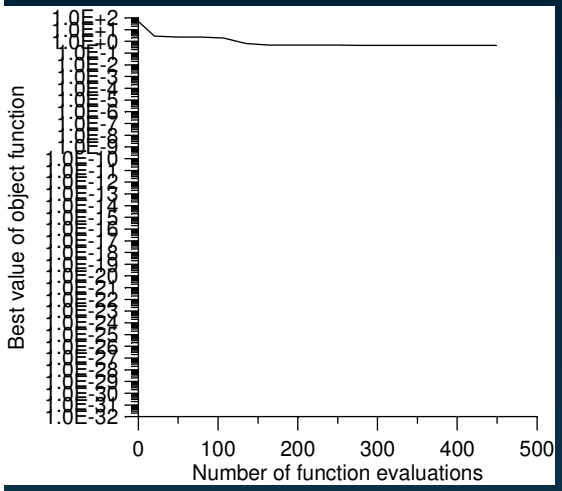


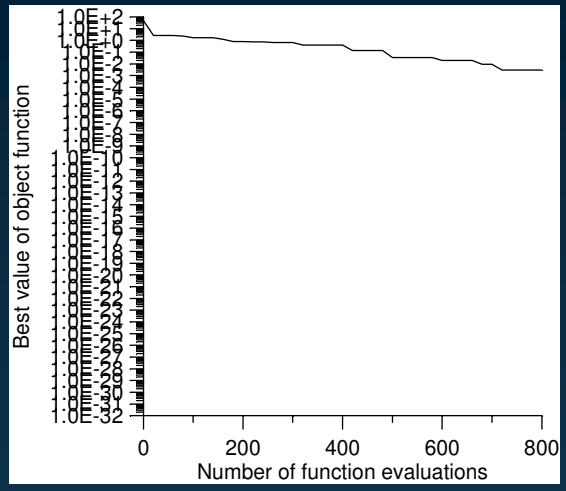


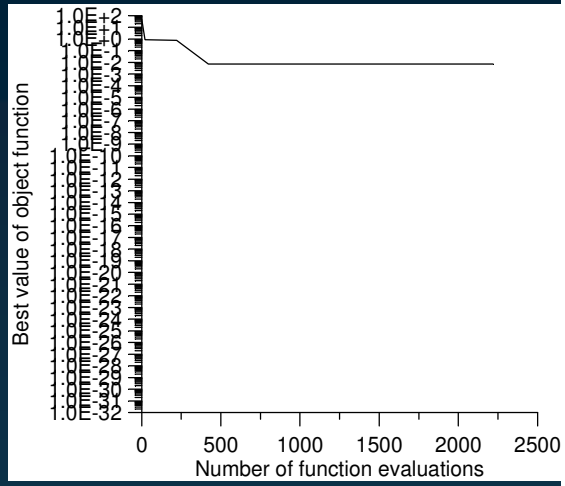


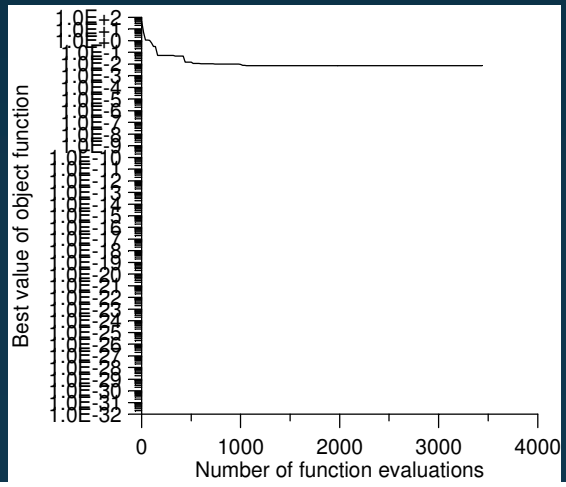
Example 1 - Griewangk's function

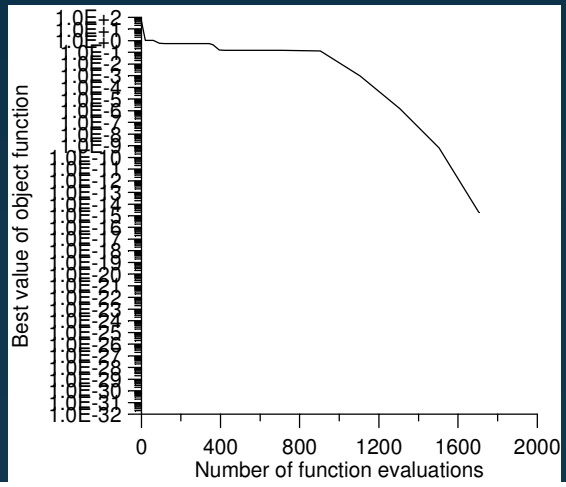
- Comparison: BFGS, DE, SA, PS, Hybrid





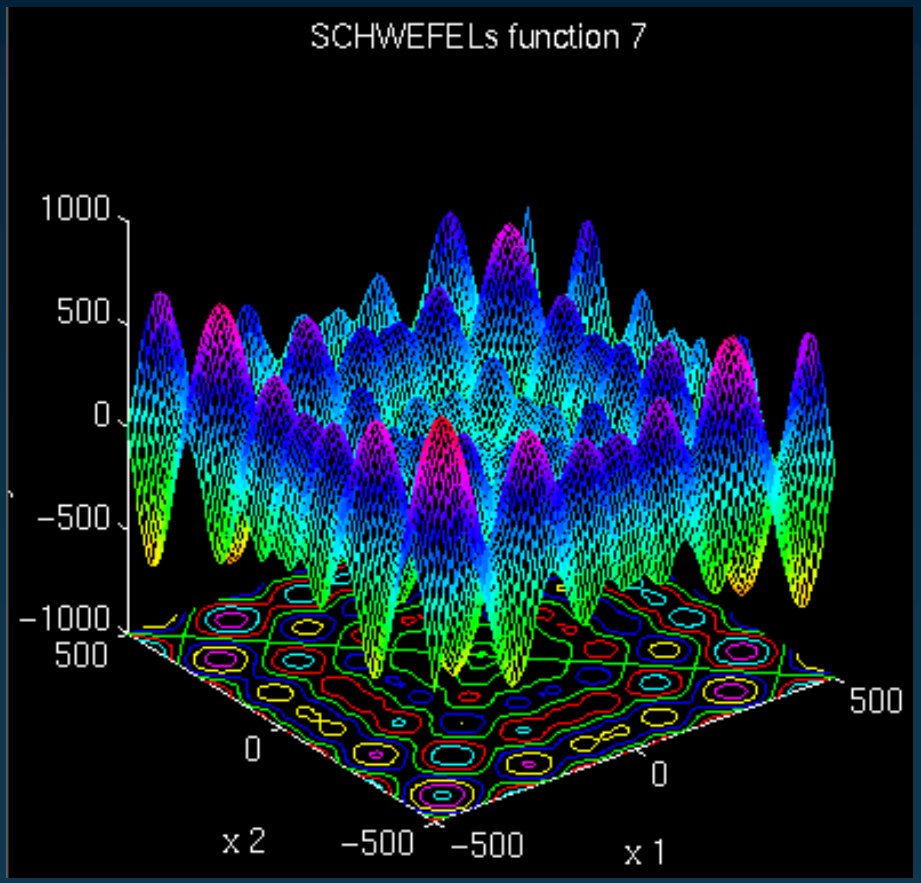






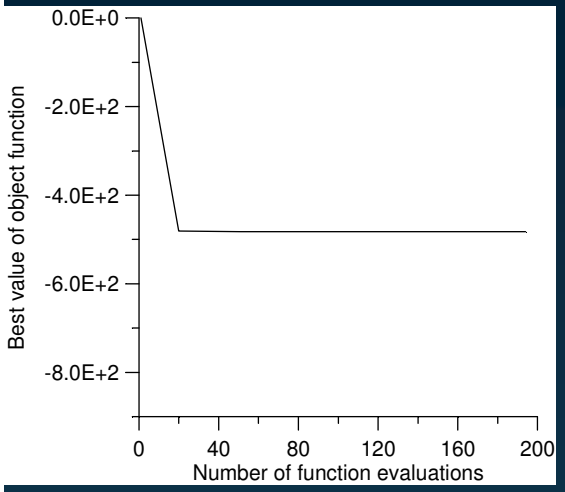
Example 2 - Schwefel's function

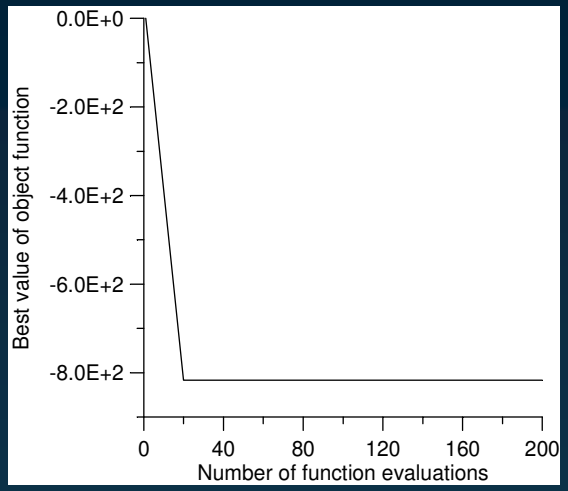
- Multiple local minima

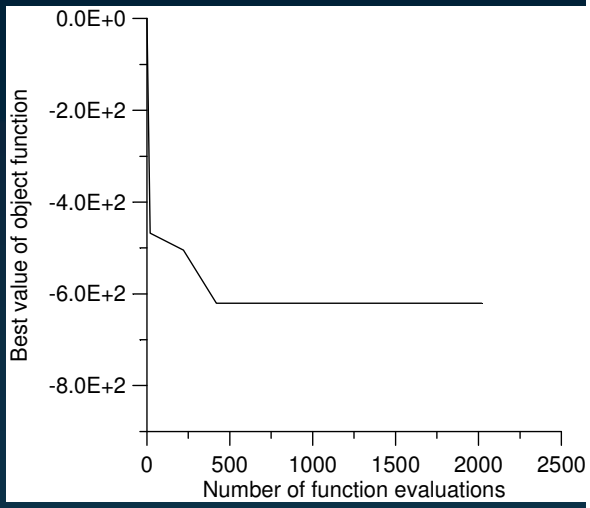


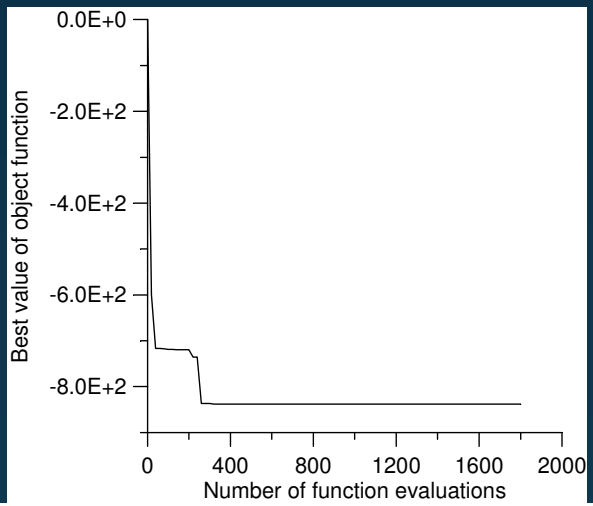
Example 2 - Schwefel's function

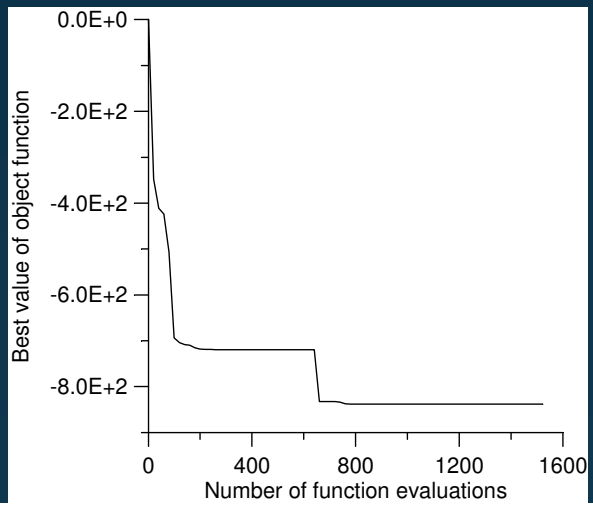
- Comparison: BFGS, DE, SA, PS, Hybrid











Statistical Analysis

Covariance Matrix

$$\mathbf{V} \equiv \begin{bmatrix} \mathrm{cov}(\hat{P}_1,\hat{P}_1) & \mathrm{cov}(\hat{P}_1,\hat{P}_2) & \ldots & \mathrm{cov}(\hat{P}_1,\hat{P}_N) \\ \mathrm{cov}(\hat{P}_2,\hat{P}_1) & \mathrm{cov}(\hat{P}_2,\hat{P}_2) & \ldots & \mathrm{cov}(\hat{P}_2,\hat{P}_N) \\ . & . & . & . \\ . & . & . & . \\ \mathrm{cov}(\hat{P}_N,\hat{P}_1) & \mathrm{cov}(\hat{P}_N,\hat{P}_2) & \ldots & \mathrm{cov}(\hat{P}_N,\hat{P}_N) \end{bmatrix} = (\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1}$$

99% Confidence Intervals ⇒ $P_j \pm 2.576\,\sigma_j$

Confidence Region ⇒ $(\hat{\mathbf{P}} - \mathbf{P})^T \mathbf{V}^{-1} (\hat{\mathbf{P}} - \mathbf{P}) \leq \chi^2_N$

where χ^2_N is the chi-square distribution with *N* degrees of freedom.

Design Of Optimum Experiments

Minimization of the Confidence Region

Maximize det ($\mathbf{J}^T\mathbf{J}$)

By using the definition of the sensitivity matrix for the case involving a single sensor, each element $\mathbf{F}_{m,n}$, $m,n = 1, \ldots, N$, of the matrix is given by $\mathbf{F} \equiv \mathbf{J}^T\mathbf{J}$

$$\mathbf{F}_{m,n} \equiv \left[\mathbf{J}^T\mathbf{J}\right]_{m,n} = \sum_{i=1}^{I} \left(\frac{\partial T_i}{\partial P_m}\right)\left(\frac{\partial T_i}{\partial P_n}\right)$$

where I is the number of measurements and N is the number of unknown parameters.

Design Of Optimum Experiments
Special Cases

Case 1. A large but fixed number of equally spaced measurements is available.

