OPTIMAL CONTROL APPLICATIONS AND METHODS
Optim. Control Appl. Meth. 2008; **00**:1–6 *Prepared using* ocaauth.cls *[Version: 2002/11/11 v1.00]*

A convexity-based homotopy method for nonlinear optimization in model predictive control

Julian Bonilla*[1,2], Moritz Diehl[2], Bart De Moor[2] and Jan Van Impe[1]

[1] *Department of Chemical Engineering (CIT/BioTec), Katholieke Universiteit Leuven, B-3001 Leuven, Belgium.*
[2] *Department of Electrical Engineering (ESAT/SCD), Katholieke Universiteit Leuven, B-3001 Leuven, Belgium.*

SUMMARY

This paper presents a convexity-based homotopy solution procedure to the nonconvex optimal control problem (OCP) arising in model predictive control. The approach deals with OCP formulations where the dynamic system involved is control-affine. The non-convex OCP is modified by introducing a penalized pseudo control term and a homotopy parameter which gradually transforms the original problem making it convex. The method solves first this convex formulation and uses the result to initialize the solution of the next problem on the homotopy path, recovering gradually the original OCP. The proposed methodology is evaluated for the benchmark control problem of an isothermal chemical reactor with Van de Vusse reactions and input multiplicity, where the method finds a global solution while local optimization techniques lead to a local minimum. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Optimal control, Convex reformulations, Homotopy methods

1. INTRODUCTION

Non-convex OCPs are of particular interest in the control community since control problems with embedded nonlinear ordinary differential equations (ODEs) are inherently non-convex. Because non-convex OCPs are generally difficult to solve, researchers develop methods and new formulations in order to deal with them efficiently and accurately. One of the features of these kind of problems is that they exhibit local minima, therefore derivative-based optimization techniques can easily lock on to a local solution if the problem is not initialized appropriately.

*Correspondence to: J. Bonilla, Katholieke Universiteit Leuven, BioTeC - Bioprocess Technology and Control, W. de Croylaan 46, B-3001 Leuven, Belgium . *Julian.Bonilla@cit.kuleuven.be*

Contract/grant sponsor: Research Council KUL: GOA AMBioRICS, CoE EF/ Views: 2660205/006 Optimization in Engineering(OPTEC), IOF-SCORES4CHEM, several PhD/postdoc & fellow grants, Flemish Government: -FWO: PhD/postdoc grants, projects -IWT: PhD Grants, McKnow-E, Eureka-Flite; Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011; EU: ERNSI; contract/grant number: G.0452.04,G.0499.04,G.0211.05,G.0226.06,G.0321.06,G.0302.07

Received 16 January 2008
Revised

Copyright © 2008 John Wiley & Sons, Ltd.

The classical global optimization approaches to deal with non-convex OCPs are based on relaxation techniques which convexify the problem [1]. The use of Branch and Bound methods (B&B) [2] combined with the convex relaxation of Bolza-type functionals has been also proposed for the solution of non-convex OCPs [3]. However, there is still a lot of work in improving the computational demand of B&B methods and finding tighter convex relaxations to functionals with embbeded ODEs [4]. On the other hand, mathematicians and practitioners have recently realized that several nonconvex optimization problems can be reformulated such that the new formulation exhibits convexity [5]. The advantages of using convex formulations lie not only in the fact that local solutions are global, but also in that they present polynomial-time convergence, and efficient and reliable methods to solve such convex problems are well developed [6].

This work addresses the task of solving non-convex OCPs employing the model structure. We have noticed that the convergence to global optima can be improved by first solving a related convex formulation, which is connected by a homotopy path to the original problem, and using the solution to initialize the original OCP. The proposed technique resembles a *continuation method* for global optimization [7] where, by means of filtering techniques, the original cost is gradually transformed into a smoother function with fewer local minimizers. An optimization algorithm is then applied to the transformed function, tracing the minimizers back to the original cost.

The paper is organized as follows: in Section 2 the basis of dynamic optimization in the *Nonlinear Model Predictive Control* (NMPC) framework is addressed, Section 3 deals with the proposed approach to solve non-convex OCPs through a particular model structure; Section 4 illustrates the applicability of the method to parameter estimation and optimal control problems; Section 5 shows the results of a closed loop simulation for a NMPC applied to a chemical reactor with input multiplicity, a discussion of the results follows.

2. NONLINEAR DYNAMIC OPTIMIZATION AND PREDICTIVE CONTROL

NMPC is a technique based on the online solution of an open loop optimal control problem over a finite time horizon. Figure 1 depicts the main signals and parameters in a NMPC strategy.

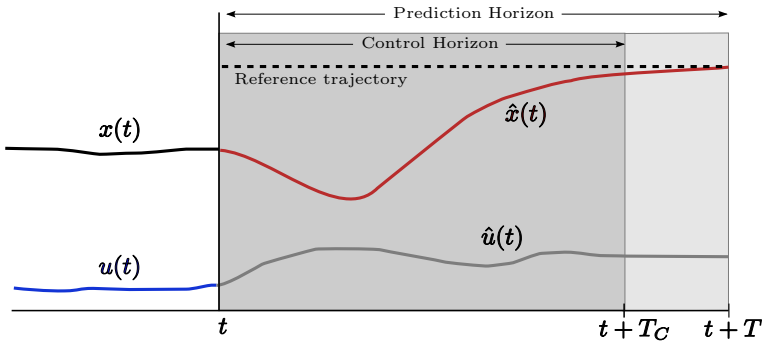

Figure 1. Model predictive control scenario

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

In figure 1 at the present time, t, the future behavior of the process states, $\hat{x}(t)$, is predicted over a time horizon, T, based on the state information, $x(t)$, and an available dynamic model of the process. The term NMPC generally refers to the formulation of predictive control using nonlinear models. Based on the prediction of the process behavior, a cost function is evaluated and optimized subject to a set of constraints imposed by the application and the designer. A future control scenario $\hat{u}(t)$ is calculated by solving the optimization problem. Ideally, there would be no process disturbances or model uncertainties and the OCP could be solved for infinite horizon, then the optimized control signal $\hat{u}(t)$ could be applied to the process like an open loop control. However, this is obviously unrealistic and some feedback needs to be incorporated. Therefore, the optimal control $\hat{u}(t)$ is implemented until a new measurement or estimate of the states becomes available. At that point, a new prediction is executed and a new OCP is formulated based on the current state shifting the entire procedure one step ahead. Thus, model predictive control is often referred to as *receding horizon control*.

