Nonlinear model predictive control: From theory to application

Article *in* Journal of the Chinese Institute of Chemical Engineers · May 2004

Rolf Findeisen
Otto-von-Guericke-Universität Magdeburg

J. Chin. Inst. Chem. Engrs., Vol. 35, No. 3, 299-315, 2004

Nonlinear Model Predictive Control: From Theory to Application

Frank Allgöwer [1], **Rolf Findeisen, and Zoltan K. Nagy**

Institute for Systems Theory in Engineering, University of Stuttgart
70550 Stuttgart, Germany

Abstract—While linear model predictive control is popular since the 70s of the past century, only since the 90s there is a steadily increasing interest from control theoreticians as well as control practitioners in nonlinear model predictive control (NMPC). The practical interest is mainly driven by the fact that today's processes need to be operated under tight performance specifications. At the same time more and more constraints, stemming for example from environmental and safety considerations, need to be satisfied. Often, these demands can only be met when process nonlinearities and constraints are explicitly taken into account in the controller design. Nonlinear predictive control, the extension of the well established linear predictive control to the nonlinear world, is one possible candidate to meet these demands. This paper reviews the basic principle of NMPC, and outlines some of the theoretical, computational, and implementational aspects of this control strategy.

Key Words : Nonlinear, Model predictive control, Receding horizon control, Model based control, Overview

INTRODUCTION

In many control problems it is desired to design a stabilizing feedback such that a performance criterion is minimized while satisfying constraints on the controls and the states. Ideally one would look for a closed solution for the feedback satisfying the constraints while minimizing the performance. However, often the closed solution can not be found analytically, even in the unconstrained case since it involves the solution of the corresponding Hamilton Jacobi-Bellmann equations. One approach to circumvent this problem is the repeated solution of an open-loop optimal control problem for a given state. The first part of the resulting open-loop input signal is implemented and the whole process is repeated. Control approaches using this strategy are referred to as model predictive control (MPC), moving horizon control or receding horizon control.

In general one distinguishes between linear and nonlinear model predictive control (NMPC). Linear MPC refers to a family of MPC schemes in which linear models are used to predict the system dynamics and considers linear constraints on the states and inputs. Note that even so the system is linear, the closed loop dynamics are nonlinear due to the presence of constraints. NMPC refers to MPC schemes that are based on nonlinear models and/or consider a non-quadratic cost-functional and general nonlinear constraints.

Since its first invention in the 70s of the last century, linear MPC has become an increasingly popular control technique used in industry (Qin and Badgwell, 2000, 2003; García *et al*., 1989; Morari and Lee, 1999; Froisy, 1994). For example (Qin and Badgwell, 2003) reports more than 4,500 applications spanning a wide range from chemicals to aerospace industries are reported. Also many theoretical and implementation issues of linear MPC theory have been studied so far (Lee and Cooley, 1996; Morari and Lee, 1999; Mayne *et al*., 2000).

Many systems are, however, inherently nonlinear. The inherent nonlinearity, together with higher product quality specifications and increasing productivity demands, tighter environmental regulations and demanding economical considerations require to operate systems over a wide range of operating conditions and often near the boundary of the admissible region. Under these conditions linear models are often not sufficient to describe the process dynamics adequately and nonlinear models must be used. This inadequacy of linear models is one of the motivations for the increasing interest in nonlinear model predictive control.

This paper focuses on the application of model predictive control techniques to nonlinear systems. It provides a review of the main principles underlying NMPC and outlines the key advantages/disadvantages of NMPC and some of the theoretical, computational, and implementational aspects. The paper is not intended as a complete review of existing NMPC techniques. Instead we refer to the following list for some

[1] To whom all correspondence should be addressed

excellent reviews (Mayne *et al*., 2000; de Nicolao *et al*., 2000; Allgöwer *et al*., 1999; Rawlings, 2000; Morari and Lee, 1999).

In the next section the basic principle of NMPC is presented, followed by a section summarizing the properties of NMPC. The theoretical aspects of NMPC like stability, robustness, and the output-feedback problem are reviewed. The paper discusses some of the computational aspects of NMPC, and points out important practical issues with respect to NMPC. An example application of NMPC to the control of a high-purity distillation column is shortly presented in the last section of the paper.