Then, each element $\mathbf{F}_{m,n}$ can be written as

$$\mathbf{F}_{m,n} = \frac{1}{\Delta t}\sum_{i=1}^{I}\left(\frac{\partial T_i}{\partial P_m}\right)\left(\frac{\partial T_i}{\partial P_n}\right)\Delta t \approx \frac{I}{t_f}\int_{t=0}^{t_f}\left(\frac{\partial T}{\partial P_m}\right)\left(\frac{\partial T}{\partial P_n}\right)dt \qquad \text{for } m,n = 1, \ldots, N$$

where t_f is the duration of the experiment and Δt is the constant time interval between two consecutive measurements. Since the number of measurements, I, is fixed, we can choose to maximize the determinant of $\mathbf{F}_I$ instead of maximizing the determinant of $\mathbf{F}$, where the elements of $\mathbf{F}_I$ are given by

$$[\mathbf{F}_I]_{m,n} = \frac{1}{t_f}\int_{t=0}^{t_f}\left(\frac{\partial T}{\partial P_m}\right)\left(\frac{\partial T}{\partial P_n}\right)dt \qquad \text{for } m,n = 1, \ldots, N$$

Design Of Optimum Experiments
Special Cases

Case 2. In addition to a large and fixed number of equally spaced measurements, the maximum value for the temperature in the region, T_{max}, is known.

$$[\mathbf{F}_I]_{m,n} = \frac{T_{max}^2}{t_f P_m P_n} \int_{t=0}^{t_f} \left(\frac{P_m}{T_{max}} \frac{\partial T}{\partial P_m} \right) \left(\frac{P_n}{T_{max}} \frac{\partial T}{\partial P_n} \right) dt \qquad \text{for } m,n = 1,\ldots,N$$

Note that the quantities inside parentheses are dimensionless. However, it is possible that T^*, and not T_{max}, is the variable suitable for the non-dimensionalization of the temperature T, i.e.,

$$[\mathbf{F}_I]_{m,n} = \frac{T_{max}^2}{t_f P_m P_n} \int_{t=0}^{t_f} \left(\frac{P_m}{T^*} \frac{\partial T}{\partial P_m} \right) \left(\frac{P_n}{T^*} \frac{\partial T}{\partial P_n} \right) \left(\frac{T^*}{T_{max}} \right)^2 dt \qquad \text{for } m,n = 1, \ldots, N$$

Design Of Optimum Experiments
Special Cases

and the design of optimum experiments is then based on the maximization of the determinant of the *dimensionless form* of $\mathbf{F}_I$:

$$[\mathbf{F}_I^*]_{m,n} = \frac{1}{t_f}\int_{t=0}^{t_f}\left(\frac{P_m}{T^*}\frac{\partial T}{\partial P_m}\right)\left(\frac{P_n}{T^*}\frac{\partial T}{\partial P_n}\right)\left(\frac{T^*}{T_{max}}\right)^2 dt$$

for $m,n = 1, \ldots, N$

Case 3. Measurements of *M* sensors are available.

$$[\mathbf{F}_I^*]_{m,n} = \frac{1}{M\,t_f}\sum_{s=1}^{M}\int_{t=0}^{t_f}\left(\frac{P_m}{T^*}\frac{\partial T_s}{\partial P_m}\right)\left(\frac{P_n}{T^*}\frac{\partial T_s}{\partial P_n}\right)\left(\frac{T^*}{T_{max}}\right)^2 dt$$

for $m,n = 1, \ldots, N$

Sequential Parameter Estimation Technique

Maximum a Posteriori Objective Function

$$S_{MAP}(\mathbf{P}) = [\mathbf{Y} - \mathbf{T}(\mathbf{P})]^T \mathbf{W} [\mathbf{Y} - \mathbf{T}(\mathbf{P})] + (\boldsymbol{\mu} - \mathbf{P})^T \mathbf{V}^{-1} (\boldsymbol{\mu} - \mathbf{P})$$

Hypotheses:

- The errors are additive, with zero mean and normally distributed.
- The statistical parameters describing the errors are known.
- There are no errors in the independent variables.
- **P** is a random vector with known mean **m** and known covariance matrix **V**.
- **P** is distributed normally.

Sequential Parameter Estimation Technique

For uncorrelated measurements:

$$\mathbf{W} = \begin{bmatrix} 1/\sigma_1^2 & & & 0 \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ 0 & & & 1/\sigma_I^2 \end{bmatrix}$$

For the minimization of $S_{MAP}(\mathbf{P})$:

$$\frac{\partial S_{MAP}(\mathbf{P})}{\partial P_1} = \frac{\partial S_{MAP}(\mathbf{P})}{\partial P_2} = \cdots = \frac{\partial S_{MAP}(\mathbf{P})}{\partial P_N} = 0$$

$$-2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P})] - 2\mathbf{V}^{-1}[\boldsymbol{\mu}-\mathbf{P}] = 0$$

where **J** is the *Sensitivity Matrix*.

Sequential Parameter Estimation Technique

$$-2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P})]-2\mathbf{V}^{-1}[\boldsymbol{\mu}-\mathbf{P}]=0$$

Linear Problems: **J** does not depend on **P** ⇒ $\mathbf{T}(\mathbf{P})=\mathbf{J}\,\mathbf{P}$

$$\mathbf{P}=[\mathbf{J}^T\mathbf{W}\,\mathbf{J}+\mathbf{V}^{-1}]^{-1}[\mathbf{J}^T\mathbf{W}\,\mathbf{Y}+\mathbf{V}^{-1}\boldsymbol{\mu}]$$

Nonlinear Problems: $\mathbf{J}\equiv\mathbf{J}(\mathbf{P})$ ⇒ $\mathbf{T}(\mathbf{P})=\mathbf{T}(\mathbf{P}^k)+\mathbf{J}^k(\mathbf{P}-\mathbf{P}^k)$

$$\mathbf{P}^{k+1}=\mathbf{P}^k+[\mathbf{J}^T\mathbf{W}\mathbf{J}+\mathbf{V}^{-1}]^{-1}\{\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P}^k)]+\mathbf{V}^{-1}(\boldsymbol{\mu}-\mathbf{P}^k)\}$$

Sequential Parameter Estimation Technique

- Utilizes the measurements in a sequential manner in order to estimate the parameters.
- Avoids matrix inversions.
- Permits the identification of improper mathematical models.
- Possible to identify if a sufficient number of transient measurements and if a sufficiently long experimental time have been used in the experiment.

Computational Algorithm For The Nonlinear Case

Step 1. Initialize the iterative procedure by setting the iteration index k to 0 and making $\mathbf{P}^0 = \boldsymbol{\mu}_m$.

Step 2. Compute the estimate for the vector of unknown parameters sequentially, for $i=0,\ldots,(I-1)$, by using

$$\mathbf{A} = \mathbf{V}_i \mathbf{J}_{i+1}^T$$

$$\Delta = \mathbf{J}_{i+1} \mathbf{A} + W_{i+1}^{-1}$$

$$\mathbf{K} = \mathbf{A}\Delta^{-1}$$

$$E_{i+1} = Y_{i+1} - T_{i+1}(\mathbf{P}^k)$$

$$\mathbf{P}_{i+1}^{k+1} = \mathbf{P}_i^{k+1} + \mathbf{K}[E_{i+1} - \mathbf{J}_{i+1}(\mathbf{P}_i^{k+1} - \mathbf{P}^k)]$$

$$\mathbf{V}_{i+1} = \mathbf{V}_i - \mathbf{K}\mathbf{J}_{i+1}\mathbf{V}_i$$

Computational Algorithm For The Nonlinear Case

Step 3. Check convergence of the values estimated sequentially with all I measurements

$$\left\| \mathbf{P}_I^{k+1} - \mathbf{P}_I^k \right\| < \varepsilon$$

If the convergence criterion is not satisfied, increment k, make

$$\mathbf{P}^k = \mathbf{P}_I^k$$

and return to step 2.

Function Estimation

CONJUGATE GRADIENT METHOD WITH ADJOINT PROBLEM

1. Direct Problem
2. Inverse Problem
3. Sensitivity Problem
4. Adjoint Problem
5. Gradient Equation
6. Iterative Procedure
7. Stopping Criterion
8. Computational Algorithm

Function Estimation

- No assumption is made regarding the functional form of the unknown.

Hilbert Space of square-integrable functions in $0 < t < t_f$

$$\int_{t=0}^{t_f} [g_p(t)]^2 dt < \infty$$

Minimize:

$$S[g_p(t)] = \int_{t=0}^{t_f} \{Y(t) - T[x_{meas}, t; g_p(t)]\}^2 dt$$

where the standard-deviation of the measurement errors was assumed as constant and known.

Conjugate Gradient Method

- Iterative method
- Simple, straightforward and powerful
- Regularization obtained from the stopping criterion
  (Iterative regularization)
- Parameter estimation and function estimation
- Linear and nonlinear inverse problems
- Capable of treating complex physical situations
- Stable for small time steps
- Easy to program
- Permits the extension to more than one unknown
- Whole-domain

Conjugate Gradient Method (Fletcher-Reeves)

Iterative Procedure:

$$g_p^{k+1}(t) = g_p^k(t) - \beta^k d^k(t)$$

Direction of Descent:

$$d^k(t) = \nabla S[g_p^k(t)] + \gamma^k d^{k-1}(t)$$

Conjugation Coefficient:

$$\gamma^k = \frac{\int_{t=0}^{t_f} \{\nabla S[g_p^k(t)]\}^2 dt}{\int_{t=0}^{t_f} \{\nabla S[g_p^{k-1}(t)]\}^2 dt}$$

with $\gamma^0 = 0$ for $k = 0$

Search step size β^k

and

Gradient Direction $\nabla S[g_p^k(t)]$

→ *Sensitivity Problem and Adjoint Problem*

Sensitivity Problem

$g_p(t)$ undergoes a variation $\varepsilon\Delta g_p(t)$

$T(x,t)$ is perturbed by $\varepsilon\Delta T(x,t)$

$$D_{\Delta g_p} T(x,t) = \lim_{\varepsilon \to 0} \frac{L_\varepsilon(g_{p\varepsilon}) - L(g_p)}{\varepsilon} = 0$$

where $L_\varepsilon(g_{p\varepsilon})$ and $L(g_p)$ are the operator forms of the direct problem, written for the perturbed $[g_p(t) + \varepsilon\Delta\, g_p(t)]$ and unperturbed $g_p(t)$ source-terms, respectively.

Sensitivity Problem

$$\frac{\partial^2 \Delta T(x,t)}{\partial x^2} + \Delta g_p(t)\,\delta(x-0.5) = \frac{\partial \Delta T(x,t)}{\partial t} \qquad \text{in } 0 < x < 1, \text{ for } t > 0$$

$$\frac{\partial \Delta T(0,t)}{\partial x} = 0 \qquad \text{at } x = 0, \text{ for } t > 0$$

$$\frac{\partial \Delta T(1,t)}{\partial x} = 0 \qquad \text{at } x = 1, \text{ for } t > 0$$

$$\Delta T(x, 0) = 0 \qquad \text{for } t = 0 \text{ , in } 0 < x < 1$$

Search Step Size

$$\min_{\beta^k} S[g_p^{k+1}(t)] = \min_{\beta^k} \int_{t=0}^{t_f} \left\{ Y(t) - T\left[x_{meas}, t; g_p^k(t) - \beta^k d^k(t) \right] \right\}^2 dt$$

Remark: Different numerical techniques can be used to compute β^k. Alternatively, the expression above can be linearized before the minimization is performed in order to obtain a closed-form expression for β^k.

Search Step Size

$$\min_{\beta^k} S[g_p^{k+1}(t)] = \min_{\beta^k} \int_{t=0}^{t_f} \left\{ Y(t) - T\left[x_{meas}, t; g_p^k(t) - \beta^k d^k(t) \right] \right\}^2 dt$$

By linearizing $T[g_p^k(t) - \beta^k d^k(t)]$ and making $d^k(t) = \Delta g_p^k(t)$

we obtain $$T[g_p^k(t) - \beta^k d^k(t)] \approx T[g_p^k(t)] - \beta^k \frac{\partial T}{\partial g_p^k} \Delta g_p^k(t)$$

Let $$\Delta T[d^k(t)] = \frac{\partial T}{\partial g_p^k} \Delta g_p^k(t)$$

Thus $$T[g_p^k(t) - \beta^k d^k(t)] \approx T[g_p^k(t)] - \beta^k \Delta T[d^k(t)]$$

Search Step Size

$$\min_{\beta^k} S[g_p^{k+1}(t)] = \min_{\beta^k} \int_{t=0}^{t_f} \{Y(t) - T[x_{meas}, t; g_p^k(t)] + \beta^k \Delta T[x_{meas}, t; d^k(t)]\}^2 dt$$

$$\beta^k = \frac{\int_{t=0}^{t_f} \left\{ T\left[x_{meas}, t; g_p^k(t) \right] - Y(t) \right\} \Delta T\left[x_{meas}, t; d^k(t) \right] dt}{\int_{t=0}^{t_f} \left\{ \Delta T\left[x_{meas}, t; d^k(t) \right] \right\}^2 dt}$$

Adjoint Problem

Lagrange Multiplier:
$\lambda(x,t)$

AUGMENTED FUNCTIONAL

$$S[g_p(t)] = \int_{t=0}^{t_f} \{Y(t) - T[x_{meas}, t; g_p(t)]\}^2 dt +$$

$$\int_{x=0}^{1} \int_{t=0}^{t_f} \lambda(x,t) \left[\frac{\partial^2 T}{\partial x^2} + g_p(t)\,\delta(x-0.5) - \frac{\partial T}{\partial t} \right] dt\,dx$$

$$D_{\Delta g_p} S[g_p(t)] = \lim_{\varepsilon \to 0} \frac{S_\varepsilon[g_{p\varepsilon}(t)] - S[g_p(t)]}{\varepsilon} = 0$$

Adjoint Problem

$$D_{\Delta g_p} S[g_p(t)] = \int_{t=0}^{t_f} \int_{x=0}^{1} 2\left\{T\left[x,t;g_p(t)\right]-Y(t)\right\}\Delta T(x,t)\,\delta(x-x_{meas})\,dx\,dt$$

$$+\int_{t=0}^{t_f} \int_{x=0}^{1} \lambda(x,t)\left[\frac{\partial^2 \Delta T}{\partial x^2}+\Delta g_p(t)\delta(x-0.5)-\frac{\partial \Delta T}{\partial t}\right] dx\,dt$$

Remark: The second integral term on the right-hand side is simplified with integration by parts and by utilizing the boundary and initial conditions of the sensitivity problem. The integral terms containing $\Delta T(x,t)$ in the resulting expression are then allowed to go to zero, in order to obtain the *adjoint problem* for the determination of the Lagrange Multiplier $\lambda(x,t)$.

Adjoint Problem

$$\frac{\partial \lambda(x,t)}{\partial t} + \frac{\partial^2 \lambda(x,t)}{\partial x^2} + 2\,[T(x,t) - Y(t)]\,\delta(x - x_{meas}) = 0$$

in $0 < x < 1$, for $0 < t < t_f$

$$\frac{\partial \lambda(0,t)}{\partial x} = 0$$

at $x = 0$, for $0 < t < t_f$

$$\frac{\partial \lambda(1,t)}{\partial x} = 0$$

at $x = 1$, for $0 < t < t_f$

$$\lambda(x,t_f) = 0$$

for $t = t_f$, in $0 < x < 1$

Adjoint Problem

Remark: Note that in the adjoint problem the value of the function $\lambda(x,t)$ is given at the final time $t = t_f$. In the conventional *initial value problem*, the value of the function is specified at time $t = 0$. However, the *final value problem* can be transformed into an *initial value problem* by defining a new time variable given by $\tau = t_f - t$.

Gradient Equation

In the process used to obtain the adjoint problem the following integral term is left:

$$D_{\Delta g_p} S[g_p(t)] = \int_{t=0}^{t_f} \lambda(0.5,t)\,\Delta g_p(t)\,dt \to 0$$

Since $g_p(t)$ belongs to the space of square-integrable functions in $0 < t < t_f$:

$$D_{\Delta g_p} S[g_p(t)] = \int_{t=0}^{t_f} \nabla S[g_p(t)]\,\Delta g_p(t)\,dt$$

$$\nabla S[g_p(t)] = \lambda(0.5,t)$$

Gradient Equation

Remark: Note that the gradient equation is null at the final time t_f. Therefore, the initial guess used for $g_p(t)$ at $t= t_f$ is never changed by the iterative procedure of the conjugate gradient method for function estimation. The estimated function can deviate from the exact solution in a neighborhood of t_f, if the initial guess used is too different from the exact $g_p(t_f)$. This apparent drawback of the method can be easily overcome by using a final time larger than that of interest, so that the effects of the initial guess are not noticeable in the time interval that the solution is sought. Another approach to overcome this difficulty is to repeat the solution of the inverse problem, by using as initial guess a previously estimated value for $g_p(t)$ in the neighborhood of t_f.