The general algorithm for NMPC can be summarized in the following steps :

1. Observe the process in order to obtain the estimate of the current states and disturbances.
2. Calculate the optimal control trajectory $\hat{u}(t)$ over the time interval T by minimizing the proposed cost subject to the problem constraints.
3. Implement the optimal control open loop until a new state estimate becomes available
4. Return to 1

The dynamic optimization problem in NMPC can be formulated as follows:
Consider the general nonlinear dynamic system

$$\begin{aligned} \dot{x}(t) &= f(x,u,t), && (1)\\ y(t) &= h(x,u,t), && (2)\\ x(0) &= x_0. && (3) \end{aligned}$$

with a set of constraints on the states and controls $\mathbb{X}(t) \in \Re^n$ and $\mathbb{U}(t) \in \Re^u$ which define the feasible region for the OCP, $\mathbb{X}(t) \times \mathbb{U}(t)$. Several assumptions can be made concerning the sets $\mathbb{U}$ and $\mathbb{X}$ which make the OCP problem less difficult to solve. For instance, box constraints of the form

$$\begin{aligned} \mathbb{X} &= \{x(t) \in \Re^n \mid x^{\min} \leq x(t) \leq x^{\max}\} \; \forall \; t \in \; [0,T], && (4)\\ \mathbb{U} &= \{u(t) \in \Re^u \mid u^{\min} \leq u(t) \leq u^{\max}\} \; \forall \; t \in \; [0,T] \end{aligned}$$

are common because actuators have limited ranges. They introduce convex constraints [9] to the nonlinear optimal control problem.

The cost to be optimized can be formulated, in general, as a Bolza-type objective over the time horizon T

$$J(x,u) = \int_0^T L(x,u,t)\,dt + E(x(T)), \qquad (5)$$

where L is the Lagrange term and is often a function of the prediction errors with respect to a desired trajectory. E is a terminal cost namely the Mayer term and can be used to enforce stability as an alternative to a terminal constraint [8].

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

By following the above considerations, the OCP is formulated as

$$\min_{x(\cdot),u(\cdot)} J(x,u), \tag{6}$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(x,u,t), \quad t \in [0,T], && (7)\\ x(0) &= x_0, && (8)\\ x(t) &\in \mathbb{X}(t), \quad t \in [0,T], && (9)\\ u(t) &\in \mathbb{U}(x,t), \quad t \in [0,T], && (10)\\ x(T) &\in \mathbb{X}_T. && (11) \end{aligned}$$

Since the continuous time OCP is an infinite dimensional optimization problem, in direct approaches the decision variables are parametrized so that the problem is formulated as a finite dimensional nonlinear programming problem (NLP). The continous time formulation for the NMPC algorithm is presented here, assuming that controls are parametrized by piecewise constant functions and the model is discretized using finite differences or higher order schemes for its numerical solution. Although, for our problem class, the direct single shooting technique is used due to its implementation simplicity, there are more efficient approaches such as collocation [10] and multiple shooting [11].

3. A CONVEXITY-BASED HOMOTOPY

In this section the proposed method is presented, mainly focus on optimal control formulations but showing that it can be easily extended to the parameter estimation problem.

3.1. The optimal control problem

Let us present the problem we are interested in by specifying further the OCP introduced previously. The particular OCP is formulated in the form

$$\min_{x(\cdot),u(\cdot)} \int_0^T \left(\| x(t) - x^{\mathrm{ref}}(t) \|_Q^2 + l(x,u,t) \right) dt + E(x(T)), \tag{12}$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(x,t) + g(x,t)u(t), \quad t \in [0,T], && (13)\\ x(0) &= x_0, && (14)\\ x(t) &\in \mathbb{X}(t), \quad t \in [0,T], && (15)\\ u(t) &\in \mathbb{U}(x,t), \quad t \in [0,T], && (16)\\ x(T) &\in \mathbb{X}_T. && (17) \end{aligned}$$

We assume that the optimization task is described by a desired state trajectory $x^{\mathrm{ref}}(t)$ ($x^{\mathrm{ref}(t)} \in C^1$). The reference trajectory need not necessarily be feasible with respect to the model equation (13). However, we consider $x^{\mathrm{ref}}(t) \in \mathbb{X}(t)$, i.e. the desired trajectory is feasible with respect to the inequality constraints.

Let us now make the following extra assumptions:

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

Assumptions

1. The set $\mathbb{U}(x^{\mathrm{ref}}(t),t)$ is convex for all $t \in [0,T]$.
2. The function $l(x^{\mathrm{ref}}(t),u,t)$ is for all $t \in [0,T]$ convex in $u \in \mathbb{U}(x^{\mathrm{ref}}(t),t)$.
3. The state penalization matrix Q in $\parallel x(t) - x^{\mathrm{ref}}(t) \parallel_Q^2$ is positive definite.

Remark 1
The above assumptions are e.g. satisfied if the controls and states are constrained by simple bounds, thus $\mathbb{U} \times \mathbb{X}$ is a hypercube, and if the Lagrange term L in the NMPC cost is given by the usual least squares objective $L(x,u,t) = \|x(t) - x^{\mathrm{ref}}(t)\|_Q^2 + \|u(t) - u^{\mathrm{ref}}(t)\|_R^2$.

