This document contains the **post-print pdf-version** of the refereed paper:

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

by *Julián Bonilla, Moritz Diehl, Filip Logist, Bart De Moor, and Jan F. M. Van Impe*

which has been archived in the university repository Lirias (<https://lirias.kuleuven.be/>) of the Katholieke Universiteit Leuven.

The content is identical to the content of the published paper, but without the final typesetting by the publisher.

When referring to this work, please cite the full bibliographic info:

J. Bonilla, M. Diehl, F. Logist, B. De Moor, J. Van Impe. 2010. A convexity-based homotopy method for nonlinear optimization in model predictive control. Optimal Control Applications and Methods, 31, 393–414.

The journal and the original published paper can be found at:
<http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291099-1514>
<http://onlinelibrary.wiley.com/doi/10.1002/oca.945/abstract>

The corresponding author can be contacted for additional info.

Conditions for open access are available at:
<http://www.sherpa.ac.uk/romeo/>

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

Julián Bonilla^{1,2}, Moritz Diehl², Filip Logist¹, Bart De Moor² 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 non-convex optimal control problems (OCPs) arising in model predictive control. The approach deals with a special class of OCP formulations where the dynamic system involved is control-affine and the objective is a penalty on deviations from a state reference trajectory. The non-convex OCP is modified by introducing a penalized pseudo state and a homotopy parameter which gradually transforms the original problem into a convex one. The method solves first this convex formulation and uses the result to initialize the solution of the next problem on the zero path, recovering 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. For the simple case with control horizon one, the method is able to find the global solution due to the convex initialization, while local optimization techniques only lead to a local minimum. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS: Optimal control, Convex reformulation, Homotopy methods

1. INTRODUCTION

Non-convex optimal control problems (OCPs) are of particular interest in the control community since control problems with embedded nonlinear ordinary differential equations

*Correspondence to: Jan Van Impe, Katholieke Universiteit Leuven, BioTeC - Bioprocess Technology and Control, W. de Croylaan 46, B-3001 Leuven, Belgium. jan.vanimpe@cit.kuleuven.be

Contract/grant sponsor: Research Council KUL: OT/09/025/TBA, GOA AMBioRICS, GOA MaNet, CoE EF/05/006 (OPTEC), IOF-SCORES4CHEM, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects: G0226.06, G0321.06, G.0302.07, G.0320.08, G.0558.08, G.0557.08, G.0588.09 research communities (ICCoS, ANMMM, MLDM); G.0377.09 IWT: PhD Grants, Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, O & O-Dsquare Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO 2007-2011); EU: ERNSI; FP7-HD-MPC (INFSO-ICT-223854), COST intelliCIS, FP7-EMBOCON (ICT-248940) Contract Research: AMINAL Other: Helmholtz: viCERP ACCM Bauknecht Hoerbiger.; contract/grant number: Bart De Moor and Jan Van Impe are full professors at the Katholieke Universiteit Leuven, Belgium. Moritz Diehl is professor at the Katholieke Universiteit Leuven. Jan Van Impe holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia.

(ODEs) are inherently non-convex. They arise in each Nonlinear Model Predictive Control (NMPC) routine where an OCP is solved at each sampling instant. 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 onto a local solution if the problem is not initialized appropriately.

The classical global optimization approaches to deal with non-convex OCPs are based on relaxation techniques which convexify the problem [1]. The use of spatial Branch and Bound methods (B&B) [2] combined with the convex relaxation of Bolza- type functionals has been also proposed for the solution of finite dimensional 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 embedded ODEs [4]. On the other hand, mathematicians and practitioners have realized that several non-convex 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 exhibit polynomial-time convergence, and that efficient and reliable methods to solve such convex problems are well developed [6].

This work proposes a method to solve nonconvex OCP with a particular structure in the cost and constraints. In particular, a control-affine model structure is assumed together with an objective which penalizes deviations from a desired reference trajectory. Many nonlinear dynamic systems of practical interest present the control-affine structure which enables the applicability of the method. Examples of these kind of systems arise for instance in the chemical industry (e.g., distillation columns and continuous stirred tank reactors -CSTRs) and mechanical engineering (e.g., car maneuvering and robot arm manipulators). The convergence to the global optimum, for these OCPs, can be improved by first solving a related convex formulation, which is connected by a homotopy path to the original problem, and using this 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 introduces the proposed approach to solve non-convex OCPs. Section 4 illustrates the applicability of the method to parameter estimation and optimal control problems. Section 5 presents the results of a closed loop simulation for an NMPC applied to a chemical reactor with input multiplicity. Finally, Section 6 summarizes the main conclusions.

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 classic NMPC

Figure 1. Model predictive control scenario.

strategy [8].

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 computed by solving the optimization problem. Ideally, there would be no process disturbances or model uncertainties and the OCP could be solved for infinite horizon. In this case, the optimized control signal $\hat{u}(t)$ could be applied to the process like an open loop control. However, this situation 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. Hence, 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 Step 1.