Stopping Criterion

$$S[g_p(t)] < \varepsilon$$

Errorless Measurements: ε is a small specified number

Measurements Containing Random Errors:

Discrepancy Principle ⇒ **Iterative Regularization**

$$\left| Y(t) - T[x_{meas}, t; g_p(t)] \right| \approx \sigma$$

where σ = standard-deviation of the measurements

$$\varepsilon = \sigma^2 t_f$$

Computational Algorithm

Suppose an initial guess $g_p^0(t)$ is available for the function $g_p(t)$. Set $k = 0$ and then:

1. Solve the direct problem and compute $T(x,t)$, based on $g_p^k(t)$.
2. Check the stopping criterion. Continue if not satisfied.
3. Knowing $T(x_{meas},t)$ and measured temperature $Y(t)$, solve the adjoint problem and compute $\lambda(0.5, t)$.
4. Knowing $\lambda(0.5, t)$, compute $S[g_p^k(t)]$.
5. Knowing the gradient $S[g_p^k(t)]$, compute g^k and the direction of descent $d^k(t)$.
6. Solve the sensitivity problem to obtain $\Delta T\,[x_{meas},t;d^k(t)]$.
7. Knowing $\Delta T\,[x_{meas},t;d^k(t)]$, compute the search step size β^k .
8. Knowing the search step size β^k and the direction of descent $d^k(t)$, compute the new estimate $g_p^{k+1}(t)$ and return to step 1.

Applications

SIMULATED MEASUREMENTS

$$Y = Y_{ex} + \omega\sigma$$

where Y_{ex} = solution of the direct problem for an *a priori* specified function or parameters

ω = random variable with normal distribution, zero mean and unitary standard deviation

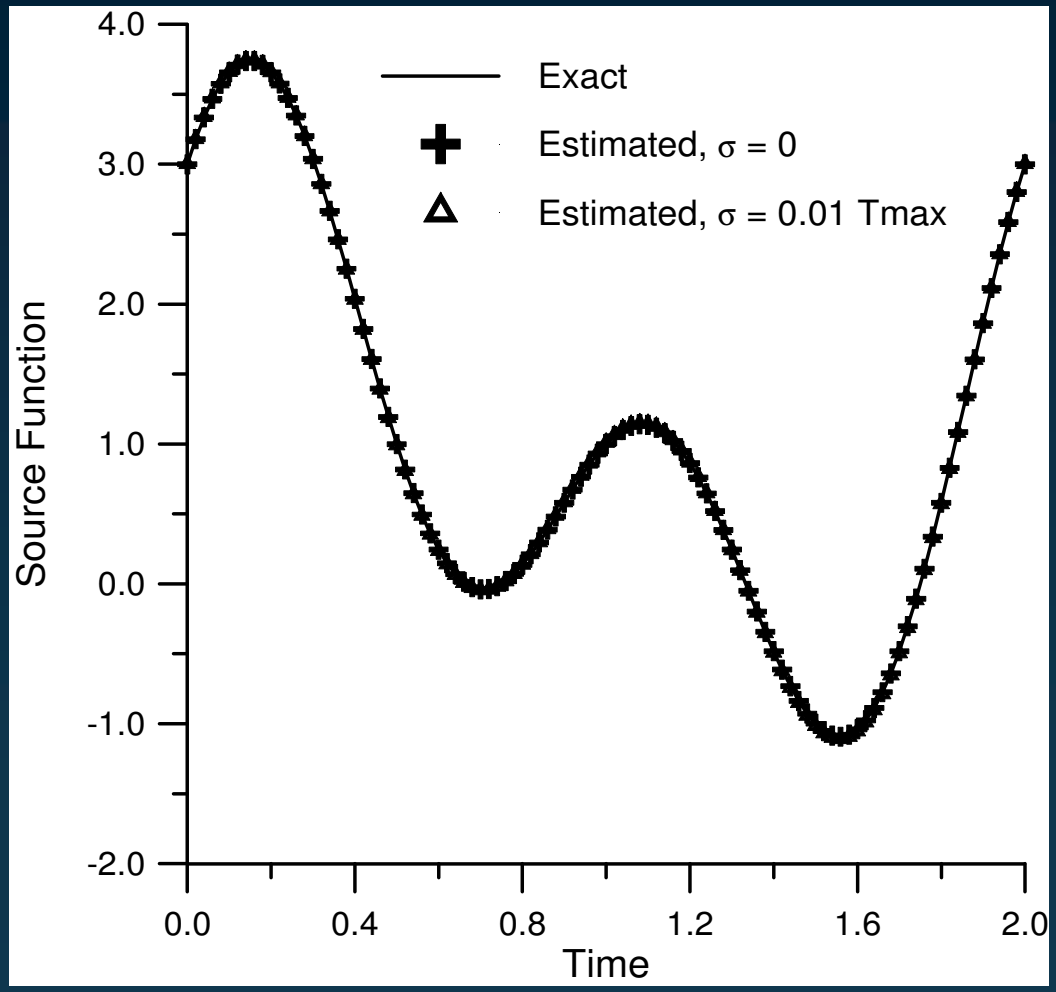
σ = standard-deviation of the measurements

Examples

Parameter Estimation *versus* Function Estimation

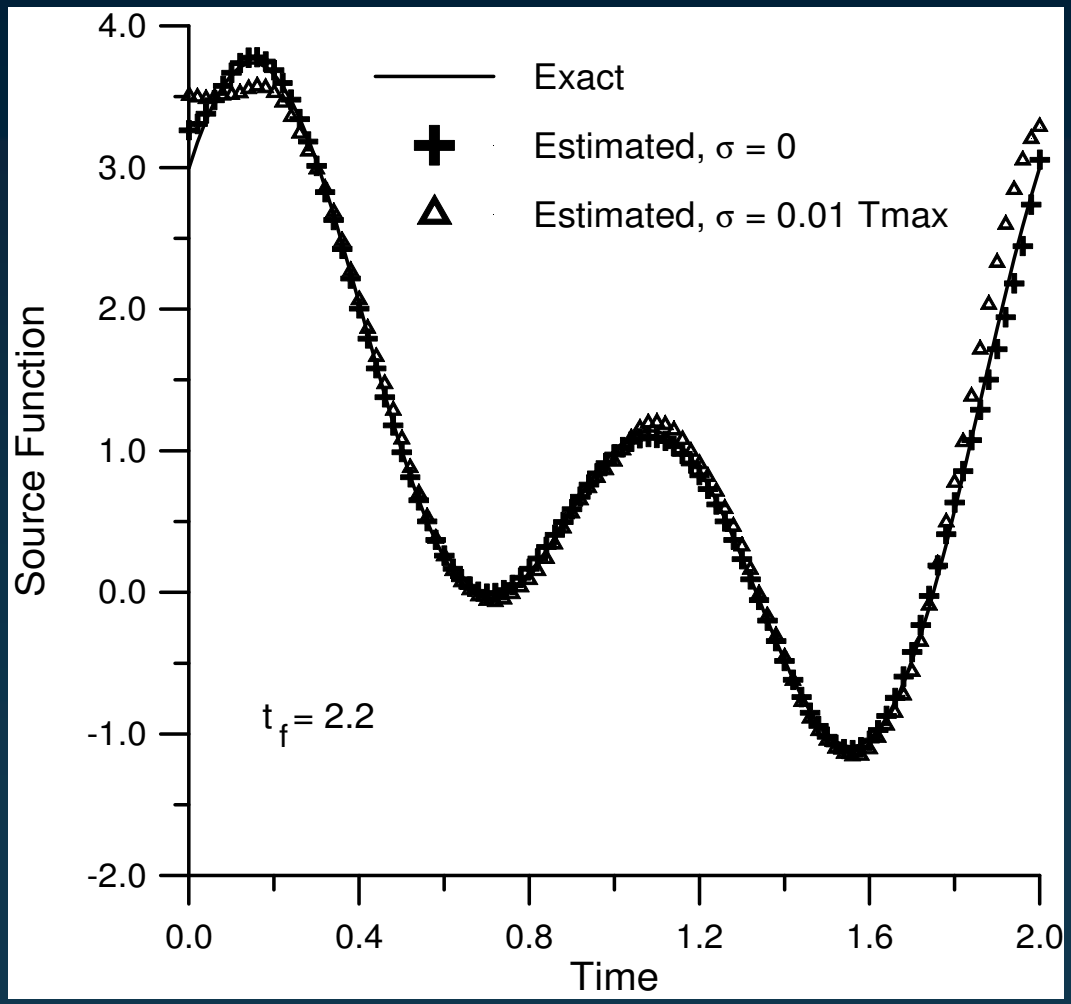
$$g_p(t) = 1 + \sin \pi t + \cos \pi t + \sin 2\pi t + \cos 2\pi t$$

Examples



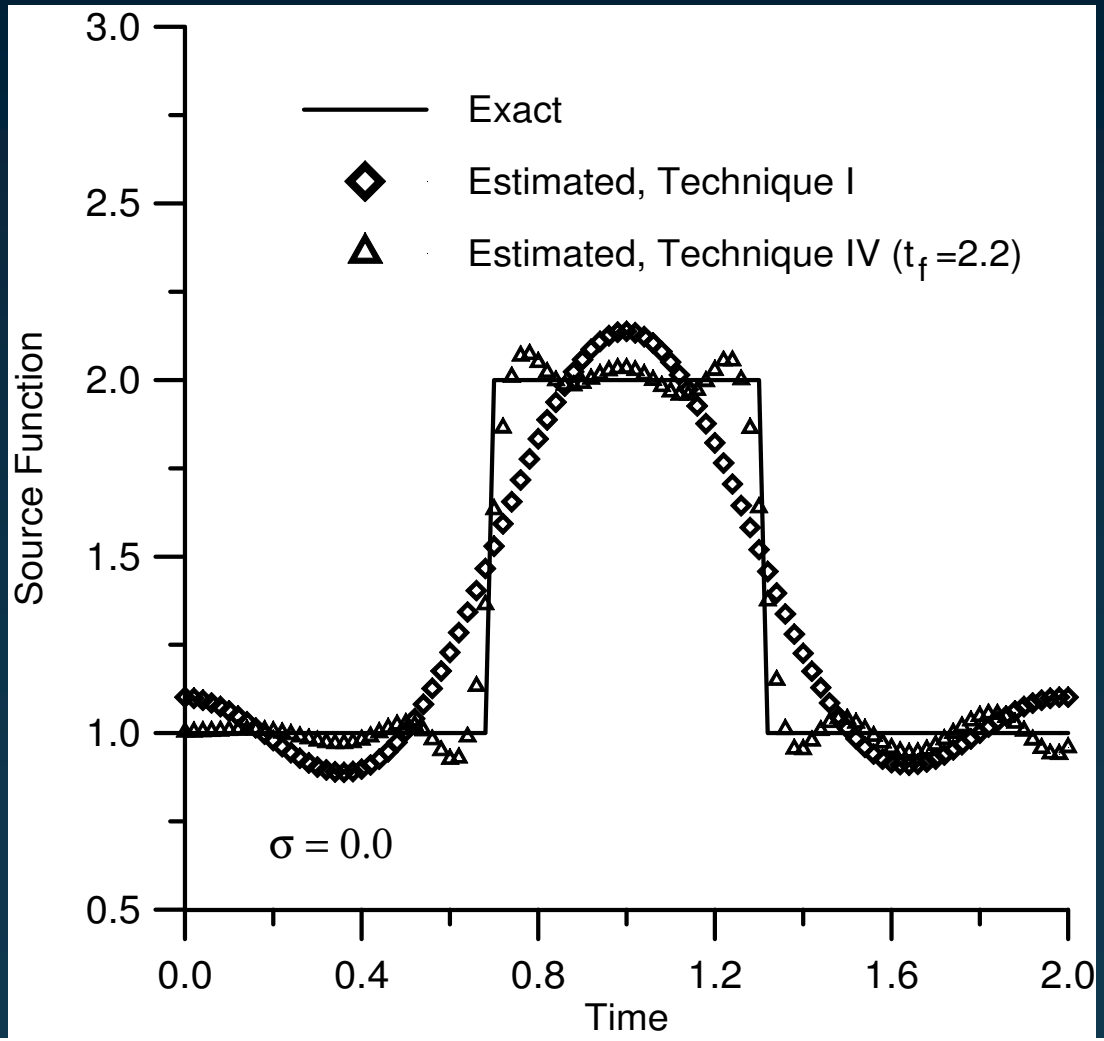
Levenberg-Marquardt

Examples



Function Estimation

Examples


Exact
Estimated, Technique I
Estimated, Technique IV (t_f =2.2)
Source Function
Time
σ = 0.0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.4
0.8
1.2
1.6
2.0

Examples

Solve the inverse problem of estimating the boundary heat flux $q(t)$. Assume that no information is available on the functional form of $q(t)$, except that it belongs to the space of square integrable functions in the domain $0 < t < 1$. Use for the inverse analysis 100 equally spaced transient measurements in $0 < t \leq 1$, of a sensor located at $x_{meas} = 0$.

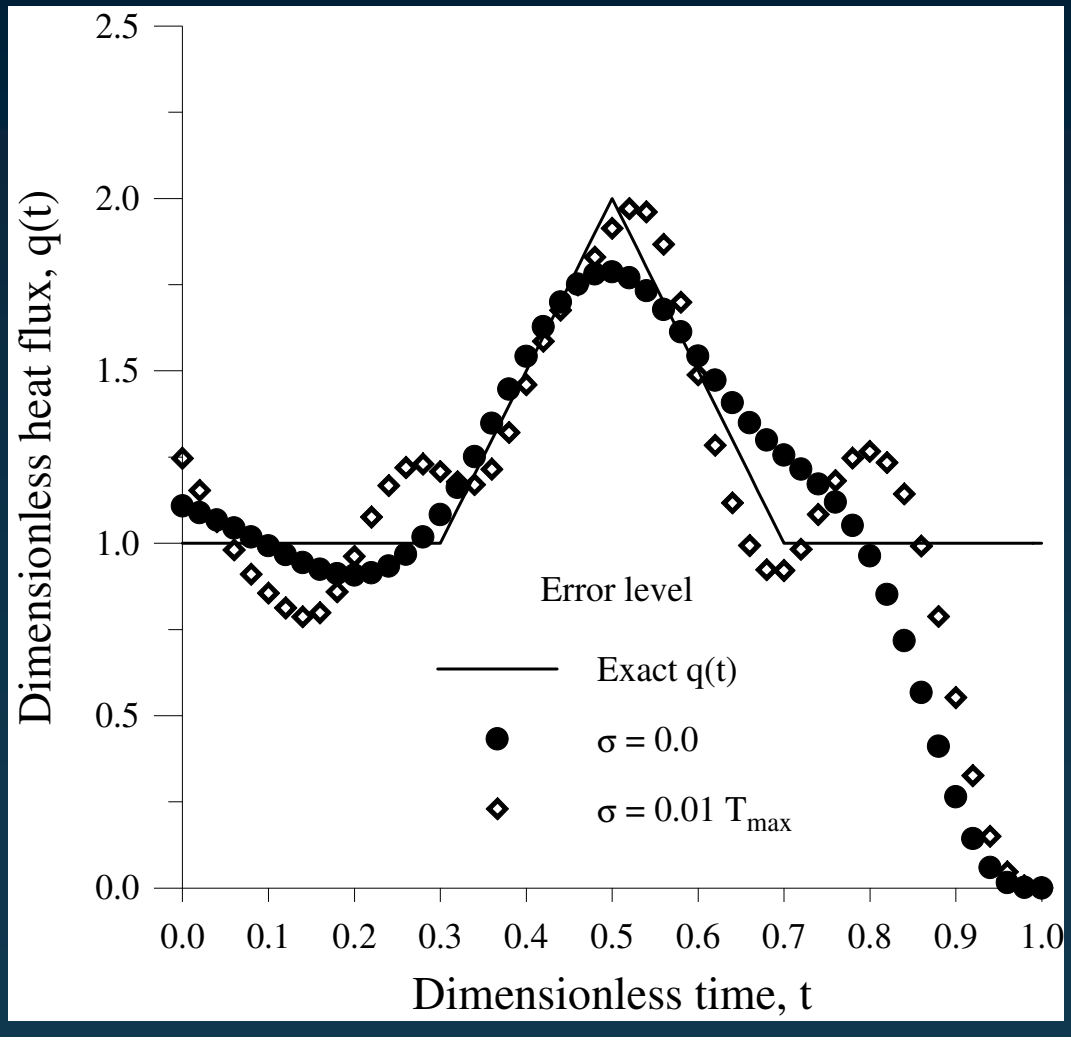
$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} \qquad \text{in } 0 < x < 1, \quad \text{for} \quad t > 0$$

$$\frac{\partial T}{\partial x} = 0 \qquad \text{at } x = 0, \quad \text{for} \quad t > 0$$