Let us now introduce a pseudo control $v(t)$ and a scalar parameter $\lambda \in (0,1)$ that interpolates between the original problem ($\lambda \to 1$) and one of its possible homotopies ($\lambda \to 0$) as follows[†]:

$$P_u(\lambda): \min_{v_0,v(\cdot),x(\cdot),u(\cdot)} \int_0^T \left(\frac{1}{\lambda} \parallel x(t) - x^{\mathrm{ref}}(t) \parallel_Q^2 + l(x,u,t) \right) dt + \frac{1}{\lambda} E(x(T)) + \frac{1}{(1-\lambda)} \left(\int_0^T \parallel v(t) \parallel dt + \parallel v_0 \parallel \right) \tag{18}$$

subject to

$$\begin{aligned}
\dot{x}(t) &= f(x,t) + g(x,t)u(t) + v(t), \quad t \in [0,T], && (19)\\
x(0) &= x_0 + v_0, && (20)\\
x(t) &\in \mathbb{X}(t), \quad t \in [0,T], && (21)\\
u(t) &\in \mathbb{U}(x(t),t), \quad t \in [0,T], && (22)\\
x(T) &\in \mathbb{X}_T. && (23)
\end{aligned}$$

Lemma 1
Following the assumptions, if $\lambda \to 0$, the above parametric optimization problem $P(\lambda)$, (18)-(23), is equivalent to a convex optimization problem. If $\lambda \to 1$ its solution approaches the solution of the original problem (12)-(17).

Clearly, the problem formulated in (18)-(23) leads to the original OCP when λ goes to one. Following this condition, $v(t)$ is strongly penalized forcing the equality constraints defined by the model and the initial condition to recover their original form (13) and (14).

On the other hand, when λ goes to zero in (18)-(23), the term $\parallel x(t) - x^{\mathrm{ref}}(t) \parallel_Q^2$ is penalized such that it tends to zero and the optimization is reduced to

$$P_u(0): \quad \min_{v_0,v(\cdot),u(\cdot)} \int_0^T \left(\parallel v(t) \parallel + l(x^{\mathrm{ref}},u,t) \right) dt + \parallel v_0 \parallel \tag{24}$$

[†]Other forms of constructing these homotopic maps are also possible

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

subject to

$$\begin{aligned} \dot{x^{\mathrm{ref}}}(t) &= f(x^{\mathrm{ref}},t) + g(x^{\mathrm{ref}},t)u(t) + v(t), \quad t \in [0,T], && (25)\\ x^{\mathrm{ref}}(0) &= x_0 + v_0, && (26)\\ u(t) &\in \mathbb{U}(x(t),t), \quad t \in [0,T]. && (27) \end{aligned}$$

Consequently, the obtained OCP problem is convex since the penalty term was selected to be a convex function of the optimization variables and the equality constraint defined by the model is affine in the optimization variables $u(t)$ and $v(t)$.

In this context, we propouse to solve the non-convex OCP with the given structure, by convexifying it through the formulation in (18)-(23) with $\lambda \to 0$. The solution of the convex homotopy is used to initialize successive OCPs moving λ towards 1 in order to recover the original formulation. Numerical examples illustrate how the method can be applied in the predictive control and parameter estimation frameworks in section 4.

3.2. Parameter estimation problem

The convex-homotopy based method resembles some smoothing methods for parameter estimation where the model is fitted in terms of some function expansion [12]. Here, the presented approach for OCPs is extended to the parameter estimation problem (PEP) where the dynamic model exhibits a special structure in the parameters to be estimated, i.e the system is affine in parameters.

Consider a PEP where it is desired to fit the dynamic behavior of the states $x(t)$ to a set of measurements $x^{\mathrm{m}}(t)$:

$$\min_{x(\cdot),p} \int_0^T \parallel x(t) - x^{\mathrm{m}}(t) \parallel_Q^2 \, dt \, + \, V(p) \qquad (28)$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(x,u,t) + g(x,u,t)p, \quad t \in [0,T], && (29)\\ x(t) &\in \mathbb{X}(t), \quad t \in [0,T], && (30)\\ p &\in \mathbb{P}. && (31) \end{aligned}$$

Let us assume that $\mathbb{P}$ is a convex set, $x^{\mathrm{m}}(t) \in \mathbb{X}(t)$ and $V(p)$ represents a convex function $\forall\, p \in \mathbb{P}$. Following the same approach presented in the OCP case, the PEP can be reformulated by introducing a pseudo control $v(t)$ and a homotopy parameter λ

$$P_p(\lambda): \quad \min_{v(\cdot),x(\cdot),p} \frac{1}{\lambda} \int_0^T \parallel x(t) - x^{\mathrm{m}}(t) \parallel_Q^2 \, dt \, + \, V(p) \, + \, \frac{1}{1-\lambda} \int_0^T \parallel v(t) \parallel \, dt \qquad (32)$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(x,u,t) + g(x,u,t)p + v(t), \quad t \in [0,T], && (33)\\ x(t) &\in \mathbb{X}(t), \quad t \in [0,T], && (34)\\ p &\in \mathbb{P}. && (35) \end{aligned}$$

Since the dynamic system in (29) is assumed to be parameter-affine, the procedure presented in the OCP case can be applied. Clearly, when $\lambda \to 1$ in (32)-(35), the original problem is

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

recovered. On the other hand, when $\lambda \to 0$ the term $\| x(t) - x^{\mathrm{m}}(t) \|_Q^2$ is penalized strongly leading to $x(t) = x^{\mathrm{m}}(t)$ and allowing the formulation of the convex problem

$$\min_{v(\cdot),p} \int_0^T (\| v(t) \|)\, dt + V(p) \tag{36}$$

subject to

$$\dot{x}^{\mathrm{m}}(t) = f(x^{\mathrm{m}}, u, t) + g(x^{\mathrm{m}}, u, t)p + v(t), \quad t \in [0, T], \tag{37}$$

$$p \in \mathbb{P}. \tag{38}$$

Consequently, it is possible to use convex optimization methods to find a good initial guess for the original PEP by solving the related convex problem in (36)-(38) and use the result to initialize the subsequents PEP on the homotopy path with $\lambda \to 1$ in order to recover the original formulation. The advantages of this approach for a nonconvex PEP are illustrated in the following section.