The dynamic optimization problem in NMPC can be formulated as follows. Consider the general nonlinear dynamic system with differential states $x(t)$ and controls $u(t)$:

$$\dot{x}(t) = f(x(t), u(t)), \quad (1)$$

$$x(0) = x_0. \quad (2)$$

A set of constraints imposed on the states and controls $\mathbb{X}$ and $\mathbb{U}$ respectively, defines the feasible region for the OCP, $\mathbb{X} \times \mathbb{U}$. Several assumptions can be made concerning the sets $\mathbb{U}$

and $\mathbb{X}$ in order to make the OCP problem less difficult to solve. For instance, box constraints defined by time invariant sets of the form

$$\begin{aligned}\mathbb{X} &= \{x \in \mathbb{R}^{n_x} \mid x^{\min} \leq x \leq x^{\max}\} \\ \mathbb{U} &= \{u \in \mathbb{R}^{n_u} \mid u^{\min} \leq u \leq u^{\max}\}\end{aligned}\quad (3)$$

are common because process inputs and outputs have limited ranges. These constraints introduce a convex feasible set [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 , as the one used in (4). There, L is the Lagrange term which is often a function of the prediction errors with respect to a desired trajectory. The terminal cost or Mayer term E can be used to enforce stability as an alternative to a terminal constraint. By following the above considerations, the general OCP is formulated as:

$$\min_{x(\cdot), u(\cdot)} \int_0^T L(x(t), u(t)) dt + E(x(T)), \quad (4)$$

subject to

$$\dot{x}(t) = f(x(t), u(t)), \quad t \in [0, T], \quad (5)$$

$$x(0) = x_0, \quad (6)$$

$$x(t) \in \mathbb{X}, \quad t \in [0, T], \quad (7)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (8)$$

$$x(T) \in \mathbb{X}_T. \quad (9)$$

3. A CONVEXITY-BASED HOMOTOPY

In this section the proposed method is presented. Although the method's main focus lies on control problems with control-affine process models, it is also shown that it can be easily extended to the parameter estimation case where the process model is affine in the parameters.

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, for a control-affine process model, is formulated in the form:

$$\text{OCP: } \min_{x(\cdot), u(\cdot)} \int_0^T \|x(t) - x^{\text{ref}}(t)\|_Q^2 dt + \|x(T) - x^{\text{ref}}(T)\|_{Q_T}^2, \quad (10)$$

subject to

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad t \in [0, T], \quad (11)$$

$$x(0) = x_0, \quad (12)$$

$$x(t) \in \mathbb{X}, \quad t \in [0, T], \quad (13)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (14)$$

$$x(T) \in \mathbb{X}_T. \quad (15)$$

We assume that the optimization task involves a desired state trajectory, $x^{\text{ref}}(t)$ ($x^{\text{ref}}(t) \in C^1$) and no knowledge of a desired input trajectory $u^{\text{ref}}(t)$ is provided. Moreover, we consider $x^{\text{ref}}(t) \in \mathbb{X}$, i.e., the desired state trajectory is feasible with respect to the inequality constraints.

Let us now make the following additional assumptions.

Assumption 1

1. The sets $\mathbb{U}$, $\mathbb{X}$ and $\mathbb{X}_T$ are convex.
2. The state penalization matrix Q is positive definite and Q_T is positive semi-definite.
3. $f(x(t))$ and $g(x(t)) \in C^1$ and the solution of $\dot{x}(t) = f(x(t)) + g(x(t))u(t)$ is uniquely determined by $u(t)$ and $x(0)$.

Remark 1

Convexity in the sets $\mathbb{X}$ and $\mathbb{U}$ is, e.g., satisfied if the controls and states are constrained by simple bounds, thus $\mathbb{U} \times \mathbb{X}$ is a hypercube.

In order to simplify the notation, we introduce the following norm:

$$\|x\|_J^2 = \int_0^T x(t)^T Q x(t) dt + x(T)^T Q_T x(T). \quad (16)$$

Additionally, as $Q \succ 0$, we say $x = y$ a.e. if $\|x - y\|_J^2 = 0$.

We present the homotopy formulation by introducing a pseudo state $x_c(t) \in \mathbb{R}^{n_x}$ and a scalar parameter $\lambda \in (0, 1)$ that interpolates between the original problem ($\lambda \rightarrow 1$) and one of its possible homotopies ($\lambda \rightarrow 0$) as follows:

$$P(\lambda) : \min_{x(\cdot), u(\cdot), x_c(\cdot)} \frac{1}{\lambda} \|x - x^{\text{ref}}\|_J^2 + \frac{1}{1 - \lambda} \|x_c - x\|_J^2 \quad (17)$$

subject to

$$\dot{x}_c(t) = f(x(t)) + g(x(t))u(t), \quad t \in [0, T], \quad (18)$$

$$x_c(0) = x_0, \quad (19)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (20)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (21)$$

$$x_c(T) \in \mathbb{X}_T. \quad (22)$$

This augmented parametric OCP defines a family of optimization problems as a function of the parameter λ . Other forms of homotopy maps are also possible where the parameter enters linearly or nonlinearly in the problem. However, our choice here is twofold. On the one hand we restrict the homotopy parameter to lie in the interval $(0, 1)$. On the other hand, there is an equivalence of the presented formulation with quadratic penalty methods for constrained optimization [10]. This equivalence allows us to illustrate convergence conditions of the parametric problem minimizer to the solution of the original OCP. In addition, we introduce the related convex problem

$$\text{CVX} : \min_{x_c(\cdot), u(\cdot)} \|x_c - x^{\text{ref}}\|_J^2 \quad (23)$$

subject to

$$\dot{x}_c(t) = f(x^{\text{ref}}(t)) + g(x^{\text{ref}}(t))u(t), \quad t \in [0, T], \quad (24)$$

$$x_c(0) = x_0, \quad (25)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (26)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (27)$$

$$x_c(T) \in \mathbb{X}_T, \quad (28)$$

where the model (24) can be seen as a Taylor series expansion of the original model (18) around the state reference trajectory, i.e., $x_c(t) = x^{\text{ref}}(t) + \delta x(t)$. The right-hand side in (24), corresponds to the zeroth order term in the expansion with a non-fixed $u(t)$ since no knowledge of a possible input reference trajectory, $u^{\text{ref}}(t)$, is provided. Higher order terms depending on $u^{\text{ref}}(t)$ and $x^{\text{ref}}(t)$ are neglected.