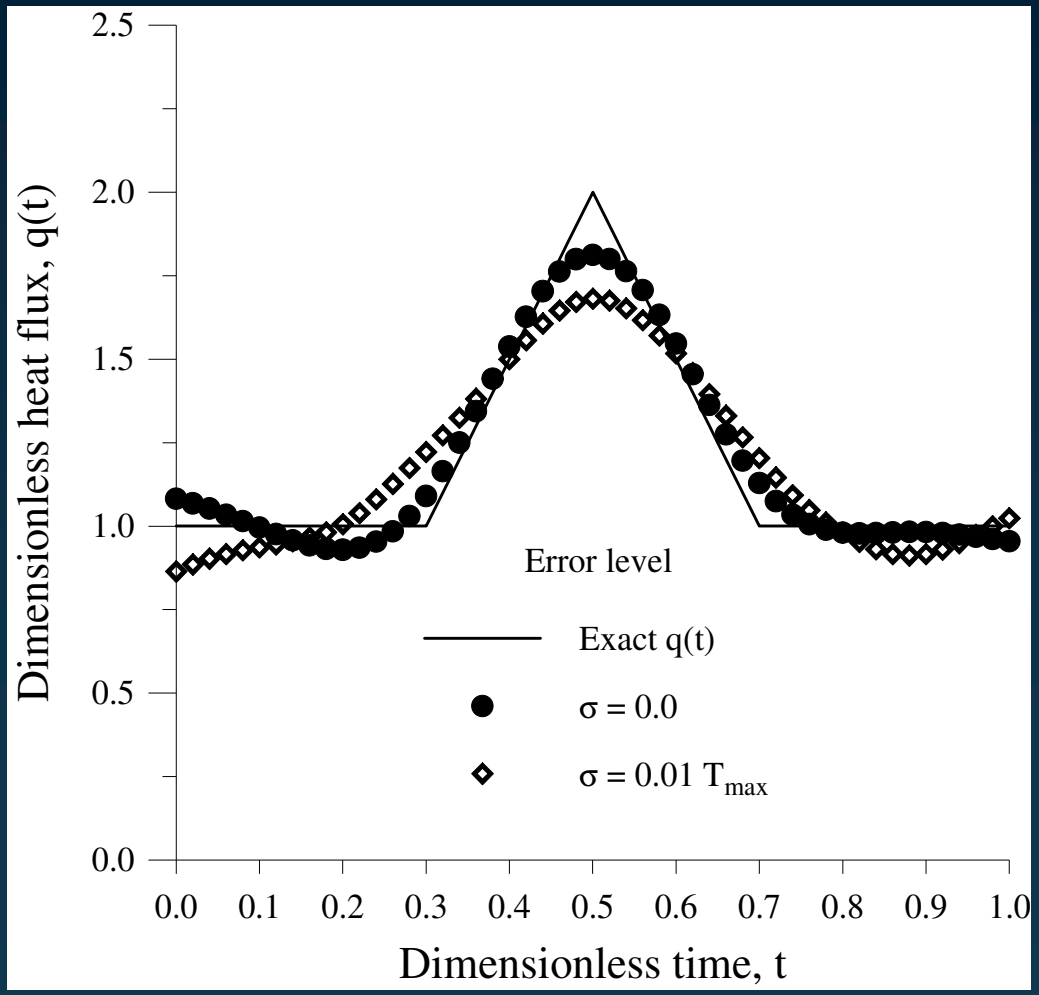
$$\frac{\partial T}{\partial x} = q(t) \qquad \text{at } x = 1, \quad \text{for} \quad t > 0$$

$$T = 0 \qquad \text{for } t = 0, \quad \text{in } 0 < x < 1$$

Examples
Effect Of The Null Gradient
At The Final Time

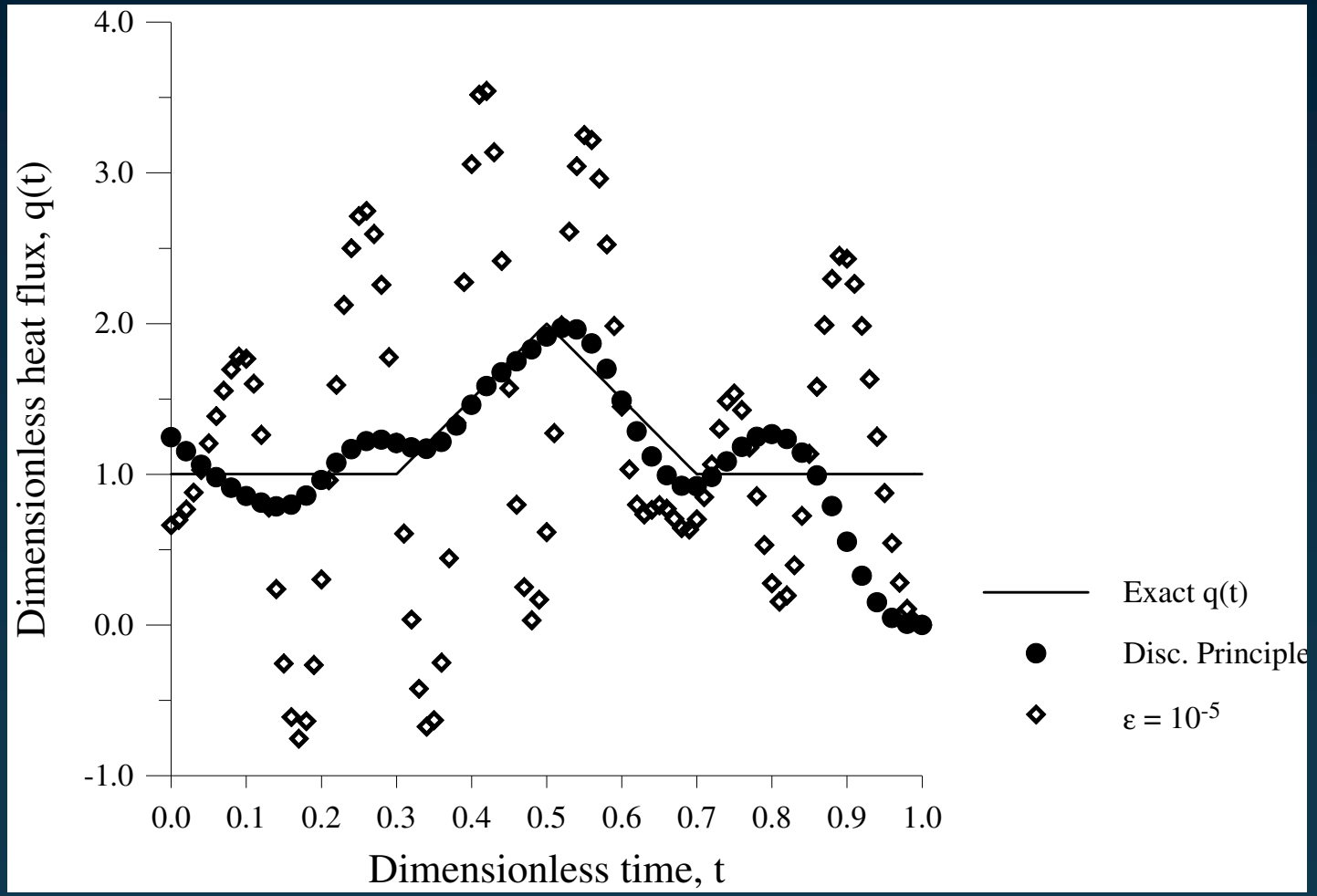


Examples
Effect Of The Null Gradient At The Final Time



Measurements taken until $t = 1.5$

Examples
Regularization Effect Of The Discrepancy Principle



Example

- Comparison of two solution techniques for the inverse problem of simultaneously estimating the spatial variations of diffusion coefficients and source terms
  - Conjugate Gradient Method with Adjoint Problem
  - Tikhonov's Regularization with Hybrid Optimizer

$$C(\mathbf{r})\frac{\partial U(\mathbf{r},t)}{\partial t} = \nabla\cdot[D(\mathbf{r})\nabla U] + \mu(\mathbf{r})U$$

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial U}{\partial x}\right) + \mu(x)U \qquad \text{in} \quad 0 < x < 1 \;,\; \textit{for}\; t > 0$$

$$\frac{\partial U}{\partial x} = 0 \qquad \text{at} \quad x = 0 \quad \text{for} \quad t > 0$$

$$D(x)\frac{\partial U}{\partial x} = 1 \qquad \text{at} \quad x = 1 \quad \text{for} \quad t > 0$$

$$U = 0 \qquad \text{for} \quad t = 0 \quad \text{in} \quad 0 < x < 1$$

Example 3

DIRECT PROBLEM

Known:

$D(x)$ and $\mu(x)$

Determine:

$U(x,t)$

INVERSE PROBLEM

Known:

Measurements of $U(x,t)$

Simultaneously Estimate:

$D(x)$ and $\mu(x)$

Example 3

TEST-CASE

Heat Conduction – Steel

$l^* = 0.050\ m$

$t_f^* = 60\ s$

50 measurements per sensor

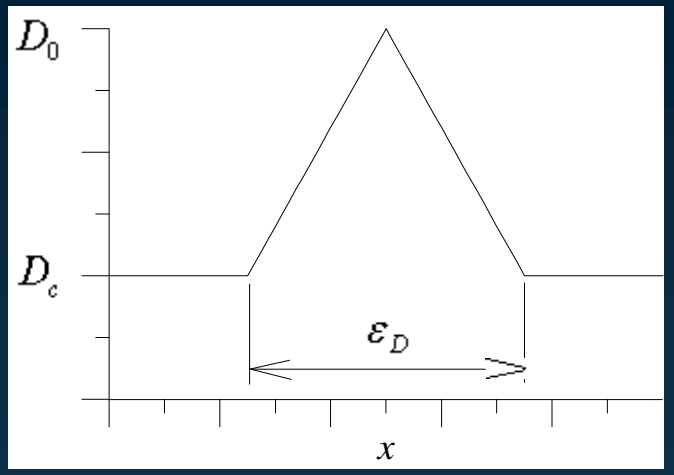
Simulated measurements

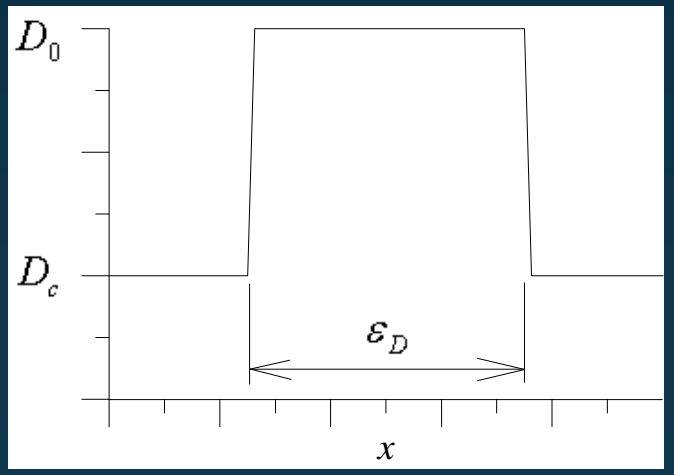
$$\rho = 7833\ {kg}/{m^3}$$

$$D_c = 54\ {W}/{m^\circ C}$$

$$c_p = 465\ {J}/{kg^\circ C}$$

$$\mu_c^* = 10^5\ {W}/{m^3{}^\circ C}$$

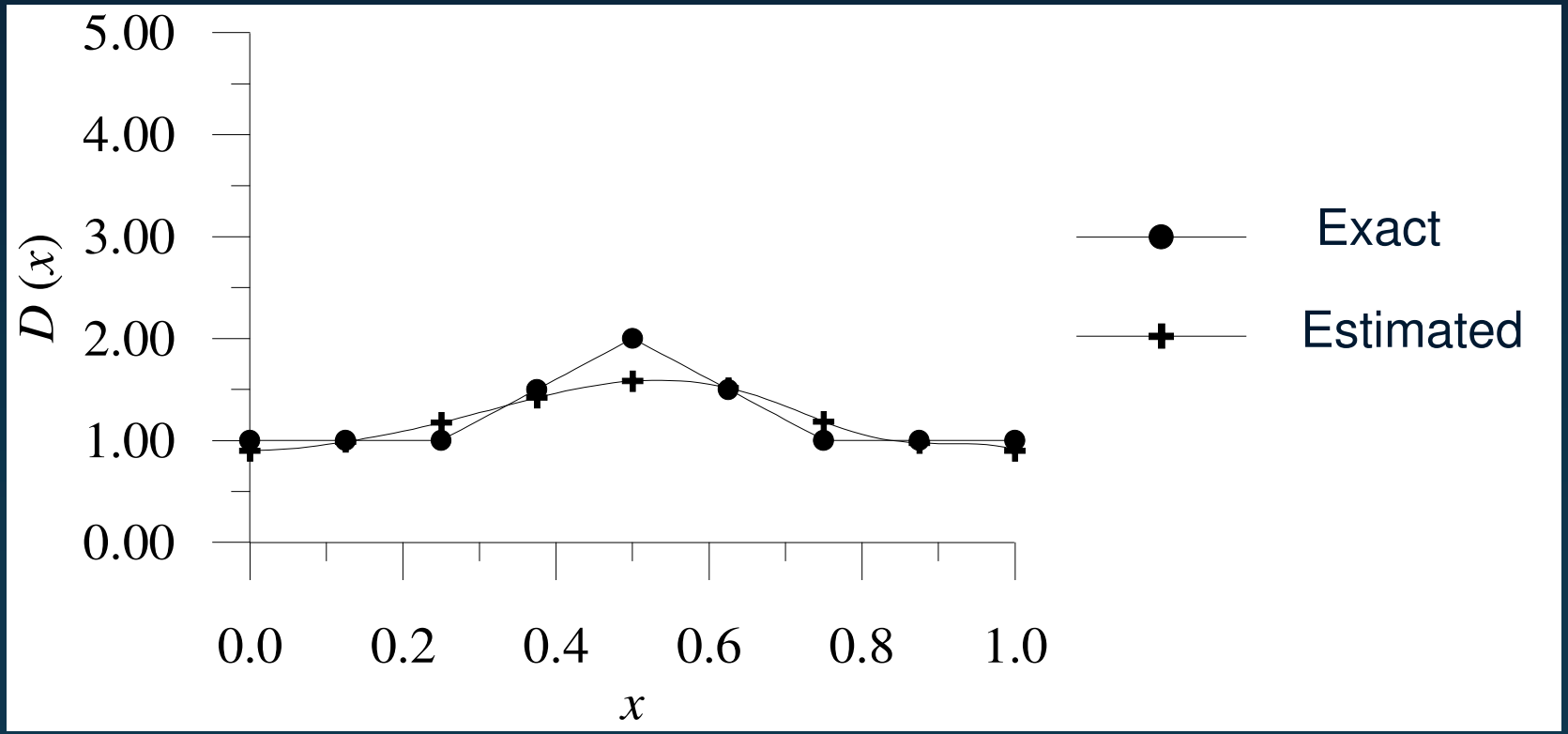




Example 3

RESULTS - CGM
Estimation of $D(x)$ for Known $\mu(x)$

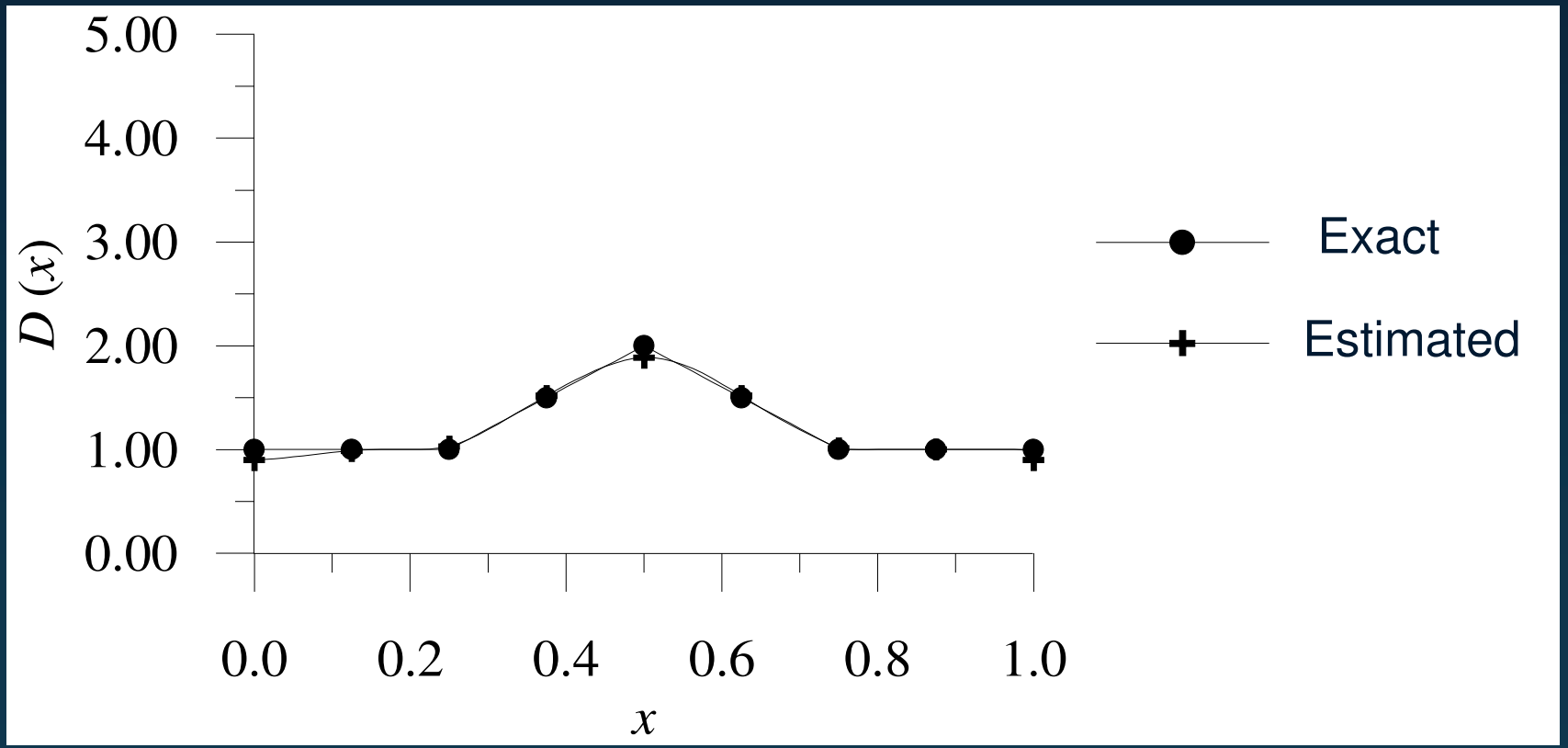
2 Non-intrusive sensors – Errorless measurements