4. CONVEXITY IN NUMERICAL EXAMPLES

In order to illustrate the applicability of the proposed method to the solution of dynamic optimization problems, we consider two study cases where the nonlinear dynamic models involved are affine in terms of the optimization variables.

4.1. Parameter Estimation Problem

Let us analyze the dynamic system

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -p & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = f(x, p, t), \tag{39}$$

which correspond to a *Harmonic Oscillator*. Notice that this autonomous system is affine in the parameter to be estimated.

Consider the optimization problem to estimate the parameter p with a Lagrange term defined in the least square sense and a vanishing Mayer term. The solution of the problem aims to match a given set of measurements $x^{\mathrm{m}}(t)$ with the model states $x(t)$ over a time interval $[0\ T]$. By defining $x(t) = [x_1(t)\ x_2(t)]^T$, the PEP is formulated as

$$\min_{x(.),p} \int_0^T \| x(t) - x^{\mathrm{m}}(t) \|_2^2\, dt, \tag{40}$$

subject to

$$\dot{x}(t) = f(x, p, t), \quad t \in [0, T]. \tag{41}$$

The measurement $x^{\mathrm{m}}(t)$ is contaminated with harmonics of the fundamental frequency $(\sqrt{p})$ in order to emphasize the presence of local minima. Figure 2 shows the non-convex cost for the problem formulated in (40)-(41). Clearly, the success of an optimization algorithm for this case depends strongly on a good initialization.

Let us to recast the optimization problem by introducing the penalty term and the proposed approach,

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

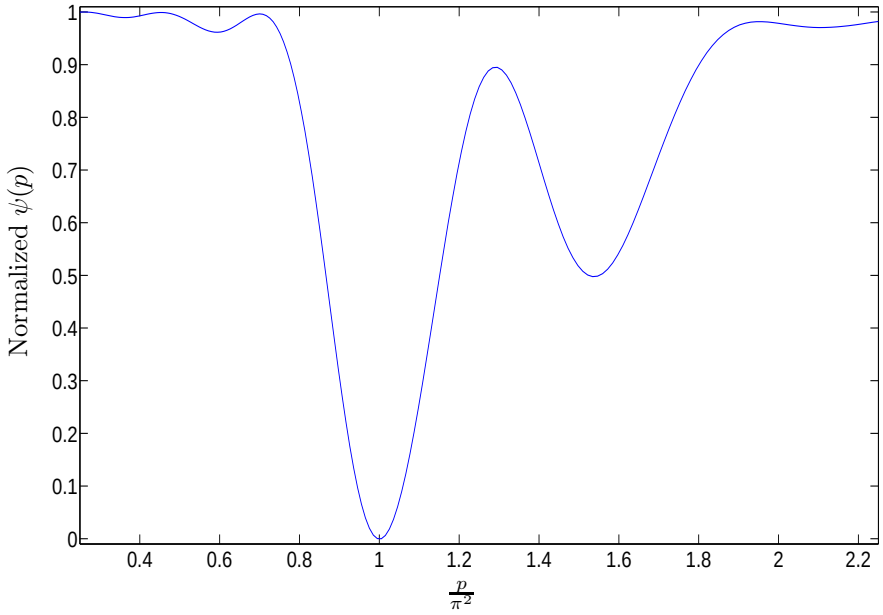

Figure 2. Non-convex cost $\psi(p)$ arising from the parameter estimation problem, when optimized over $x(.)$: $\psi(p) = \min_{x(.)} \int_0^T \parallel x(t) - x^{\mathrm{m}}(t) \parallel_2^2 \, dt$ s.t. $\dot{x} = f(x, p, t)$.

$$\min_{x(.),v(.),p} \int_0^T \left(\frac{1}{\lambda} \parallel x(t,p) - x^{\mathrm{ref}}(t) \parallel_2^2 + \frac{1}{(1-\lambda)} \parallel v(t) \parallel_2^2 \right) dt \tag{42}$$

subject to

$$\dot{x}(t) \quad = \quad f(x, p, t) + v(t) \quad t \in [0, T]. \tag{43}$$

Figure 3 shows the cost to minimize when the parameter λ moves between zero and one. This plot is obtained by minimizing first over x in such a way that the cost $\psi(x^*, p, \lambda)$ in figure 3 is already optimal in x. Small values of λ lead to a convex problem, while values of λ close to one lead to the original cost. Consequently, the original problem can be solved by solving first the convex reformulation which provides a global solution and a good initial guess for the original non-convex PEP.

4.2. Optimal Control Problem

Many nonlinear dynamic systems of practical interest present the affine structure which enables the applicability of the method, examples of this arise in the chemical industry such as distillation columns and continuous stirred tank reactors (CSTR) and mechanical engineering like in car maneuvering and robot arm manipulators.

Consider the benchmark problem presented in [13], where it is anylized an isothermal CSTR

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

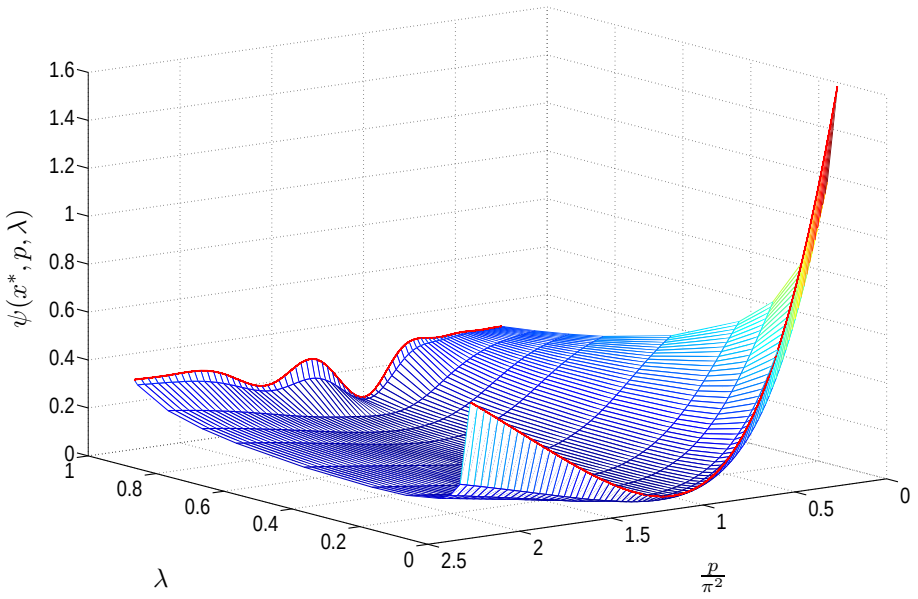

Figure 3. Cost function for all problems between the original problem and the convex problem on the homotopy path for the parameter estimation example