The following additional technical assumptions on the presented optimization problems are introduced:

Assumption 2

There exist unique global solutions for the problems OCP (10)-(15), $P(\lambda)$ (17)-(22) and CVX (23)-(28), namely $(x_{\text{ocp}}^*(t), u_{\text{ocp}}^*(t))$, $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ and $(x_{c\text{cvx}}^*(t), u_{\text{cvx}}^*(t))$, respectively.

Assumption 3

1. The parametric problem solution $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ converges pointwise to $(x_{P(1)}^*(t), u_{P(1)}^*(t), x_{cP(1)}^*(t))$, when λ goes to one, i.e.,

$$\lim_{\lambda \rightarrow 1} (x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t)) = (x_{P(1)}^*(t), u_{P(1)}^*(t), x_{cP(1)}^*(t)) \quad \text{pointwise.} \quad (29)$$

2. The parametric problem solution $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ converges pointwise to $(x_{P(0)}^*(t), u_{P(0)}^*(t), x_{cP(0)}^*(t))$, when λ goes to zero, i.e.,

$$\lim_{\lambda \rightarrow 0} (x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t)) = (x_{P(0)}^*(t), u_{P(0)}^*(t), x_{cP(0)}^*(t)) \quad \text{pointwise.} \quad (30)$$

Let us introduce the following lemmata whose proofs are provided in the Appendix:

Lemma 1

The solution of the parametric problem (17)-(22) approaches the solution of the original OCP (10)-(15) in the limit when λ goes to one, i.e.,

$$(x_{P(1)}^*(t), u_{P(1)}^*(t)) = (x_{\text{ocp}}^*(t), u_{\text{ocp}}^*(t)), \quad \forall t \in [0, T] \quad \text{a.e.} \quad (31)$$

$$x_{cP(1)}^*(t) = x_{\text{ocp}}^*(t), \quad \forall t \in [0, T] \quad \text{a.e..} \quad (32)$$

Lemma 2

The solution of the parametric optimization problem $P(\lambda)$ is in the limit, when λ goes to zero, equivalent to the solution of the convex optimization problem in (23)-(28), i.e.,

$$(x_{cP(0)}^*(t), u_{P(0)}^*(t)) = (x_{cvx}^*(t), u_{cvx}^*(t)), \forall t \in [0, T] \quad \text{a.e.} \quad (33)$$

$$x_{P(0)}^*(t) = x^{\text{ref}}(t), \forall t \in [0, T] \quad \text{a.e.} \quad (34)$$

Lemma 3

If $x^{\text{ref}}(t)$ is a feasible trajectory for some unknown controls $\bar{u}(t)$, i.e., $\dot{x}^{\text{ref}}(t) = f(x^{\text{ref}}(t)) + g(x^{\text{ref}}(t))\bar{u}(t)$, then $\bar{u}(t)$ is obtained exactly by solution of the convex problem (23)-(28). Moreover, $(x_{P(\lambda)}^*, u_{P(\lambda)}^*) = (x^{\text{ref}}, \bar{u})$ for all $\lambda \in [0, 1]$.

Convergence of the parametric problem solution to the global solution of the original OCP is guaranteed only if we are able to find the global solution for each $P(\lambda)$ as shown in the appendix. This condition is only easily satisfied for the first problem on the homotopy path namely $P(0)$ where a convex problem is addressed as shown in Lemma 2.

In this context, we propose to solve the non-convex OCP with the given structure, by convexifying it through the formulation in (17)-(22) with $\lambda = 0$. The solution of the convex homotopy is used to initialize successive OCPs when 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. We note that the method is a heuristic that is not guaranteed to find the global solution of the non-convex original problem. However, in view of Lemma 3, it is able to find the correct controls for a trajectory that is perfectly trackable, this situation is presented in Figure 2a. For small values of $\|x - x^{\text{ref}}\|_J^2$ the situation in Lemma 3 is nearly recovered as presented in Figure 2b. Nevertheless, complications can appear for higher values of $\|x - x^{\text{ref}}\|_J^2$ such as bifurcations as illustrated in Figure 2c.

3.2. Parameter estimation problem

The convex-homotopy based approach resembles some smoothing methods for parameter estimation where the model is fitted in terms of some function expansion [11]. Here, the

Figure 2. Some possible behaviours of the homotopy path in the space (x, u, x_c) . Notice that the benevolent situation in (b) is covered with Assumptions 2 and 3 for small trajectory changes. However, if Assumption 2 does not hold, complications like bifurcations or emergence of a new local minima can appear as in (c).

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 the parameters. Additionally, it is assumed that measurements of the state vector are available.

Consider a PEP[†] where it is desired to fit the dynamic behavior of the states $x(t)$ to a set of measurements $x^m(t)$:

$$\min_{x(\cdot), p} \|x - x^m\|_J^2 \quad (35)$$

subject to

$$\dot{x}(t) = f(x(t), u^m(t)) + g(x(t), u^m(t))p, \quad t \in [0, T], \quad (36)$$

$$x(t) \in \mathbb{X}, \quad t \in [0, T], \quad (37)$$

$$p \in \mathbb{P}. \quad (38)$$

Let us assume that $\mathbb{P}$ and $\mathbb{X}$ are convex sets, $x^m(t) \in \mathbb{X}$ and a fixed input $u^m(t)$ corresponding with the measurement data $x^m(t)$ is provided. Following the same approach presented in the OCP case, the PEP can be reformulated by introducing a pseudo state $x_c(t)$ and a homotopy parameter λ :

$$P_p(\lambda) : \min_{x_c(\cdot), x(\cdot), p} \frac{1}{\lambda} \|x - x^m\|_J^2 + \frac{1}{1 - \lambda} \|x_c - x\|_J^2 \quad (39)$$

subject to

$$\dot{x}_c(t) = f(x(t), u^m(t)) + g(x(t), u^m(t))p, \quad \forall t \in [0, T], \quad (40)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (41)$$