PRINCIPLE OF NMPC

Model predictive control is formulated as the repeated solution of a (finite) horizon open-loop optimal control problem subject to system dynamics and input and state constraints. Figure 1 depicts the basic principle of model predictive control. Based on measurements obtained at time t, the controller predicts the dynamic behavior of the system over a prediction horizon T_p in the future and determines (over a control horizon $T_c \le T_p$) the input such that a predetermined open-loop performance objective is minimized. If there were no disturbances and no model-plant mismatch, and if the optimization problem could be solved over an infinite horizon, then the input signal found at $t = 0$ could be applied open-loop to the system for all $t \ge 0$. However, due to disturbances and model-plant mismatch the actual system behavior is different from the predicted one. To incorporate feedback, the optimal open-loop input is implemented only until the next sampling instant. The sampling time between the new optimization can vary in principle. Typically, it is, however, fixed, *i.e*., the optimal control problem is re-evaluated after the sampling time, δ. Using the new system state at time $t + \delta$, the whole procedure—prediction and optimization—is repeated, moving the control and prediction horizon forward.

In Fig. 1 the open-loop optimal input is depicted as arbitrary function of time. To allow a numerical solutions of the open-loop optimal control problem the input is usually parametrized by a finite number of "basis" functions, leading to a finite dimensional optimization problem. In practice often a piecewise constant input is used, leading to T_c/δ decisions for the input over the control horizon.

The determination of the applied input based on the predicted system behavior allows the direct inclusion of constraints on states and inputs as well as the minimization of a desired cost function. However, since often a finite prediction horizon is chosen and thus the predicted system behavior will in general differ from the closed-loop one, precaution must be taken to achieve closed-loop stability and reasonable closed-loop performance. Summarizing, a standard NMPC scheme works as follows:

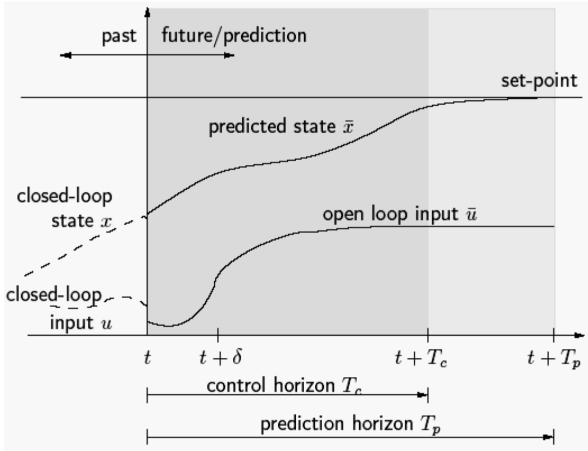

Fig. 1. Principle of model predictive control.

(1) Obtain estimates of the states of the system.
(2) Calculate an optimal input minimizing the desired cost function over the prediction horizon using the system model for prediction.
(3) Implement the first part of the optimal input until the next sampling instant.
(4) Continue with (2).

In the following we discuss the mathematical formulation of NMPC.

Mathematical formulation of NMPC

In this section only the continuous time formulation of NMPC is presented. However, most of the approaches outlined have dual discrete time counterparts, see for example (Mayne *et al*., 2000; de Nicolao *et al*., 2000; Rawlings, 2000; Allgöwer *et al*., 1999; Rawlings *et al*., 1994). Consider the class of continuous time systems described by the following nonlinear differential equation

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0\,, \tag{1}$$

subject to input and state constraints of the form:

$$u(t) \in U,\ \forall t \ge 0\,, \tag{2}$$

$$x(t) \in X,\ \forall t \ge 0\,. \tag{3}$$

Here $x(t) \in \mathbf{R}^n$ and $u(t) \in \mathbf{R}^m$ denote the vector of states and inputs, respectively. Furthermore, the input constraint set U is assumed to be compact and X is connected. For example U and X are often given by box constraints of the form:

$$U := \{u \in \mathbf{R}^m \mid u_{min} \le u \le u_{max}\}\,, \tag{4}$$

$$X := \{x \in \mathbf{R}^n \mid x_{min} \le x \le x_{max}\}\,, \tag{5}$$

with the constant vectors u_{min}, u_{max} and x_{min}, x_{max}.