with the Van de Vusse reactions

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C, \qquad 2A \xrightarrow{k_3} D. \tag{44}$$

The CSTR dynamics is governed by the set of nonlinear equations

$$\dot{C}_a(t) = \frac{F(t)}{V}(C_{a,0}(t) - C_a(t)) - k_1 C_a(t) - k_3 C_a^2(t), \tag{45}$$

$$\dot{C}_b(t) = k_1 C_a(t) - k_2 C_b(t) - \frac{F(t)}{V} C_b(t). \tag{46}$$

$F(t)$ represents the feed flow of the reactant A, V is the reactor volume, and $C_a(t)$ and $C_b(t)$ represent the concentrations of the reactant and the intermediate, respectively (see Figure 4). k_1, k_2 and k_3 are constant reaction rates. It is assumed that the reactor volume remains constant and the feed consists of pure A with a concentration $C_{a,0}(t)$. The intermediate concentration $C_b(t)$ is controlled by manipulating the feed rate.

Table I summarizes the parameters and nominal conditions for this benchmark problem. This particular system presents input multiplicity as illustrated in Figure 5. In consequence, the formulation of a cost in terms of the errors with respect to a desired trajectory can generate two possible control values. A local optimization technique can easily be attracted by a local minimum, delivering a suboptimal solution, if it is initialized inappropriately. However, for

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

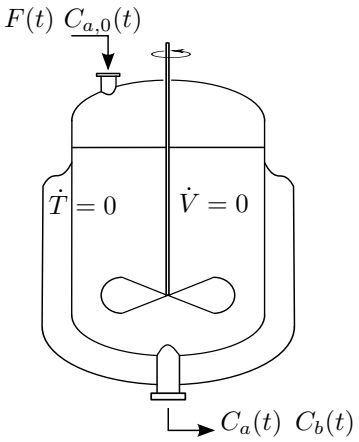

Figure 4. Isothermic CSTR

Table I. CSTR parameters and nominal conditions

Parameter	Value
k_1	50 1/h
k_2	100 1/h
k_3	10 L/(gmol.h)
V	1 L
$F_{\min}$	0 L/h
$F_{\max}$	200 L/h
$\bar{F}$	180.95 L/h
$\bar{C}_{a,0}$	10 gmol/L
$\bar{C}_a$	6.181 gmol/L
$\bar{C}_b$	1.1 gmol/L

this particular application, global optimization approaches are able to find a better solution as proposed in [14]. Our approach suggests a computationally lesss demanding heuristic to address the problem of finding a global minima.

Let us define $x(t) = [C_a(t)\ C_b(t)]^T$ and $u(t) = \frac{F(t)}{V}$. The OCP for the CSTR output concentration can be formulated as

$$\min_{x(.),u(.)} \int_0^T \left(\|x(t) - x^{\text{ref}}(t)\|_Q^2 + \|u(t) - u^{\text{ref}}(t)\|_R^2 \right) dt \tag{47}$$

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

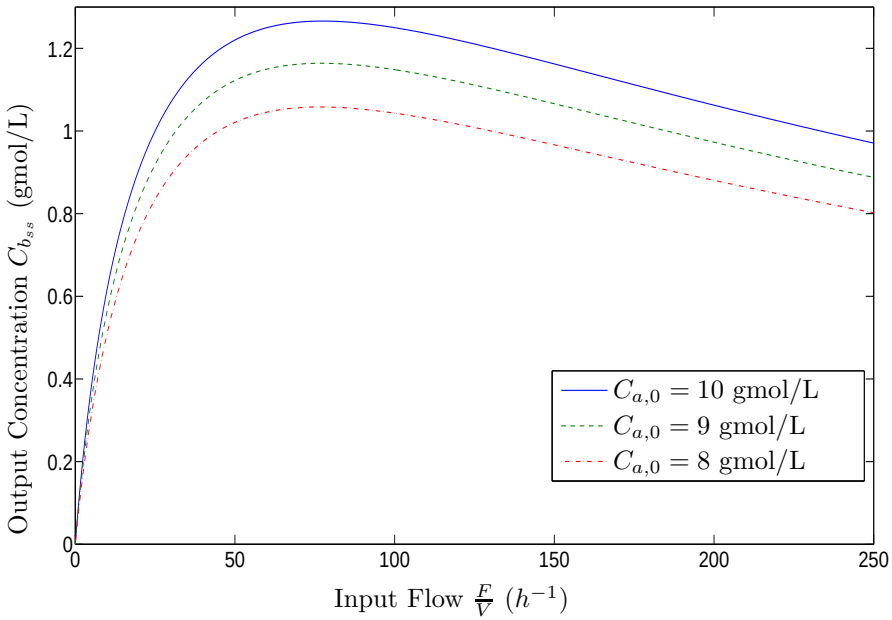

Figure 5. Static Gain for the CSTR process

subject to

$$x(0) = x_0 \tag{48}$$
$$\dot{x}(t) = f(x,t) + g(x,t)u(t) \quad t \in [0,T], \tag{49}$$
$$x(t) \in \mathbb{X}, \quad t \in [0,T], \tag{50}$$
$$u(t) \in \mathbb{U}, \quad t \in [0,T], \tag{51}$$

where $x^{\text{ref}}(t)$ and $u^{\text{ref}}(t)$ represent the state and input trajectories to be followed over the time horizon $[0,\ T]$.