$$p \in \mathbb{P}. \quad (42)$$

Since the dynamic system in (36) is parameter-affine, the procedure presented in the OCP case can be applied. When λ tends to one in (39)-(42), the original problem is recovered. On the other hand, when it goes to zero, the term $\|x - x^m\|_J^2$ is penalized strongly leading to $x(t) = x^m(t)$ and allowing the formulation of the convex problem:

$$\min_{x_c(\cdot), p} \|x_c - x^m\|_J^2 \quad (43)$$

subject to

$$\dot{x}_c(t) = f(x^m(t), u^m(t)) + g(x^m(t), u^m(t))p, \quad t \in [0, T], \quad (44)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (45)$$

$$p \in \mathbb{P}. \quad (46)$$

Consequently, it is possible to use convex optimization methods to find an initial guess for the original PEP by solving the related convex problem in (43)-(46). The result is used to

[†]Notice that the problem is formulated such that the initial condition is also optimized.

initialize the subsequent PEPs on the homotopy path with λ towards one in order to recover the original formulation. Although other less general convex formulations have been proposed based on the idea of evaluating the system at the measurements [12], the proposed convex problem exhibits less sensitivity with respect to high frequency noise in the measurements [13]. Moreover, the use of the homotopy refines the convex solution towards the original parameter estimation problem. The advantages of this approach for a non-convex 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 non-penalized optimization variables.

4.1. Parameter Estimation Problem

Let us analyze the dynamic system

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -p & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = f(x(t), p), \quad (47)$$

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

Consider the optimization problem to estimate the parameter p with a Lagrange term defined in the least squares sense and a vanishing Mayer term. The solution of the problem aims to match a given set of measurements $x^m(t)$ 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:

$$\min_{x(\cdot), p} \|x - x^m\|_J^2 \quad \text{subject to} \quad \dot{x}(t) = f(x(t), p), \quad t \in [0, T], \quad (48)$$

with $Q(t) = 1, \forall t \in [0, T]$ and $Q_T = 0$.

The measurement $x^m(t)$ is contaminated with harmonics of the fundamental frequency, $\sqrt{p}$, in order to emphasize the presence of local minima. Figure 3 illustrates the non-convex cost

$$\psi(p) = \min_{x(\cdot)} \|x - x^m\|_J^2 \quad \text{subject to} \quad \dot{x}(t) = f(x(t), p), \quad t \in [0, T]. \quad (49)$$

Clearly, the success of an optimization algorithm for this case depends strongly on parameter initialization in the neighborhood of the solution.

The problem is formulated as in (39)-(42), without constraints on p and $x_c(t)$, i.e., $\mathbb{P} = \mathbb{R}^{n_p}$ and $\mathbb{X} = \mathbb{R}^{n_x}$. Figure 4 depicts the cost

$$\psi(p, \lambda) = \min_{x(\cdot), x_c(\cdot)} \frac{1}{\lambda} \|x - x^m\|_J^2 + \frac{1}{1 - \lambda} \|x_c - x^m\|_J^2 \quad \text{subject to} \quad \dot{x}_c(t) = f(x(t), p), \quad t \in [0, T], \quad (50)$$

Figure 3. Non-convex cost $\psi(p)$ arising from the parameter estimation problem (48), when optimized over $x(\cdot)$.

to be minimized with respect to p when the parameter λ moves between zero and one. This plot is obtained by minimizing first over $x(t)$ and $x_c(t)$ in such a way that the cost $\psi(p, \lambda)$ in Figure 4 is already optimal in $x(t)$ and $x_c(t)$. Notice that λ going to zero leads 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 an initial guess for the original PEP (48).

4.2. The Optimal Control Problem

Since the continuous time OCP is an infinite dimensional optimization problem, the decision variables are parameterized in direct solution approaches so that the problem can be formulated as a finite dimensional nonlinear programming problem (NLP). Here, we assume that the controls are first parameterized by piecewise constant functions and that the model is discretized using finite differences or higher order schemes for its numerical solution. Although, for the current study, the direct single shooting technique is used due to its implementation simplicity, more efficient approaches such as collocation [14] and multiple shooting [15] can also be exploited.

Consider the benchmark problem presented in [16], where an isothermal CSTR with the Van de Vusse reactions

is analyzed. The CSTR dynamics are governed by the set of nonlinear differential equations

Figure 4. Cost function for a set of objective functions between the original problem and the convex problem on the homotopy path for the PEP (39)-(42).

$$\dot{C}_a(t) = \frac{F(t)}{V}(C_{a,0}(t) - C_a(t)) - k_1 C_a(t) - k_3 C_a^2(t), \quad (52)$$

$$\dot{C}_b(t) = k_1 C_a(t) - k_2 C_b(t) - \frac{F(t)}{V} C_b(t), \quad (53)$$

where $F(t)$ represents the feed flow, while V is the reactor volume. $C_a(t)$ and $C_b(t)$ represent the concentrations of the reactant A and the intermediate B , respectively (see Figure 5). Constants k_1 , k_2 and k_3 are reaction rates. It is assumed that the reactor volume, V , remains constant and that 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 exhibits input multiplicity as illustrated in Figure 6. Consequently, the formulation of a cost in terms of the errors with respect to a desired trajectory can generate two possible control values. Hence, a local optimization technique can easily be attracted by a local minimum, delivering a suboptimal solution, if it is initialized inappropriately. However, for this particular application, global optimization approaches are able to find a better solution as proposed in [17]. Our approach suggests a computationally less demanding heuristic to address the problem of finding a global solution to this problem.