Example 3

RESULTS - CGM
Estimation of $D(x)$ for Known $\mu(x)$

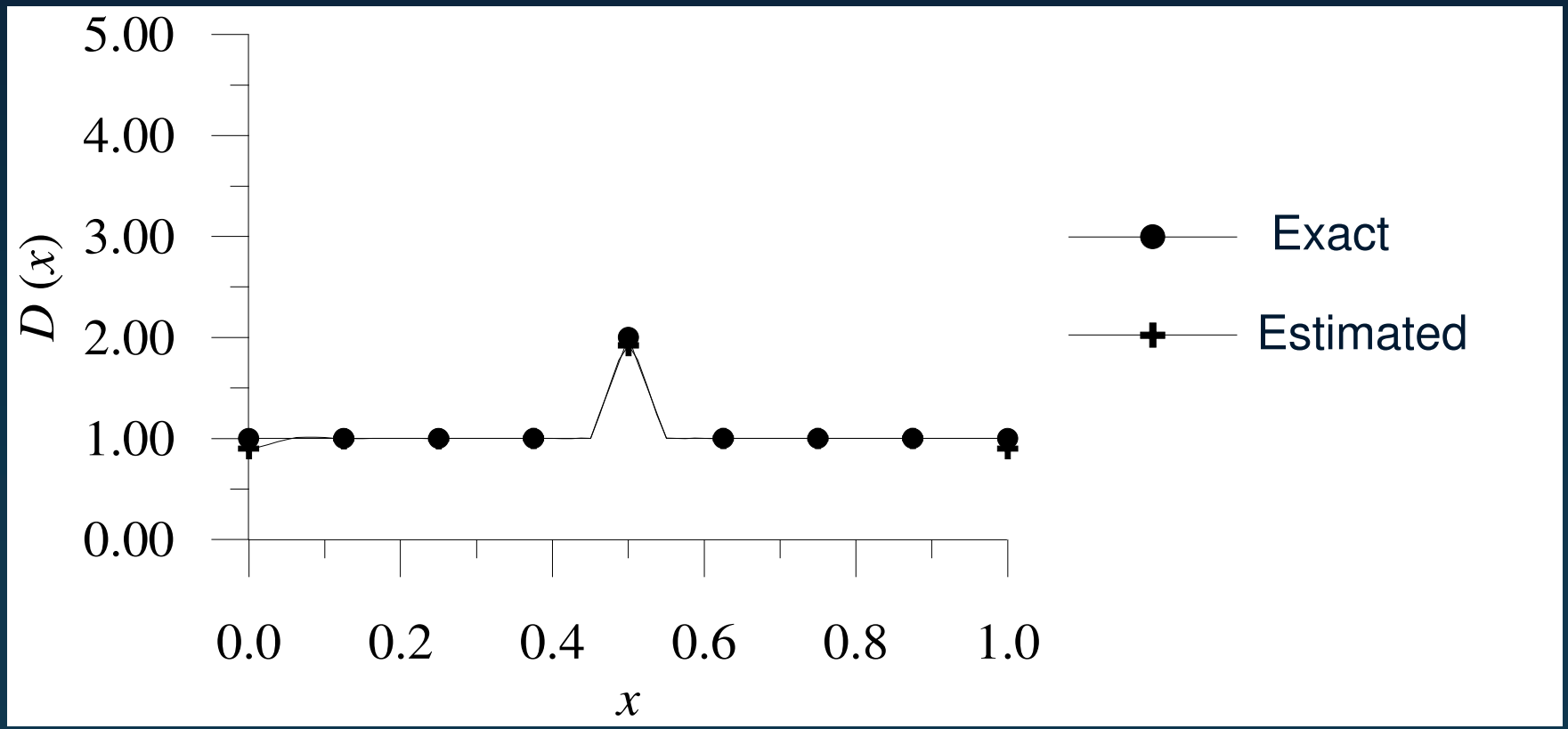
10 Sensors – Errorless measurements



Example 3

RESULTS - CGM
Estimation of $D(x)$ for Known $\mu(x)$

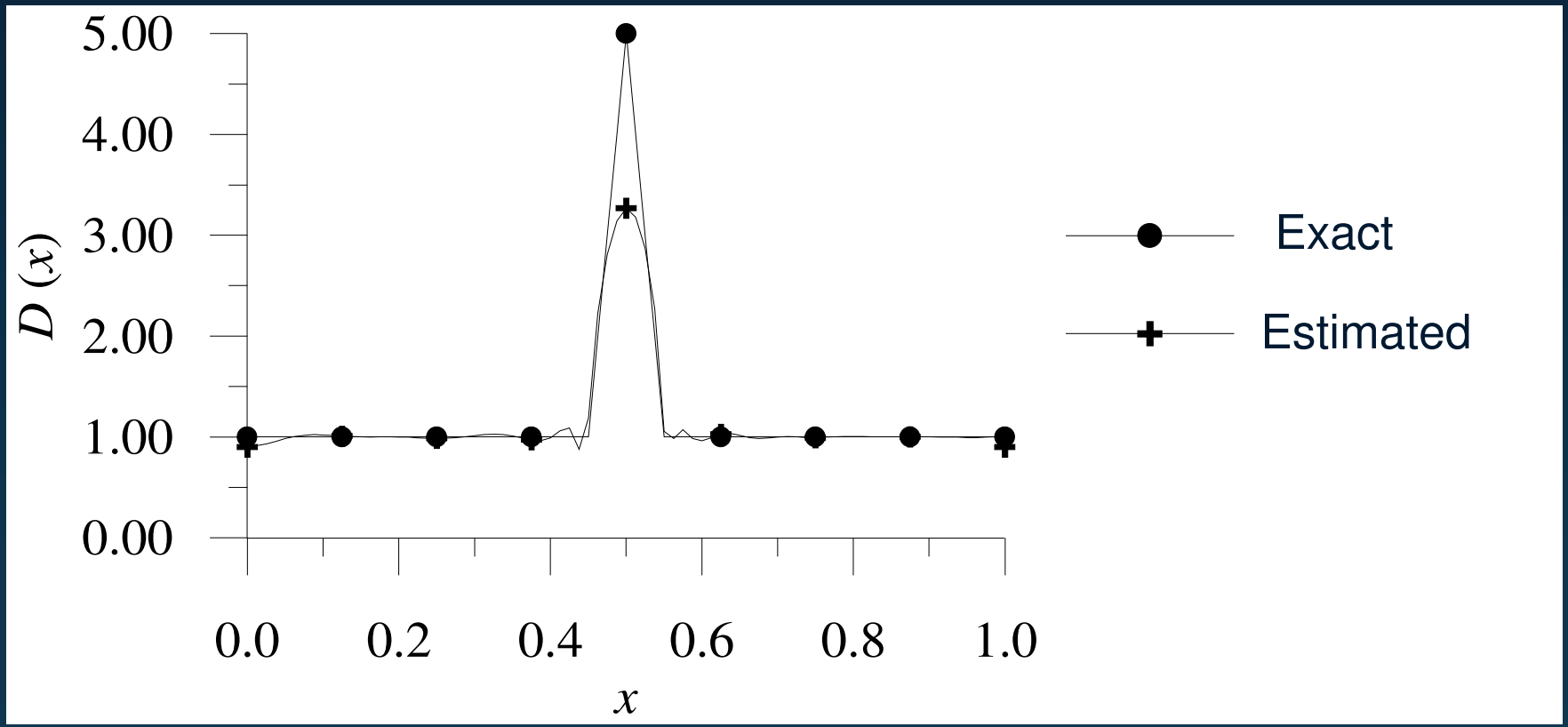
80 Sensors – Errorless measurements



Example 3

RESULTS - CGM
Estimation of $D(x)$ for Known $\mu(x)$

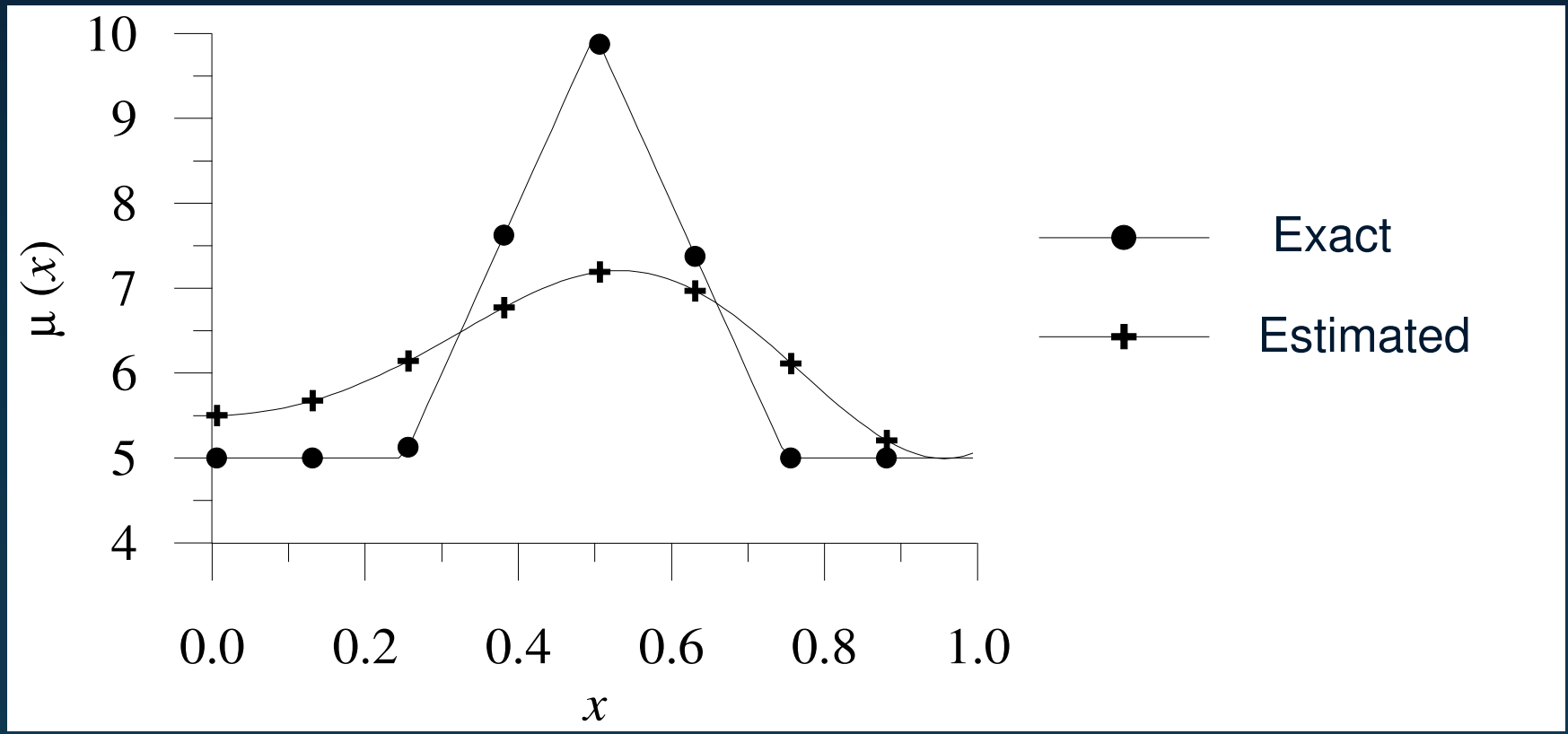
80 Sensors – Errorless measurements



Example 3

RESULTS - CGM
Estimation of $\mu(x)$ for Known $D(x)$

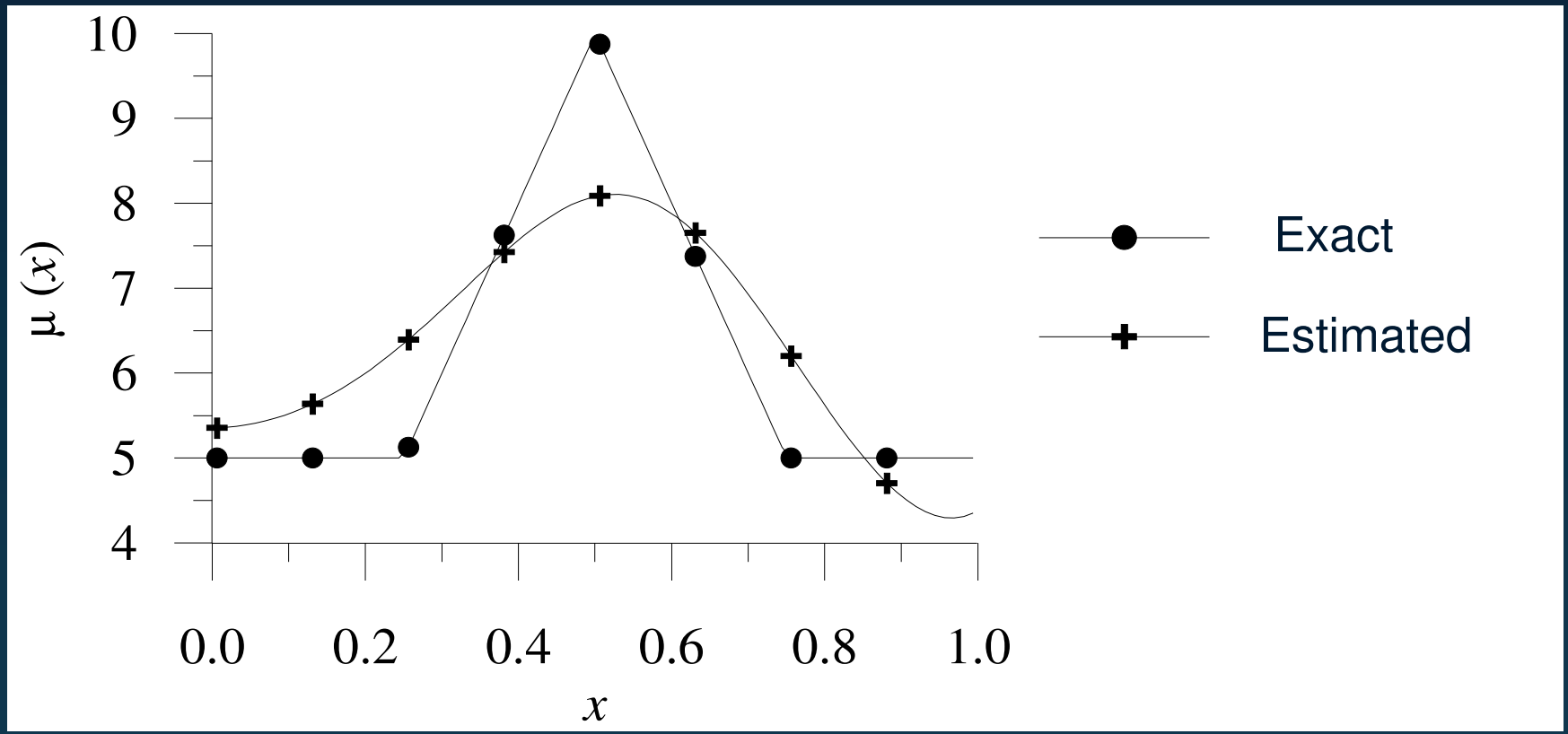
2 Non-intrusive sensors – Errorless measurements