The presented optimal control problem is solved for a prediction horizon $T = 0.06$ h and discretized with a sampling rate of 0.002 h (7.2 s) but the control is assumed constant over the entire time horizon as in [14]. The differences between model states and state trajectory are penalized in a different proportion along the time interval $[0,\ T]$. The discrete version of the penalization matrix Q is set to $diag(0.0001\ 0.01)$ until the end of the horizon where it is changed to $diag(0.0001\ 10000)$. The control penalization R is set to 100 along the entire time horizon.

Figure 6 shows the cost to be optimized when a step change on the reference trajectories $(x^{\text{ref}}(t),\ u^{\text{ref}}(t))$ is performed. The original problem has one global and one local solution, the last one outside of the feasible region. Clearly, if a local optimization technique starts with an initial guess close to the constraints, the solution to the original NLP locks on to the upper bound of $u(t)$ in its way to reach the local minimum.

The previous OCP is modified using the proposed approach by introducing the penalty term

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

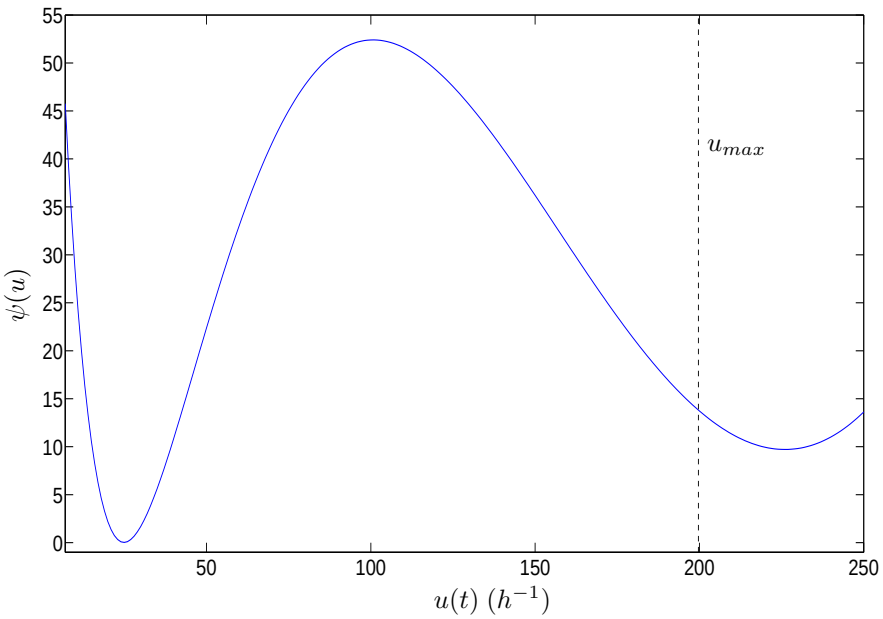

Figure 6. Original cost to be optimized for the CSTR open loop optimal control problem when optimized over $x(.)$: $\psi(u) = \min_{x(.)} \int_0^T \left(\| x(t) - x^{\text{ref}}(t) \|_2^2 + \|u(t) - u(t)^{\text{ref}}\|_2^2 \right) dt$ s.t. $\dot{x} = f(x,t) + g(x,t)u(t)$.

and the homotopy parameter λ leading to the OCP

$$\min_{x(.),u(.),v(.)} \int_0^T \left(\frac{1}{\lambda} \|x(t) - x^{\text{ref}}(t)\|_{\mathcal{Q}}^2 + \|u(t) - u^{\text{ref}}(t)\|_{\mathcal{R}}^2 \right) dt \; + \; \frac{1}{1-\lambda} \int_0^T \|v(t)\|_2^2 dt, \tag{52}$$

subject to

$$\begin{aligned} x(0) &= x_0 + v_0, && (53)\\ \dot{x}(t) &= f(x,t) + g(x,t)u(t) + v(t), \quad t \in [0,T], && (54)\\ x(t) &\in \mathbb{X}, \quad t \in [0,T], && (55)\\ u(t) &\in \mathbb{U}, \quad t \in [0,T]. && (56) \end{aligned}$$

Notice that only the non-convex terms in the cost are penalized with the homotopy parameter. Figure 7 shows the homotopic maps for different values of λ in the range [0.9 1]. For values below 0.9, it has been graphically noticed that the problem is already convex. It is remarkable that the optimum for the convex homotopy ($u^* = 26.95h^{-1}$) is quite close to the global optimum for the original problem ($u^* = 25h^{-1}$). Consequently, it is possible to attain the global optimum for this application, if the OCP is convexified using the proposed approach and the local technique is used for increasing values of λ using the previous optimal solutions as initial guess.

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

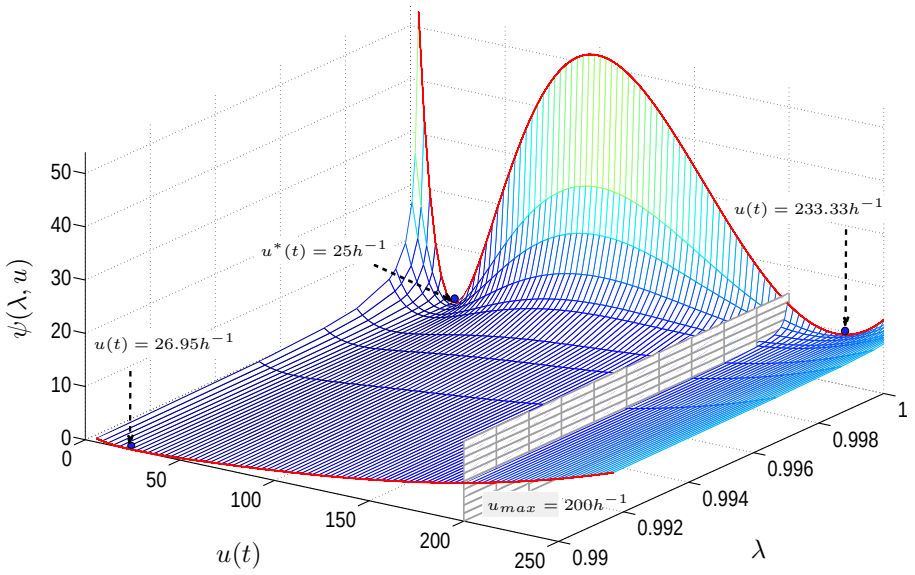

Figure 7. Cost functions for all the problems between the original OCP and its convex formulation on the homotopy path for the CSTR control example