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

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

$$\min_{x(\cdot), u(\cdot)} \|x - x^{\text{ref}}\|_J^2 \quad (54)$$

subject to

$$x(0) = x_0 \quad (55)$$

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \quad t \in [0, T], \quad (56)$$

$$x(t) \in \mathbb{X}, \quad t \in [0, T], \quad (57)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (58)$$

where $x^{\text{ref}}(t)$ represents the state trajectory and no reference trajectory for the controls is provided. In the numerical solution, the control $u(t)$ is discretized piecewise constant with N_u intervals.

Figure 5. Isothermal CSTR.

Figure 6. Steady state gain for the CSTR process for different values of the input concentration $C_{a,0}$.

The presented nonlinear dynamics is discretized with a sampling rate of 0.002 h (7.2 s) and solved for a prediction horizon $T = 0.06$ h (30 samples) with a constant control along the prediction horizon [17], i.e., the control horizon, N_u is set to one. 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 $\text{diag}(1e^{-3} \ 2e^{-1}) \text{ L}^2/(\text{gmol}^2\text{h})$ while Q_T is $\text{diag}(1e^{-3} \ 80) \text{ L}^2/(\text{gmol}^2)$.

The cost defined by

$$\psi(u) = \min_{x(\cdot)} \|x - x^{\text{ref}}\|_J^2 \quad (59)$$

subject to (55) and (56) is illustrated in Figure 7. It corresponds to the cost to minimize when a step change on the reference trajectory, $x^{\text{ref}}(t)$, is performed. The original problem has one global and one local solution. 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 $u(t)$ leading to the local minimizer, $u_{\text{lm}}(t)$.

The OCP (54)-(58) is modified using the proposed approach by introducing the norm on $x_c(t) - x(t)$ and the homotopy parameter λ leading to the parametric problem (17)-(22) with $\mathbb{X}_T = \mathbb{R}^{n_x}$. Figure 8 shows the homotopy maps for different values of λ in the interval $(0, 1)$ where $\psi(\lambda, u)$ is analogous to (50), i.e., the cost value in (17) is first minimized with respect to $x(t)$ and $x_c(t)$. Notice that the extremes of the map correspond to a convex function, and to the cost given by the original problem in Figure 7, respectively. In this example it is possible to

Figure 7. Original cost for the CSTR open loop optimal control problem when first optimized over $x(t)$. The value u_{lm} corresponds to the local minimizer

attain the global optimum, if the OCP (54)–(58) is convexified with the proposed approach and the local technique is employed for increasing values of λ performing each time the initialization with the previously computed optimal solution.

Figure 9 shows the behavior of $\|x_c - x\|_J^2$ as a function of the homotopy parameter. Notice that it tends to zero as λ goes to one, recovering the original OCP. This plot corresponds to the values of $\|x_c - x\|_J^2$ on the zero path illustrated in Figure 8.

Additionally, the condition presented in *Lemma 3* is illustrated in Figure 10, where the problem is assumed in steady state and neither changes in the reference nor disturbances are applied. Following this condition the optimal state trajectory for the OCP satisfies $x_{ocp}^*(t) = x^{ref}(t)$ meaning that the system remains in steady state. As stated in *Lemma 3*, the optimal solution of the original OCP (54)–(58) equals the solution provided by the convex cost given by $P(0)$. Hence, in order to obtain the global solution for the original OCP for $x_{ocp}^*(t) = x^{ref}(t)$ it would be enough to solve the convex OCP $P(0)$. Notice that the original OCP for $\lambda = 1$ is still highly non-convex.

5. CLOSED LOOP SIMULATIONS

In order to illustrate the applicability of the method in the NMPC framework, a closed loop simulation study is performed. Notice that in NMPC the NLP resulting from (54)–(58) has to be solved at each sampling instant. A nonlinear open loop observer is implemented in

Figure 8. Cost functions for a set of the problems between the original OCP and its convex formulation on the homotopy path for the CSTR control example when a step change in the reference is performed. The zero path has been obtained by minimizing each one of the parametric OCPs, $P(\lambda)$, for discrete values of λ increasing monotonically, leading to $u^*(t)$ when $\lambda \rightarrow 1$.

Figure 9. Behavior of the norm on $\|x_c - x\|_j^2$ as a function of the homotopy parameter for the CSTR example. This plots corresponds to the zero path values in Figure 8

order to estimate the possible process disturbances and model states. The state reference trajectory is calculated from the desired process output (intermediate concentration $C_b(t)$) and the estimated disturbance as proposed in [18]. No knowledge of the input reference trajectory is provided. At each sampling instant the problem is solved first using the homotopy approach

Figure 10. Cost functions for a set of the problems between the original OCP and its convex formulation on the homotopy path for the CSTR control example. In this case $x_{ocp}^*(t) = x^{ref}(t)$ leading to $u_{P(\lambda)}^* = u_{ocp}^* \forall t \in [0, T]$.

with $\lambda = 0$, i.e., the convex problem is addressed, and after, the original NLP ($\lambda = 1$) is solved using as initial guess the solution provided by the convex OCP. Consequently, one quadratic programming problem (QP) and one NLP are solved every 7.2 s. The former one to obtain the initial guess for $u(t)$ and the latter to refine the approximated solution. Nevertheless, whenever necessary, methods to follow the zero path can be incorporated [19].