In NMPC the input applied to the system is usually given by the solution of the following finite horizon open-loop optimal control problem, which is solved at every sampling instant:

Problem 1:

Find $\min_{\bar{u}(\cdot)} J(x(t), \bar{u}(\cdot))$

subject to: $$\dot{\bar{x}}(\tau) = f(\bar{x}(\tau),\ \bar{u}(\tau)),\ \bar{x}(t) = x(t)\,, \tag{6}$$

$$\bar{u}(\tau) \in U,\ \forall\, \tau \in [t,\ t+T_c]\,, \tag{7}$$

$$\bar{u}(\tau) = \bar{u}(t+T_c), \forall\, \tau \in [t+T_c, t+T_p]\,, \tag{8}$$

$$\bar{x}(\tau) \in X,\ \forall\, \tau \in [t,\ t+T_p]\,, \tag{9}$$

with the cost functional

$$J(x(t),\ \bar{u}(\cdot)) := \int_t^{t+T_p} F(\bar{x}(\tau), \bar{u}(\tau))\, \mathrm{d}\tau.$$

Here T_p and T_c are the prediction and the control horizon with $T_c \le T_p$. The bar denotes internal controller variables and $\bar{x}(\cdot)$ is the solution of Eq. (6) driven by the input signal $\bar{u}(\cdot) : [t,\ t + T_p] \rightarrow U$ under the initial condition $x(t)$. The distinction between the real system variables and the variables in the controller is necessary, since even in the nominal case the predicted values will not be the same as the actual closed-loop values. The difference in the predicted and the real values is due to determination of the applied input via a re-optimization (over a moving finite horizon T_c) at every sampling instant.

The cost functional J is defined in terms of the stage cost F, which specifies the performance. The stage cost can for example arise from economical and ecological considerations. Often, a quadratic form for F is used:

$$F(x,\ u) = (x - x_s)^{\mathrm{T}}\ Q(x - x_s) + (u - u_s)^{\mathrm{T}}\ R(u - u_s).$$

Here x_s and u_s denote a desired reference trajectory, that can be constant or time-varying. The deviation from the desired values is weighted by the positive definite matrices Q and R. In the case of a stabilization problem (no tracking), *i.e.*, $x_s = u_s = const$, one can assume, without loss of generality, that $(x_s, u_s) = (0, 0)$ is the steady state to stabilize.

The state measurement enters the system via the initial condition in Eq. (6) at the sampling instants, *i.e.*, the system model used to predict the future system behavior is initialized by the actual system state. Since all state information is required for the prediction, the full state must be either measured or estimated. Equation (8) fixes the input beyond the control horizon to $\bar{u}(t+T_c)$.

In the following, optimal solutions of Problem 1 are denoted by $\bar{u}^*(\,\cdot\,; x(t)) : [t,\ t+T_p] \rightarrow U$. The open-loop optimal control problem is solved repeatedly at the sampling instants $t_j = j\delta, j = 0, 1, \cdots$, and the input applied to the system is given by the sequence of optimal solutions of Problem 1:

$$u(t) := \bar{u}^*(t;\ x(t_j))\,,$$

where t_j is the closest sampling instant to t with $t_j \le t$. Thus, the nominal closed-loop system is given by:

$$\dot{x}(t) = f(x(t),\ \bar{u}^*(t;\ x(t_j)))\,. \tag{10}$$

The optimal cost of Problem 1 as a function of the state is referred to as value function V and is given by:

$$V(x) = J(x,\ \bar{u}^*(\,\cdot\,;\ x))\,. \tag{11}$$

The value function plays a central role in the stability analysis of NMPC, since it often serves as a Lyapunov function candidate (Mayne *et al.*, 2000; Allgöwer *et al.*, 1999).

PROPERTIES OF NMPC

From a theoretical and practical point of view, one would like to use an infinite prediction and control horizon, *i.e.*, T_p and T_c in Problem 1 are set to ∞. This would lead to a minimization of cost up to infinity. However, normally the solution of a nonlinear infinite horizon optimal control problem can not be calculated (sufficiently fast). For this reason finite prediction and control horizons are considered. In this case the actual closed-loop input and states will differ from the predicted open-loop ones, even if no model plant mismatch and no disturbances are present. At the sampling instants the future is only predicted over the prediction horizon. At the next sampling instant the prediction horizon moves forward, allowing to obtain more information thus leading to a mismatch of the trajectories.