5. CLOSED LOOP SIMULATIONS

In order to illustrate the applicability of the method in the NMPC framework, a closed loop simulation study was performed. Notice that in NMPC the NLP resulting from (47)-(51) has to be solved at each sampling instant.

A nonlinear open loop observer was implemented in order to estimate the possible process disturbances and model states. The state and input reference trajectories are calculated from the desired process output (intermediate concentration $C_b(t)$) and the estimated disturbance as proposed in [15]. At each sampling instant the problem is solved first using the homotopy approach with $\lambda = 0.9$ and after, the original NLP ($\lambda = 1$) is solved using as initial guess the solution provided by the convex homotopy. Consequently, two NLPs are solved every 7.2s, one to get a good initial guess and the original OCP which converges to the global minimum with that good initialization.

Figure 8 presents the results for a tracking and disturbance rejection test using both, a NLP solver based on sequential quadratic programming (SQP)[‡] and the proposed convexity-based homotopy approach. The simulation is performed for both strategies under the same conditions, however, the homotopic approach finds the global solution while the local optimization technique, which is each time initialized with the previous solution, finds a local optimum.

[‡]the MATLAB NLP solver was used for this simulation study

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

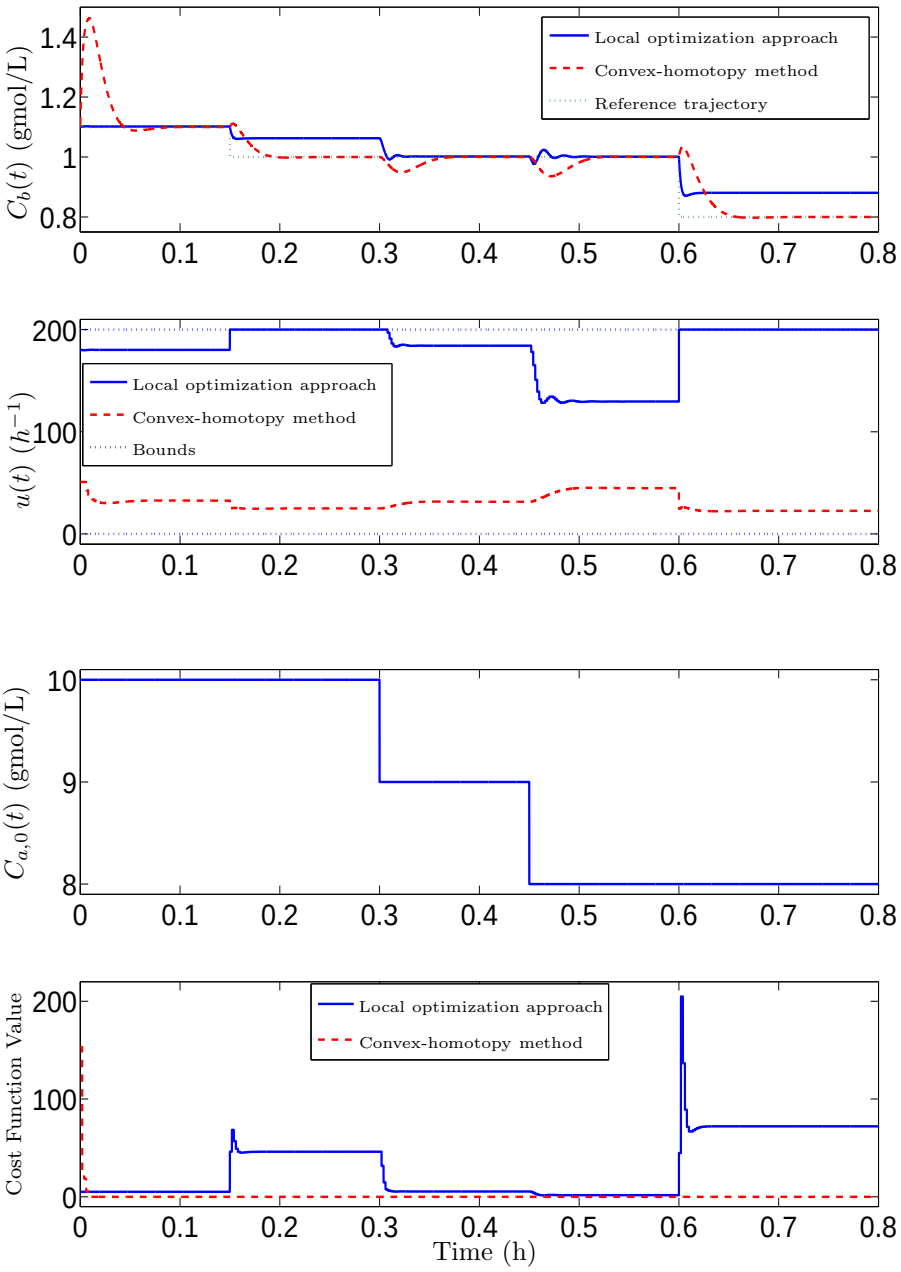

Figure 8. Tracking and disturbance test for the NMPC applied to the CSTR using a local optimizer and convexity-based homotopy approach