In order to evaluate the performance of the closed loop NMPC with the presented method, a test scenario is proposed. The test consists of reference changes and disturbance rejection. During the first 0.05 h of simulation, the system reaches steady state. Hereafter, The output reference trajectory is changed at 0.15 h and 0.6 h. On the other hand, disturbances are applied by changing the feed concentration at 0.3 h and 0.45 h. Figure 11 presents the results for a tracking and disturbance rejection test using both, an NLP solver based on sequential quadratic programming (SQP)[‡] and the proposed convexity-based homotopy approach. Both algorithms are initialized with the same initial state $[x_1, x_2]^T = [6.18, 1.1]^T$ gmol/L. On the one hand, to calculate the first control move, the proposed approach does not require initialization in $u(t)$ since the first problem to solve is convex. On the other hand, the NLP solver requires an initialization which clearly affects the convergence to a solution. In this simulation study, the control is initialized to $u(t) = 180.94h^{-1}$ which corresponds to the control value for the given initial state. The proposed approach immediately finds the right control, causing the strong change in the control action and the big peak in $C_b(t)$. Notice that the homotopy

[‡]The MATLAB® (The Mathworks Inc., Natick) NLP, *fmincon* solver is used for this simulation study

Figure 11. Tracking and disturbance test for the NMPC applied to the CSTR using a local optimizer and convexity-based homotopy approach. The problem is solved with a control horizon $N_u = 1$. The convex approach is switched on at $t=0.05$ hours with the same initialization as in the local approach.

Figure 12. Computational demand for the studied test scenario with $N_u = 1$.

approach finds the global solution while the local optimization technique, which is each time initialized with the previous solution, finds only a local optimum. For the tracking problem, evaluated by performing the changes in reference as mentioned, the NLP locks onto the upper bound of $u(t)$. On the other hand, the homotopy approach finds the global solution as illustrated in the cost function plot. Similar behavior can be noticed when evaluating the disturbance rejection properties. When the process is perturbed, the classical technique finds a local solution while the proposed method is able to compute the global solution. Notice also 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 are solved. The test has been executed using a Intel Core 2 Duo microprocessor at 1.6 GHz running a linux distribution with 2 GB of RAM. Figure 12 presents the computational demand of the evaluated algorithms. In both cases, the computational demand is below the time constraint of 7.2 s, making the approach feasible for implementation.

The OCP is also solved when increasing the degrees of freedom to $N_u = 30$. Figure 13 shows the optimal control provided by the local optimization technique and the proposed method for the case studied in Figure 8 but using 30 control moves. The convex initialization combined with the local optimizer is still able to provide what we believe to be the global solution with a considerably lower objective[§]. Additionally, the closed loop behavior of the process is evaluated with the test scenario described previously and an increased control horizon, $N_u = 30$. Figure 15 depicts the behavior of the NMPC when input concentration and reference changes are performed. In this case, the proposed method is still able to compute the appropriate control sequence, while the local optimization technique provides a suboptimal

[§]For this non-convex problem in 30 variables it is nearly impossible to verify that the global solution was found.

Figure 13. Optimal control trajectory provided by the optimization methods when a step change in the reference is performed at 0.15h. The control horizon is set to $N_u = 30$.

Figure 14. Computational demand for the studied test scenario in Figure 15 with increased control horizon $N_u = 30$.

solution which drives the system to the wrong steady state value exhibiting steady state errors. The computational demand of the simulation is presented in Figure 14 where the time employed by both algorithms is still smaller than the sampling period T_s of 7.2 s.

Figure 15. Tracking and disturbance test for the NMPC applied to the CSTR using a local optimizer and convexity-based homotopy approach. The problem is solved with a control horizon $N_u = 30$. The convex approach is switched on at $t=0.05$ hours with the same initialization as in the local approach.

6. CONCLUSIONS

We have presented a convexity-based homotopy method to initialize non-convex OCPs arising in optimal control problems where the dynamic model involved is affine in the controls and the objective is a penalty on deviations from a desired state reference. It has been demonstrated that the method can be easily extended to parameter estimation problems 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 and optimal control where the proposed approach provides an initial guess to the original OCP, helping local optimization methods to find 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 onto a local solution. Finally, computational aspects of the method have been addressed in order to illustrate its feasibility. In summary, the presented approach promises to become an attractive heuristic for a smart initialization of a specific class of non-convex optimal control problems, using the power of convex optimization.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the anonymous referees for their insightful comments that helped to improve the paper.

APPENDIX

In this section Lemmata 1, 2 and 3 are proven using Assumptions 1, 2 and 3. The proof introduced here follows the formulation presented in [20] for quadratic penalty methods in constrained optimization.

Proof of Lemma 1

Consider the augmented OCP:

$$\min_{x(\cdot), u(\cdot), x_c(\cdot)} \|x - x^{\text{ref}}\|_J^2 \quad (60)$$

subject to

$$\dot{x}_c(t) = f(x(t)) + g(x(t))u(t), \quad t \in [0, T], \quad (61)$$

$$x_c(0) = x_0, \quad (62)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (63)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (64)$$

$$x_c(T) \in \mathbb{X}_T, \quad (65)$$

$$x_c(t) = x(t) \quad t \in [0, T]. \quad (66)$$

whose solution exactly matches the solution of the original problem (10)–(15), since the added equality constraint (66) can be trivially satisfied by the new added variable $x_c(t)$. This additional degree of freedom is introduced in order to provide the same number of optimization variables for the original OCP and the parametric one.

We proceed with the proof by reformulating the parametric OCP (17)–(22) such that the homotopy term penalizing the original cost is factorized:

$$P(\lambda) : \min_{x(\cdot), u(\cdot), x_c(\cdot)} \frac{1}{\lambda} (\|x - x^{\text{ref}}\|_J^2 + \mu(\lambda)\|x_c - x\|_J^2) \quad (67)$$

subject to

$$\dot{x}_c(t) = f(x(t)) + g(x(t))u(t), \quad t \in [0, T], \quad (68)$$

$$x_c(0) = x_0, \quad (69)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (70)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (71)$$

$$x_c(T) \in \mathbb{X}_T. \quad (72)$$

and

$$\mu(\lambda) = \frac{\lambda}{1 - \lambda}. \quad (73)$$

Formulation (67)–(72) resembles the cost used in quadratic penalty methods where the second term in the right-hand side of (67) corresponds to an equality constraint. The scaling factor $\frac{1}{\lambda}$, in (67) is neglected since scaling has no effects on the solution. Moreover, for the case we are interested in, i.e., λ increasing towards 1, this factor tends to one. Notice that (73) defines an equivalence between λ going to 1 and μ going to $+\infty$.