Example 3

RESULTS - CGM
Estimation of $\mu(x)$ for Known $D(x)$

10 Sensors – Errorless measurements

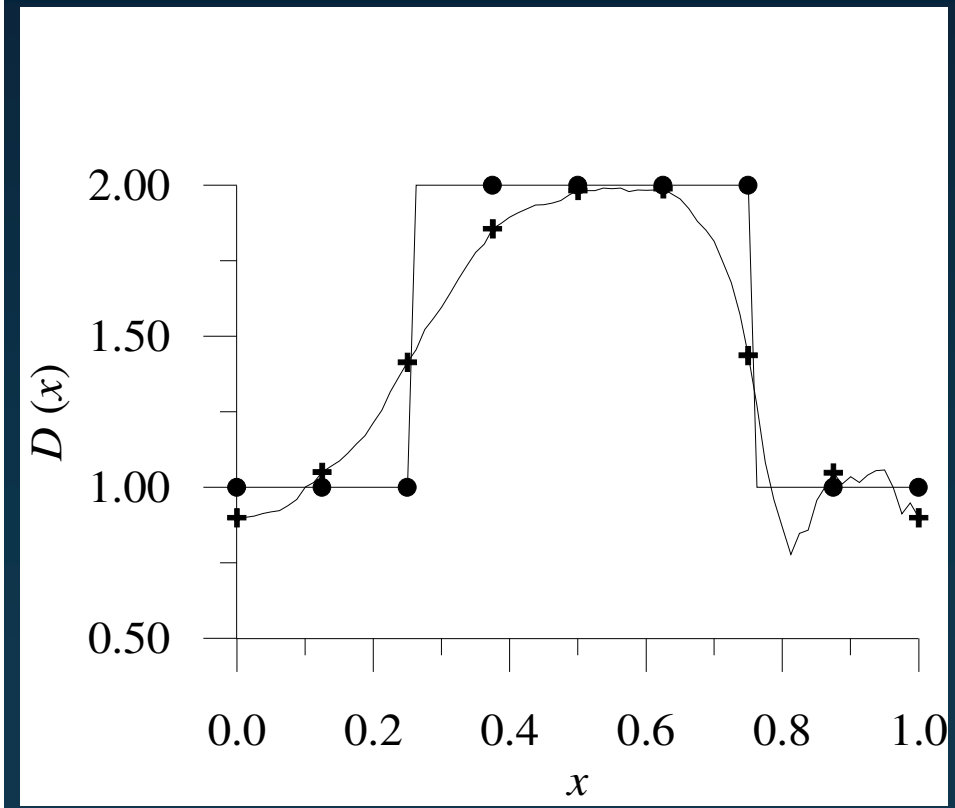


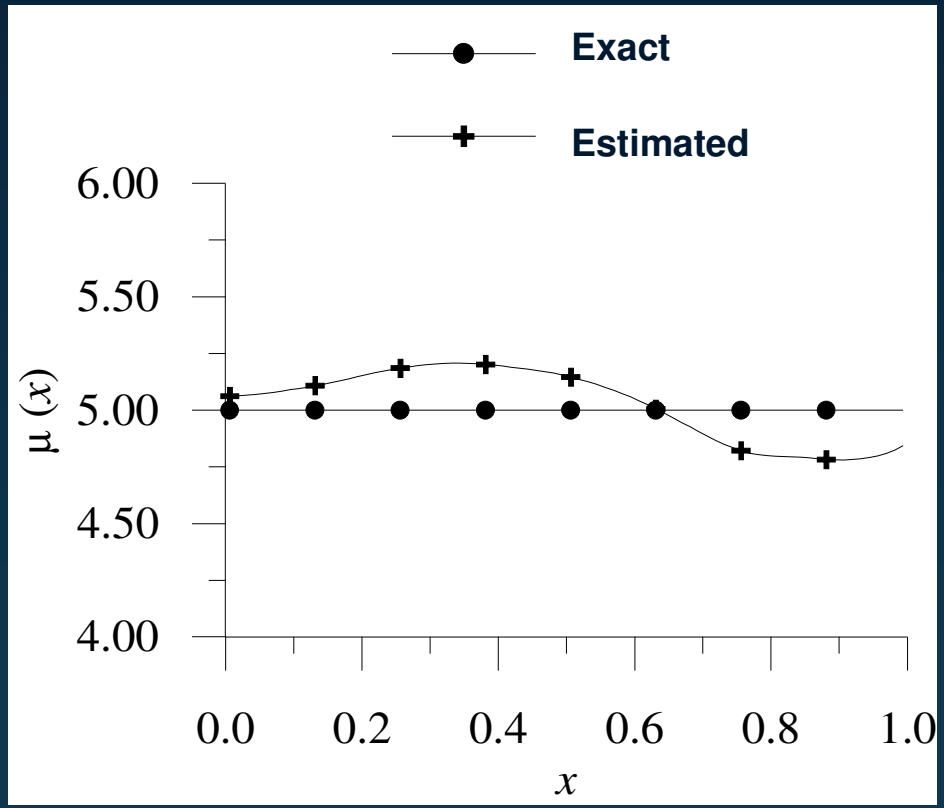
Example 3

RESULTS - CGM

Simultaneous estimation of $\mu(x)$ and $D(x)$

2 Non-intrusive sensors – Measurements with σ = 0.01 Y_{max}



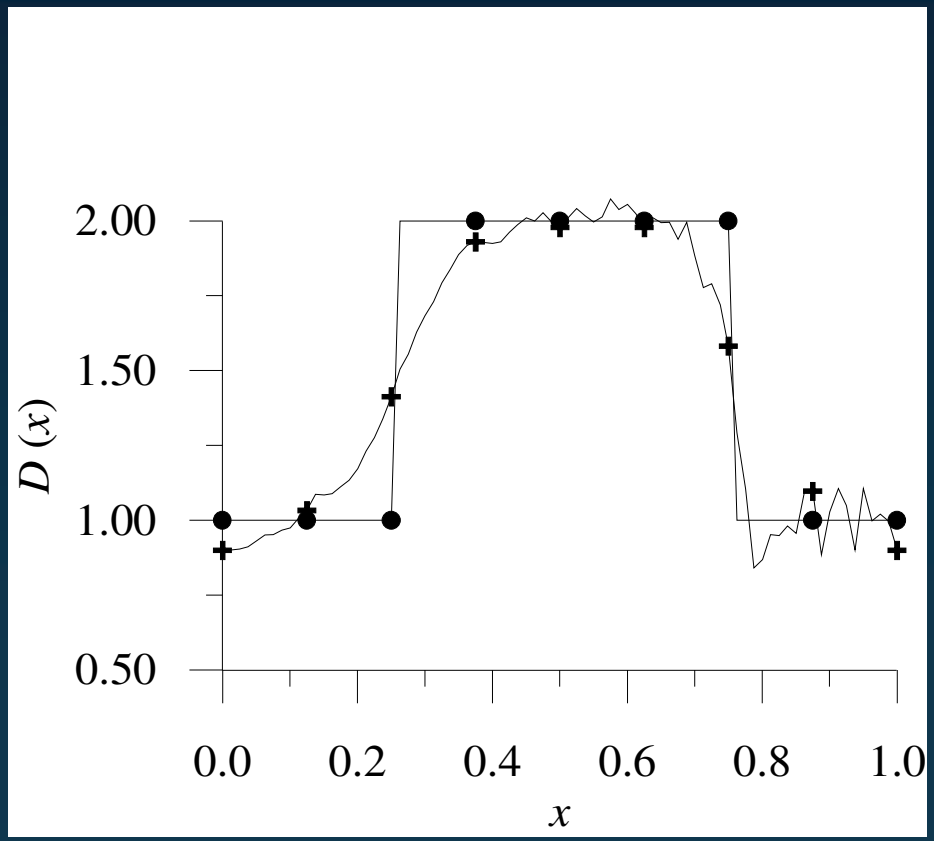


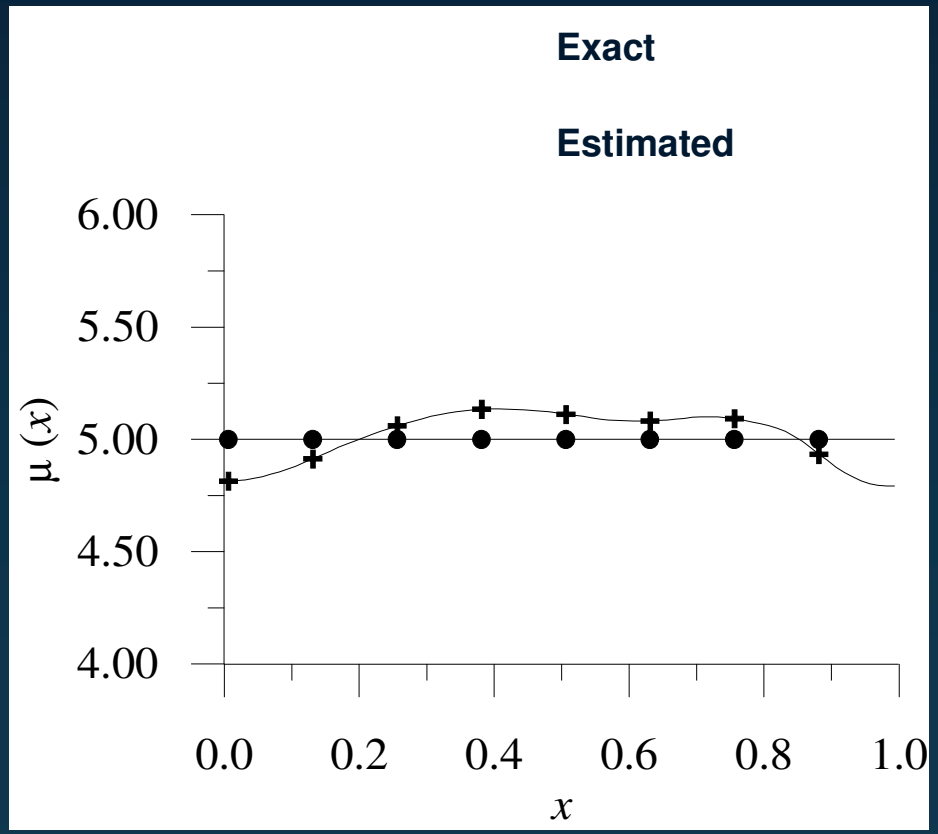
Example 3

RESULTS - CGM

Simultaneous estimation of $\mu(x)$ and $D(x)$

10 Sensors – Measurements with σ = 0.01 Y_{max}



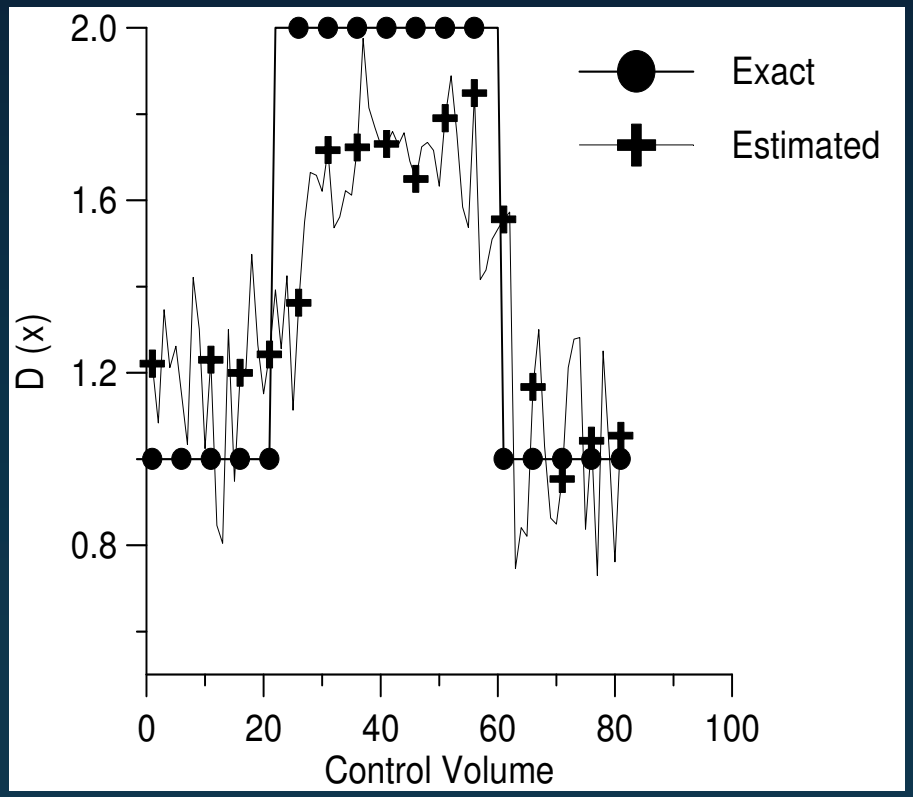


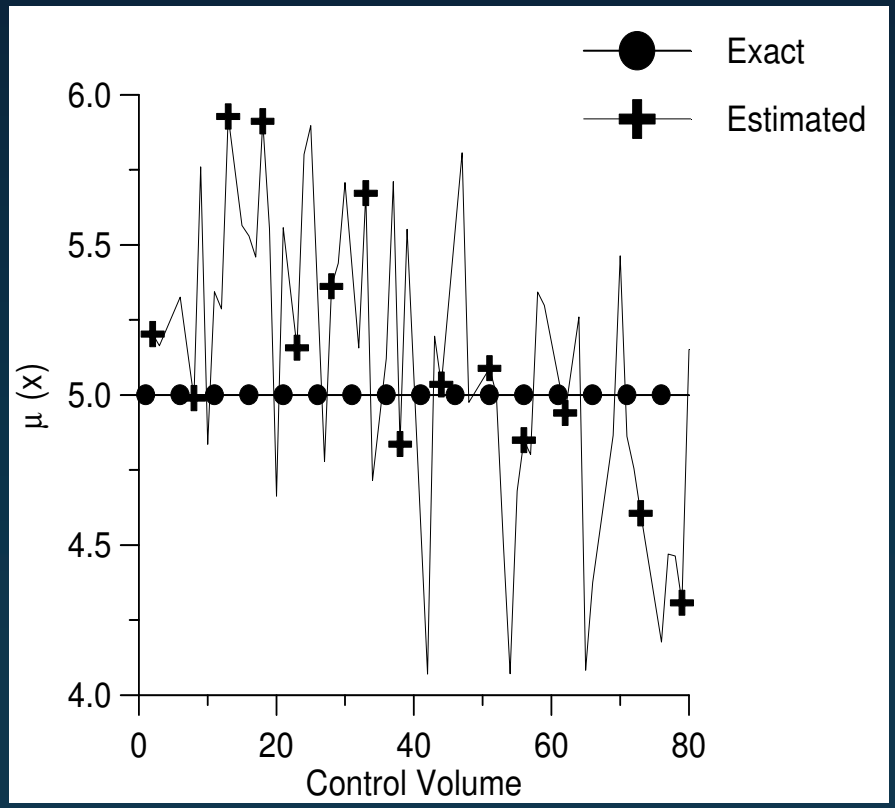
Example 3

RESULTS - TIKHONOV

Simultaneous estimation of $\mu(x)$ and $D(x)$

10 Sensors – Measurements with $\sigma = 0.01$ $Y_{max} - \alpha_1 = 0$



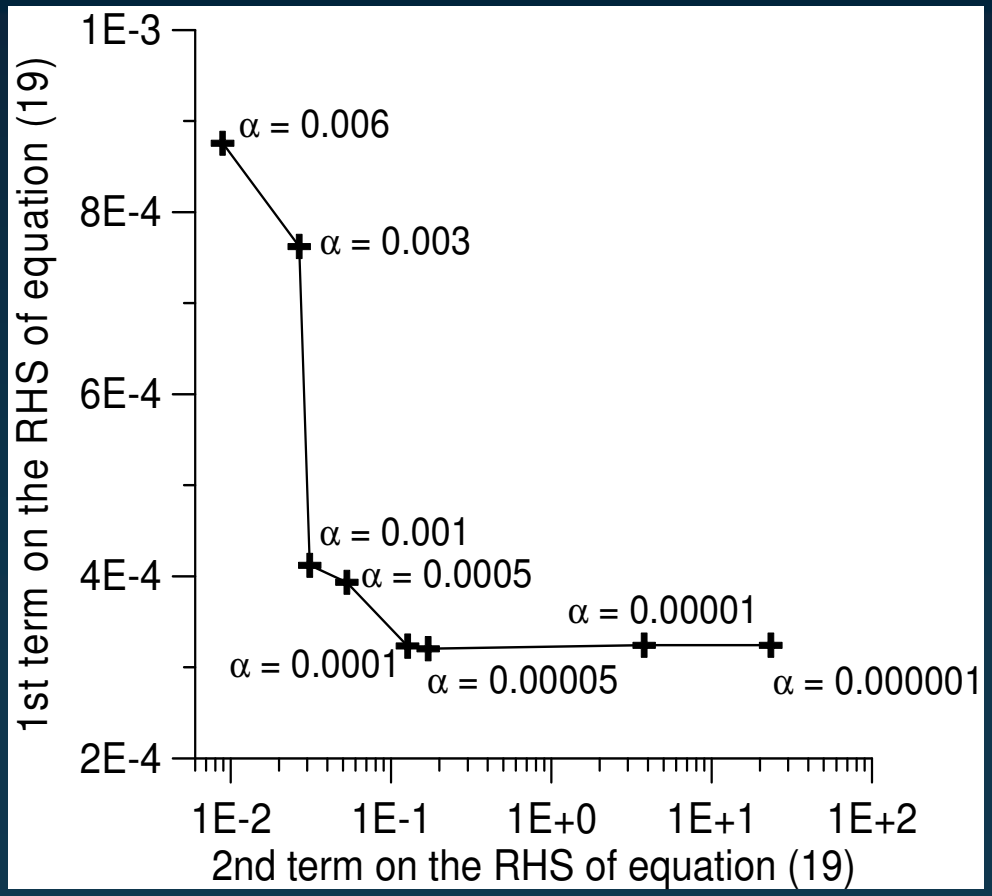


Example 3

RESULTS - TIKHONOV

Simultaneous estimation of $\mu(x)$ and $D(x)$

10 Sensors – Measurements with σ = 0.01 Y_{max}

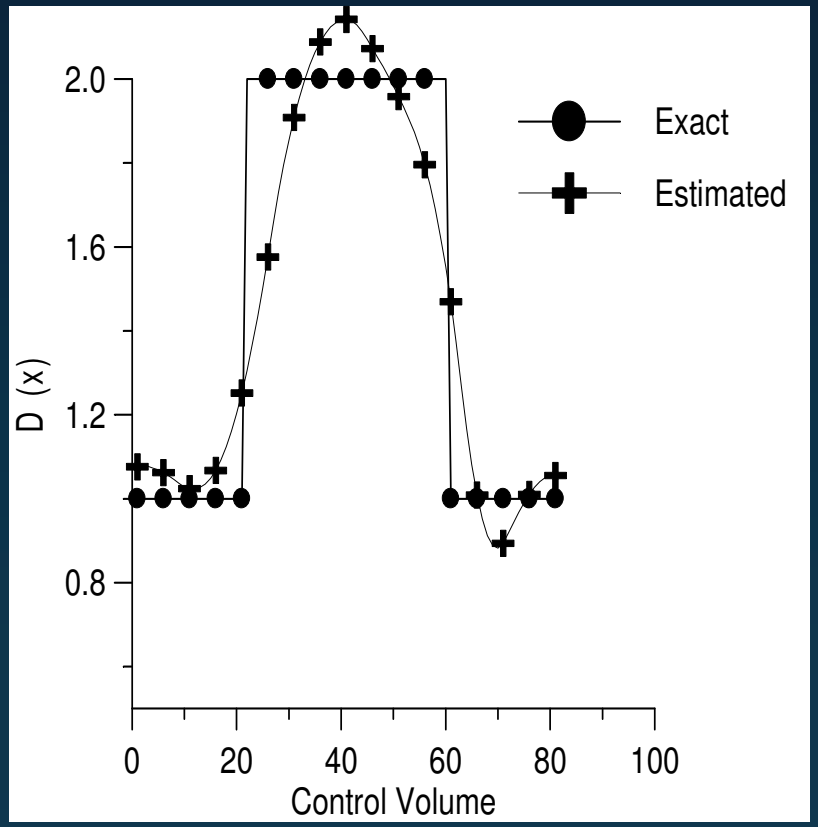


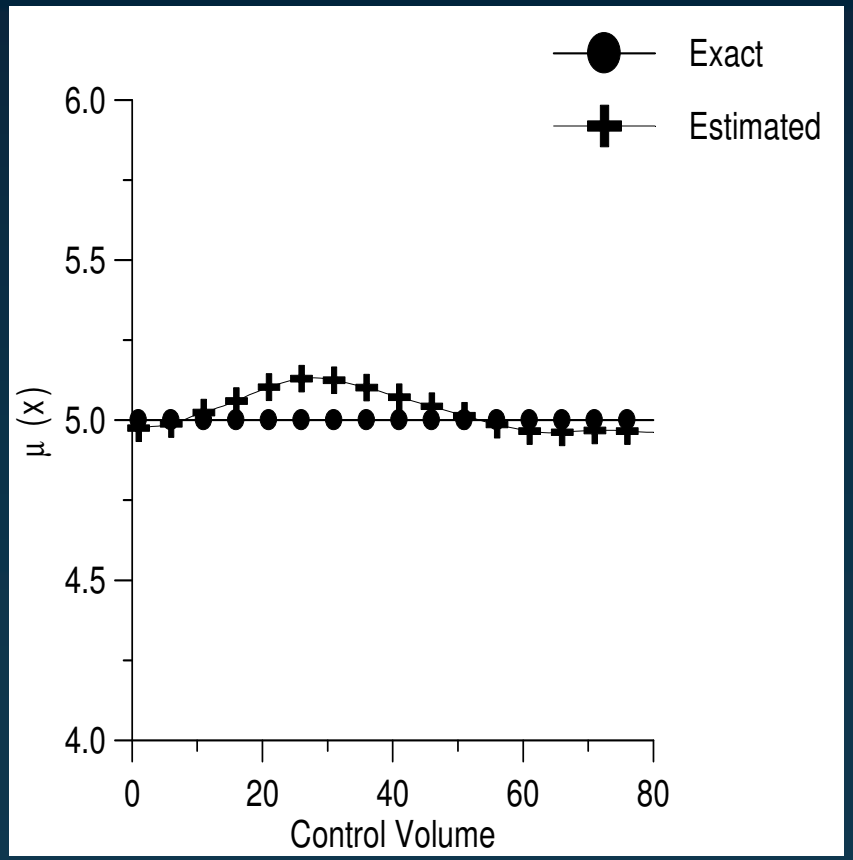
Example 3

RESULTS - TIKHONOV

Simultaneous estimation of $\mu(x)$ and $D(x)$

10 Sensors – Measurements with $\sigma = 0.01\, Y_{max}$ – $\alpha_1 = 0.0001$