Copyright © 2008 John Wiley & Sons, Ltd.
Prepared using ocaauth.cls

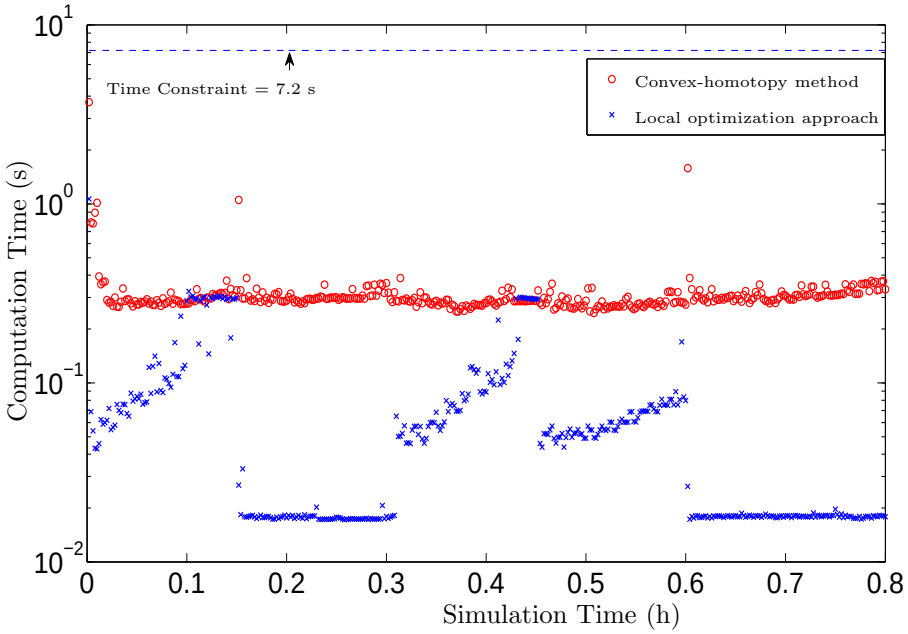

Figure 9. Computational demand over one simulation test

The output reference trajectory is changed at 0.15 h and 0.6 h, in both cases the solution to the original NLP, using the local optimization technique, locks on to upper bound of $u(t)$. On the other hand the homotopy approach finds the global solution as illustrated in the cost function plot. Additionally, the process is disturbed by changing the feed concentration at 0.3 h and 0.45 h, where the classical technique finds a local solution. Notice that for this cost formulation, a local minimum leads to steady state errors with respect to the reference trajectory.

For the 0.8 h simulation, 400 NLP problems were solved. The test was executed using a Intel Core 2 Duo microprocessor at 1.6 GHz running a linux distribution with 2 GB of RAM. Figure 9 presents the computational demand of the evaluated algorithms. It is remarkable that, for all the cases, the computational demand was below the time constraint of 7.2 s, making the approach feasible for implementation.

6. CONCLUSIONS

We have presented a convexity-based homotopy method to initialize non-convex NLPs arising in optimal control problems where the dynamic model involved is affine in the controls. It has been demonstrated that the method can be easily extended to NLPs for parameter estimation when the dynamic model is affine in the parameters to be estimated. We have investigated the applicability of the method to two particular examples in the field of parameter estimation

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls

and optimal control where the proposed approach provides an initial guess to the original NLP, leading to the global optimum. The validity of the method for NMPC applications has been illustrated through a benchmark problem in chemical process control where the approach was able to find the optimal control trajectory while a classical technique locks on to a local solution. Finally, computational aspects of the method are addressed in order to illustrate its feasibility for the CSTR optimal control problem. In summary, our new approach promises to become an attractive heuristic for smart initialization of non-convex optimal control problems, using the power of convex optimization.

REFERENCES

1. McCormick G. Computability of global solutions to factorable nonconvex programs: part I - convex underestimating problems. *Mathematical Programming* 1976; **10**(1976):147-175.
2. Horst R, Tuy H. *Global Optimization: Deterministics Approaches* (3rd edn) Springer-Verlag: Berlin, 1996.
3. Singer AB, Barton PI. Global optimization with nonlinear ordinary differential equations. *Journal of Global Optimization* 2006; **34**(2006):159–190.
4. Chachuat B, Singer AB, Barton PI. Global Solution of Dynamic Optimization and Mixed-Integer Dynamic Optimization. *Ind. Eng. Chem. Res.* 2006; **45**(25):8373–8392.
5. Boyd S, Crusius C, and Hansson A. Control applications of nonlinear convex programming. *Journal of Process Control* 1998; **8**(5-6):313–324.
6. Nesterov Y, Nemirovski A. *Interior Points Polinomial Methods in Convex Programming* (1st edn). SIAM Publications: Philadelphia, 1995.
7. Moré J, Wu Z. Global continuation for distance geometry problems. *SIAM journal on optimization* 2006; **7**(3):814–836.
8. Allgöwer F, Findeisen R, Nagy Z. Nonlinear model predictive control: from theory to application. *J. Chin. Inst. Chem. Engrs* 2004; **35**(3):299–315.
9. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press, Cambridge, 2006.
10. Biegler L. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Comput. Chem. Engng.* 1984; **8**(1984):243–248.
11. Bock H, Plitt K. A multiple shooting algorithm for direct solution of optimal control problems. *In Proc.9th IFAC world congress* 1984;
12. Ramsay J, Hooker G, Campbell D, Cao J. Parameter estimation for differential equations: a generalized smoothing approach. *J. R. Statist. Soc. B.* 2007; **69**(5):741–796.
13. Sistu PB, Bequette BW. Model predictive control of processes with input multiplicities. *Chemical Engineering Science* 1995; **50**(6):921–936.
14. Long C, Polisetty P, Gatzke E. Nonlinear model predictive control using a global optimization approach. *Journal of Process Control* 2006; **16**(2006):365–643.
15. Rossiter JA. *Model-Based Predictive Control, A practical Approach*. CRC Press: 2003.

Copyright © 2008 John Wiley & Sons, Ltd.

Prepared using ocaauth.cls