Now, suppose that $(x_{\text{ocp}}^*(t), u_{\text{ocp}}^*(t), x_{c\text{ocp}}^*(t))$ is a global solution to the OCP (60)–(66), hence $x_{c\text{ocp}}^*(t) = x_{\text{ocp}}^*(t)$. Suppose also that $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ is a global solution to the parametric OCP (17)–(22). Notice that any solution to the OCP (60)–(66) lies in the feasible set of the parametric OCP, i.e., (68)–(72) are satisfied by $(x_{\text{ocp}}^*(t), u_{\text{ocp}}^*(t), x_{c\text{ocp}}^*(t)) \forall t \in [0, T]$.

Since $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ is a unique global solution of the parametric problem and $\|x_{c\text{ocp}}^* - x_{\text{ocp}}^*\|_J^2 = 0$, the following inequality holds:

$$\|x_{P(\lambda)}^* - x^{\text{ref}}\|_J^2 + \mu(\lambda)\|x_{cP(\lambda)}^* - x_{P(\lambda)}^*\|_J^2 \leq \|x_{\text{ocp}}^* - x^{\text{ref}}\|_J^2 \quad (74)$$

After dividing (74) by $\mu(\lambda)$ and re-arranging terms, we obtain

$$\|x_{cP(\lambda)}^* - x_{P(\lambda)}^*\|_J^2 \leq \frac{1}{\mu(\lambda)} (\|x_{\text{ocp}}^* - x^{\text{ref}}\|_J^2 - \|x_{cP(\lambda)}^* - x_{P(\lambda)}^*\|_J^2). \quad (75)$$

Following Assumption 3 and taking the limit when $\lambda \rightarrow 1$ in (75) yields

$$\|x_{cP(1)}^* - x_{P(1)}^*\|_J^2 \leq \lim_{\lambda \rightarrow 1} \frac{1}{\mu(\lambda)} \left(\|x_{ocp}^* - x^{\text{ref}}\|_J^2 - \|x_{cP(\lambda)}^* - x_{P(\lambda)}^*\|_J^2 \right) = 0. \quad (76)$$

Due to Assumption 1, $Q \succ 0$ and $Q_T \succeq 0$, hence, the left-hand-side of inequality (76) cannot be negative. Consequently,

$$\|x_{cP(1)}^* - x_{P(1)}^*\|_J^2 = 0, \quad \forall t \in [0, T] \quad \text{a.e.} \quad (77)$$

This condition corresponds to the additional constraint (66) imposed to the augmented OCP in (60) with the new variable $x_c(t)$.

Additionally, taking the limit in (74) gives

$$\lim_{\lambda \rightarrow 1} \left(\|x_{P(\lambda)}^* - x^{\text{ref}}\|_J^2 + \mu(\lambda) \|x_{cP(\lambda)}^* - x_{P(\lambda)}^*\|_J^2 \right) \leq \|x_{ocp}^* - x^{\text{ref}}\|_J^2. \quad (78)$$

The non-negativity of $\mu(\lambda)$ directly implies positiveness of the limit in left-hand side of (78), yielding

$$\|x_{P(1)}^* - x^{\text{ref}}\|_J^2 \leq \|x_{ocp}^* - x^{\text{ref}}\|_J^2. \quad (79)$$

Hence, due to the uniqueness of the solution the following equalities hold,

$$x_{P(1)}^*(t) = x_{ocp}^*(t), \quad \text{a.e.} \quad (80)$$

$$x_{cP(1)}^*(t) = x_{P(1)}^*(t), \quad \text{a.e.} \quad (81)$$

$$x_{ocp}^*(t) = x_{cocp}^*(t), \quad \text{a.e.} \quad (82)$$

and $u_{P(1)}^*$ is uniquely determined by the set of constraints (68)-(72) which is the same set as the original OCP (61)-(66). Consequently,

$$(x_{P(1)}^*(t), u_{P(1)}^*(t), x_{cP(1)}^*(t)) = (x_{ocp}^*(t), u_{ocp}^*(t), x_{cocp}^*(t)), \quad \text{a.e.} \quad (83)$$

meaning that the solution of the parametric OCP (17)-(22) converges to the solution of the original OCP (10)-(15) in the limit when λ goes to one.

Proof of Lemma 2

The proof of Lemma 2 follows the same structure as the proof of Lemma 1. Consider an equivalent formulation of the convex optimization problem (23)-(28):

$$\min_{x_c(\cdot), u(\cdot), x(\cdot)} \|x_c - x^{\text{ref}}\|_J^2 \quad (84)$$

subject to

$$\dot{x}_c(t) = f(x(t)) + g(x(t))u(t), \quad t \in [0, T], \quad (85)$$

$$x_c(0) = x_0, \quad (86)$$

$$x_c(t) \in \mathbb{X}, \quad t \in [0, T], \quad (87)$$

$$u(t) \in \mathbb{U}, \quad t \in [0, T], \quad (88)$$

$$x_c(T) \in \mathbb{X}_T, \quad (89)$$

$$x(t) = x^{\text{ref}}(t), \quad \forall t \in [0, T]. \quad (90)$$

whose solution exactly matches the solution of the original convex problem (23)-(28) since the added equality (90) is satisfied by the introduced extra variable $x(t)$.