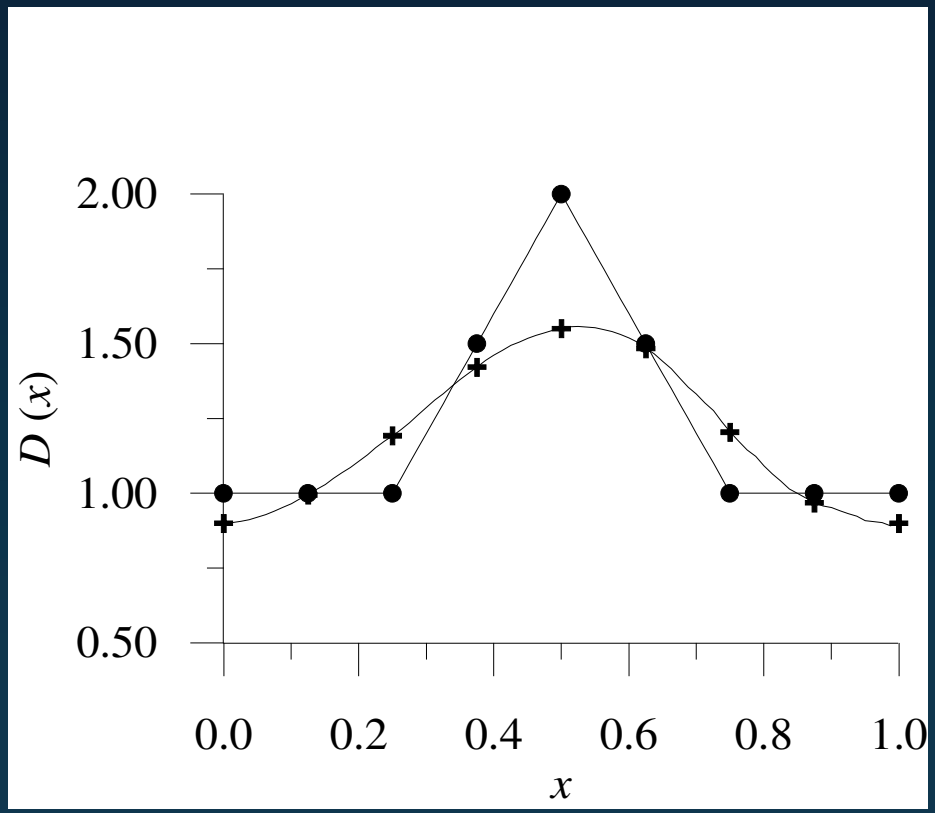


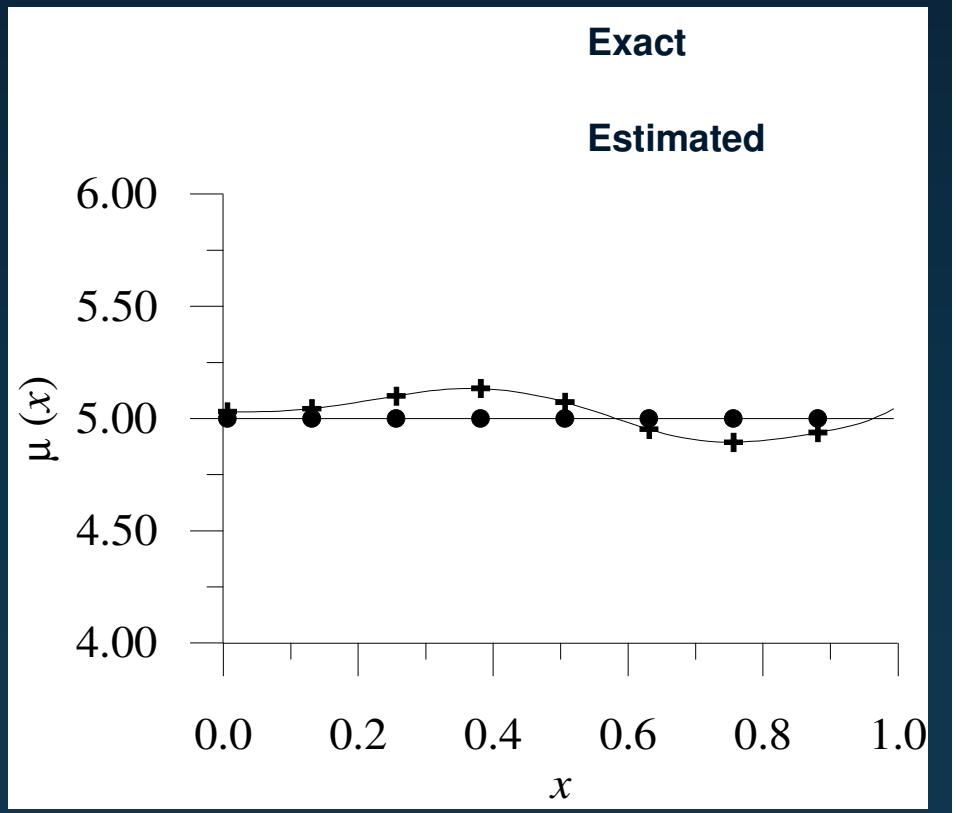
Example 3

RESULTS - CGM

Simultaneous estimation of $\mu(x)$ and $D(x)$

2 Non-intrusive sensors – Measurements with σ = 0.01 Y_{max}



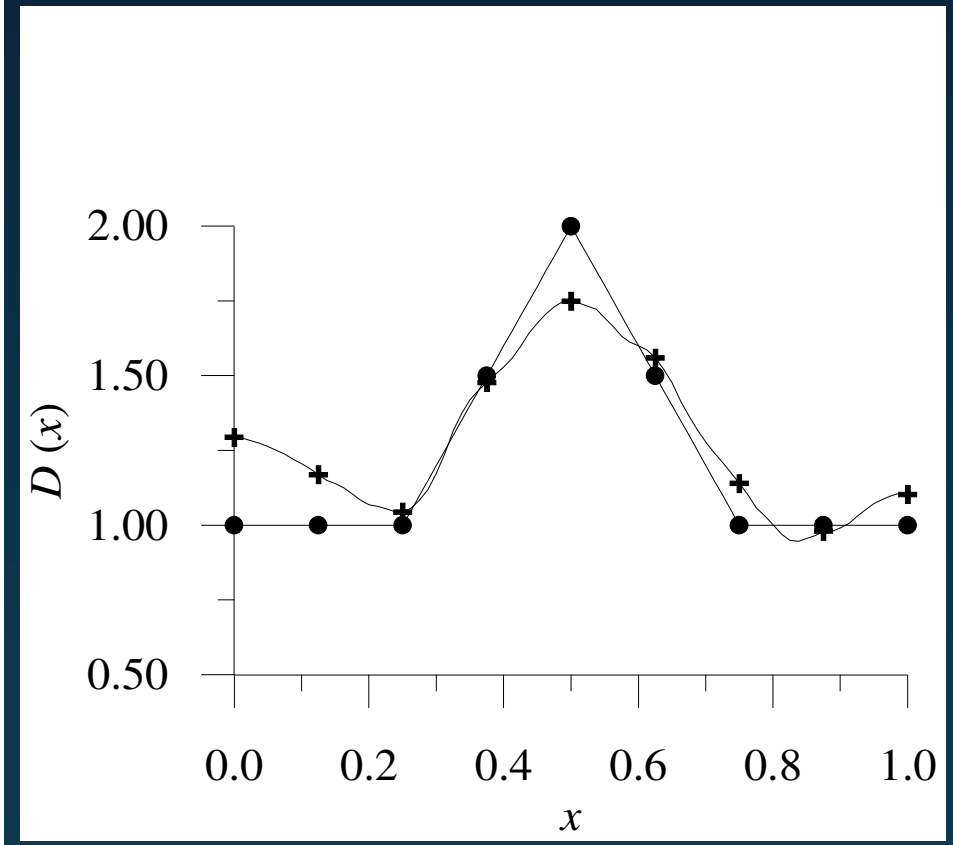


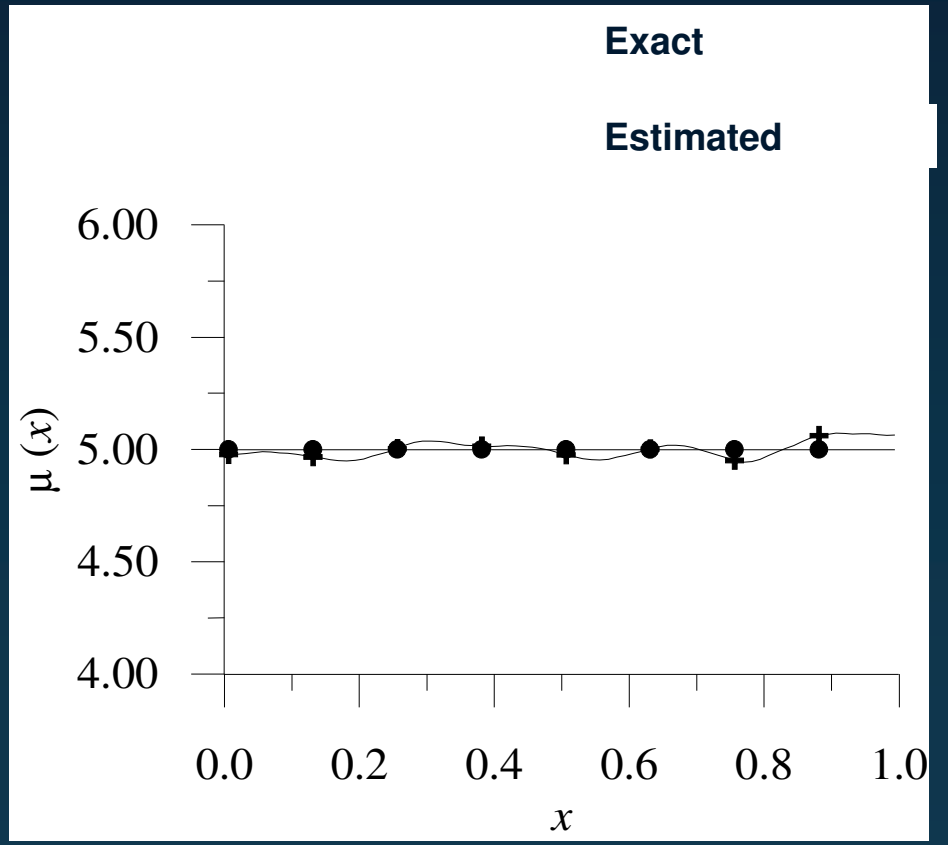
Example 3

RESULTS - TIKHONOV

Simultaneous estimation of $\mu(x)$ and $D(x)$

2 Non-intrusive sensors – Measurements with σ = 0.01 Y_{max}



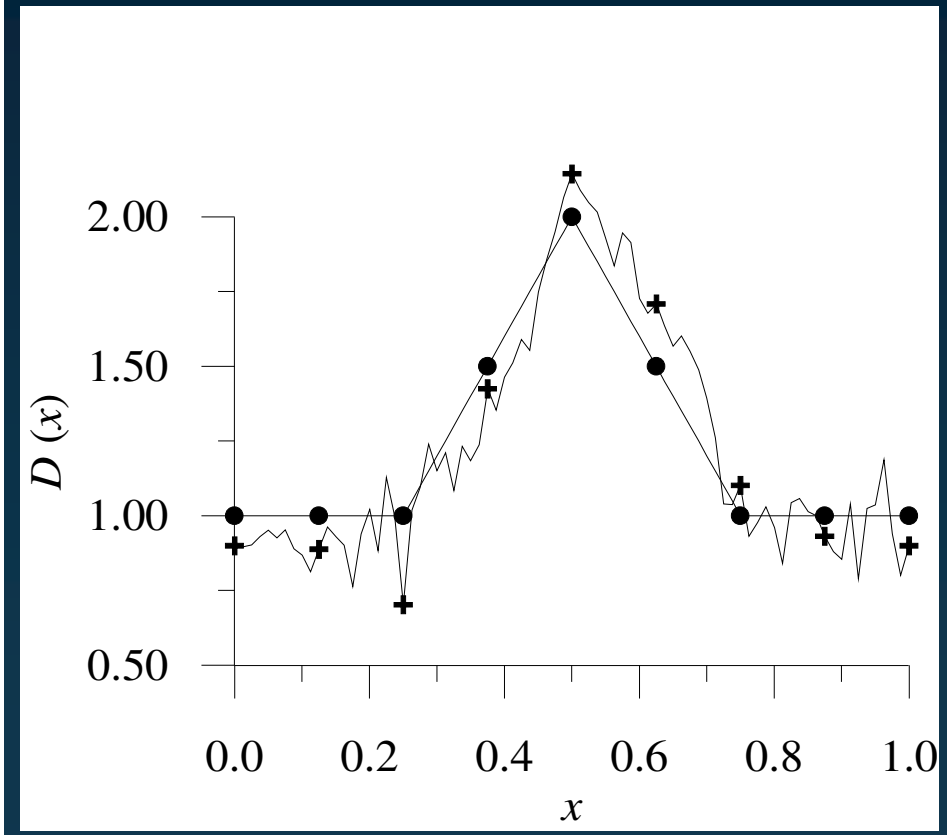


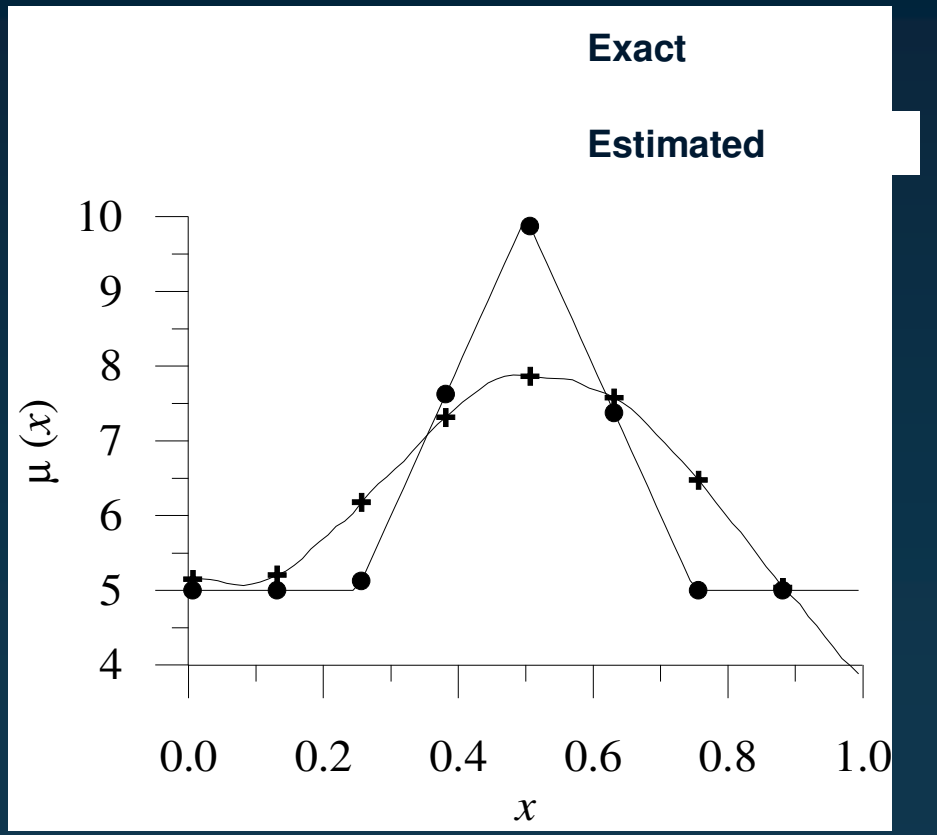
Example 3

RESULTS - CGM

Simultaneous estimation of $\mu(x)$ and $D(x)$

50 Sensors – Measurements with $\sigma = 0.01$ Y_{max}



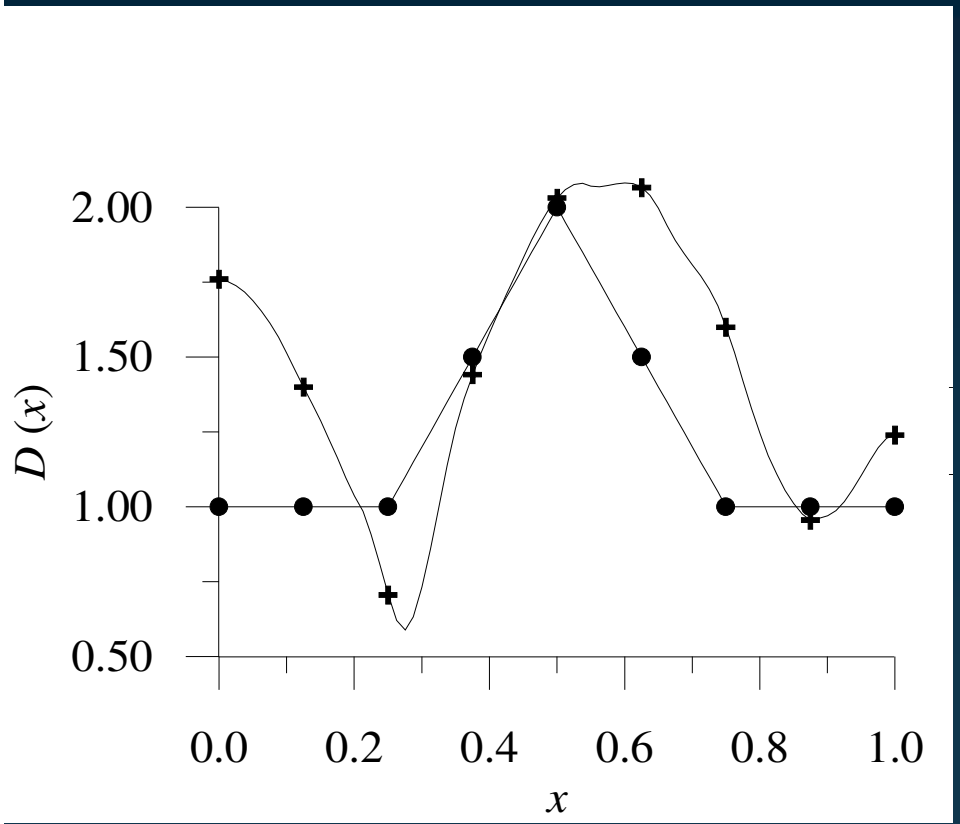


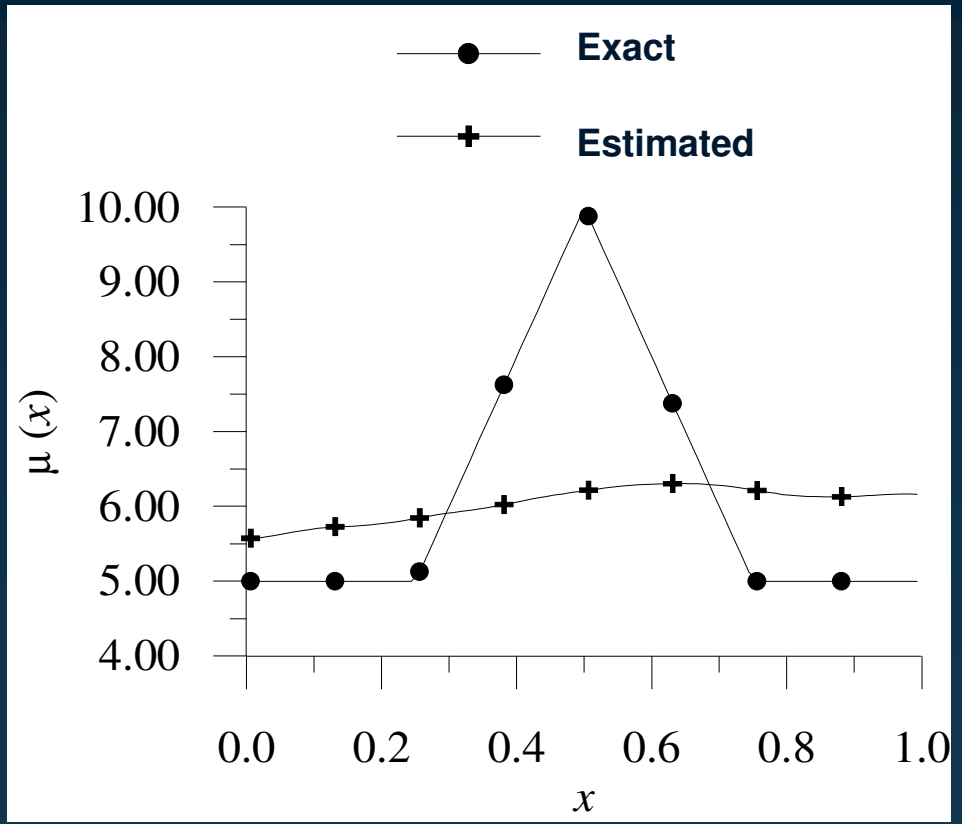
Example 3

RESULTS - TIKHONOV

Simultaneous estimation of $\mu(x)$ and $D(x)$

50 Sensors – Measurements with σ = 0.01 Y_{max}





Acknowledgments

- The authors are grateful to the financial support of:
  - CAPES, CNPq and FAPERJ.
  - UTA, University of Texas at Arlington.
  - FIU, Florida International University.
  - UFRJ, Federal University of Rio de Janeiro.
  - IME, Military Institute of Engineering.