Let us reformulate the parametric OCP as:

$$P(\lambda) : \min_{x(\cdot), u(\cdot), x_c(\cdot)} \frac{1}{1-\lambda} (\alpha(\lambda) \|x - x^{\text{ref}}\|_J^2 + \|x_c - x\|_J^2) \quad (91)$$

subject to (68) to (72) and

$$\alpha(\lambda) = \frac{1-\lambda}{\lambda}. \quad (92)$$

The cost in (91) corresponds to a quadratic penalty formulation where the first norm in the cost is considered an equality constraint. Notice also that α goes to $+\infty$ as λ goes to zero. Again, the scaling factor $\frac{1}{1-\lambda}$ has no influence on the cost in the limit case.

Assume that $(x_{\text{cvx}}^*(t), u_{\text{cvx}}^*(t), x_{c\text{cvx}}^*(t))$ is a unique global solution to the convex OCP (84)-(90), and $(x_{P(\lambda)}^*(t), u_{P(\lambda)}^*(t), x_{cP(\lambda)}^*(t))$ is a unique global solution to the parametric OCP $P(\lambda)$. Notice that any solution to the convex OCP in (84)-(90) is a feasible trajectory for the parametric OCP $P(\lambda)$ i.e., (18) to (22) are satisfied by $(x_{\text{cvx}}^*(t), u_{\text{cvx}}^*(t), x_{c\text{cvx}}^*(t)) \forall t \in [0, T]$.

Since the reformulated parametric problem (91) exhibits the same structure as the problem in (67) in the proof of Lemma 1, the same procedure presented there can be applied here, leading to:

$$\|x_{P(0)}^* - x^{\text{ref}}\|_J^2 = 0 \quad (93)$$

and

$$\|x_{P(0)}^* - x^{\text{ref}}\|_J^2 \leq \|x_{c\text{cvx}}^* - x^{\text{ref}}\|_J^2. \quad (94)$$

where $(x_{P(0)}^*(t), u_{P(0)}^*(t), x_{cP(0)}^*(t))$ is a limit trajectory according to Assumption 3. Hence, the following equalities hold

$$x_{cP(0)}^*(t) = x_{c\text{cvx}}^*(t), \quad \text{a.e.} \quad (95)$$

$$x_{P(0)}^*(t) = x^{\text{ref}}(t), \quad \text{a.e.} \quad (96)$$

$$x_{\text{cvx}}^*(t) = x^{\text{ref}}(t), \quad \text{a.e.} \quad (97)$$

and $u_{P(0)}^*$ is uniquely determined by the set of constraints (68)-(72) which is the same set in the original convex problem (85)-(89). Consequently,

$$(x_{P(0)}^*(t), u_{P(0)}^*(t), x_{cP(0)}^*(t)) = (x_{\text{cvx}}^*(t), u_{\text{cvx}}^*(t), x_{c\text{cvx}}^*(t)), \quad \text{a.e.} \quad (98)$$

meaning that the solution of the parametric OCP (17)-(22) converges to the solution of the original convex OCP (23)-(28) in the limit when λ goes to zero.

Proof of Lemma 3

If $x_{\text{ocp}}^*(t) = x^{\text{ref}}(t)$, $(x^{\text{ref}}(t), \bar{u}(t), x_{\text{ocp}}^*(t))$ is a solution of the OCP (60)–(66). Moreover, $(x^{\text{ref}}(t), u_{\text{ocp}}^*(t), x_{\text{ocp}}^*(t))$ minimizes the objective of the parametric OCP in (17) with function value zero. Due to Assumption 2, we have thus $(x_{\text{P}(\lambda)}^*(t), u_{\text{P}(\lambda)}^*(t), x_{\text{cP}(\lambda)}^*(t)) = (x^{\text{ref}}(t), \bar{u}(t), x^{\text{ref}}(t))$ for all $\lambda \in [0, 1]$.

REFERENCES

1. McCormick G. Computability of global solutions to factorable nonconvex programs: part I - convex underestimating problems. *Mathematical Programming* 1976; **10**(1):147–175.
2. Horst R, Tuy H. *Global Optimization: Deterministic 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**(2):159–190.
4. Chachuat B, Singer AB, Barton PI. Global solution of dynamic optimization and mixed-integer dynamic optimization. *Industrial Engineering Chemical Research* 2006; **45**(25):8373–8392.
5. Boyd S, Crusius C, 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 Polynomial 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. *Journal of the Chinese Institute of Chemical Engineers* 2004; **35**(3):299–315.
9. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press, 2006.
10. Gould N. On the convergence of sequential penalty function methods for constrained minimization. *SIAM Journal of Numerical Analysis*. 1989; **26**(1):107–128.
11. Ramsay J, Hooker G, Campbell D, Cao J. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B*. 2007; **69**(5):741–796.
12. Ljung L. Prediction error estimation methods, *Circuits Systems Signal Processing*, 2002; **21**(1):11–21.
13. Bonilla J, Diehl M, De Moor B, Van Impe J. A nonlinear least squares estimation procedure without initial parameter guesses. In *Proceedings of the 47th IEEE Conference on Decision and Control* 2008: 5519–5524.
14. Biegler L. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers and Chemical Engineering* 1984; **8**:243–248.
15. Bock H, Plitt K. A multiple shooting algorithm for direct solution of optimal control problems. In *Proceedings of the 9th IFAC world congress* 1984.
16. Sistu PB, Bequette BW. Model predictive control of processes with input multiplicities. *Chemical Engineering Science* 1995; **50**(6):921–936.
17. Long C, Polisetty P, Gatzke E. Nonlinear model predictive control using a global optimization approach. *Journal of Process Control* 2006; **16**(6):635–643.
18. Rossiter JA. *Model-Based Predictive Control, a Practical Approach*. CRC Press: 2003.
19. L. T. Watson, S. C. Billups, A. P. Morgan. HOMPACK: A suite of codes for globally convergent homotopy algorithms. *ACM transactions on Mathematical Software*. 1987; **13**(3):281–310.
20. Nocedal J. Wright S. *Numerical Optimization*, Springer, NY: 2006.