The difference of the predicted values and the closed-loop values has two immediate consequences. Firstly, the actual goal to compute a feedback such that the performance objective over the infinite horizon of the closed-loop is minimized is not achieved. In general it is by no means true that the repeated minimization over a moving finite horizon leads to an optimal solution for the infinite horizon problem. The solutions will often differ significantly if a short finite horizon is chosen. Secondly there is in general no guarantee that the closed-loop system will be stable. It is indeed easy to construct examples for which the closed-loop becomes unstable if a short finite horizon is chosen. Hence, when using finite prediction horizons the problem must be modified to guarantee stability, as outlined in section about stability.

The key characteristics and properties of NMPC are:

- NMPC allows the direct use of nonlinear models for prediction.
- NMPC allows the explicit consideration of state and input constraints.
- In NMPC a specified time domain performance criteria is minimized on-line.
- In NMPC the predicted behavior is in general different from the closed loop behavior.
- For the application of NMPC typically a real-time solution of an open-loop optimal control problem is necessary.
- To perform the prediction the system states must be measured or estimated.

Many of these properties can be seen as advantages as well as drawbacks of NMPC. The possibility to directly use a nonlinear model is advantageous if a detailed first principles model is available. In this case often the performance of the closed loop can be increased significantly without much tuning. Nowadays first principle models of a plant are often derived even before a plant is build. Especially in the process industry there is a strong desire to use (rather) detailed models from the first design up to the operation of the plant for reasons of consistence and cost minimization. On the other side, if no first principle model is available, it is often impossible to obtain a good nonlinear model based on identification techniques. In this case it is better to fall back to other control strategies like linear MPC.

Basing the applied input on the solution of an optimal control problem that must be solved on-line is advantageous and disadvantageous at the same time. First, and most important, this allows to directly consider constraints on states and inputs which are often difficult to handle otherwise. Furthermore, the desired cost objective, the constraints and even the system model can in principle be adjusted on-line without making a complete redesign of the controller necessary. However, solving the open-loop optimal control problem, if attacked blindly, can be difficult or even impossible for large systems.

THEORETICAL ASPECTS OF NMPC

In this section different theoretical aspects of NMPC are discussed. Besides, the question of nominal stability of the closed-loop, which can be considered as somehow mature, remarks on robust NMPC strategies as well as the output feedback problem are given. Note that we do not distinguish between so called instantaneous NMPC schemes, for which it is assumed that the computation of the optimal input requires no time and $\delta = 0$, and sampled-data NMPC which recalculate the optimal input signal only at the sampling instants as described above.

Stability

One of the key questions in NMPC is, whether a finite horizon NMPC strategy does guarantee stability of the closed-loop or not. The key problem with a finite prediction and control horizon results from the difference between the predicted open-loop and the resulting closed-loop behavior. Ideally, one would seek for an NMPC strategy which achieves closed-loop stability independent of the choice of the parameters and, if possible, approximates the infinite horizon NMPC scheme as good as possible. An NMPC strategy that achieves closed-loop stability independent of the choice of the performance parameters is often referred to as NMPC approach with guaranteed stability. Different approaches to achieve closed-loop stability using finite horizon lengths exist. Here only the key ideas are reviewed and no detailed proofs are given. Moreover, no attempt is made to cover all existing methods. Most of the technical details are left out for reasons of a clear presentation.

Without loss of generality it is assumed that the origin ($x = 0$ and $u = 0$) is the steady state to stabilize. Furthermore, the prediction horizon is set equal to the control horizon, *i.e.*, $T_p = T_c$, to simplify the presentation.

Infinite horizon NMPC

Probably the most intuitive way to achieve stability is to use an infinite horizon cost, *i.e.*, T_p in Problem 1 is set to ∞. In this case, the open-loop input and state trajectories computed as the solution of the NMPC optimization Problem 1 at a specific sampling instant are in fact equal to the closed-loop trajectories of the nonlinear system. Thus, the remaining parts of the trajectories at the next sampling instant are still optimal (end pieces of optimal trajectories are optimal). This also implies convergence of the closed-loop. Detailed derivations can be found in (Mayne and Michalska, 1990; Mayne *et al*., 2000; Keerthi and Gilbert, 1988, 1985).

Finite horizon NMPC schemes with guaranteed stability

Different possibilities to achieve closed-loop stability using a finite horizon length exist. Most of these approaches modify the standard NMPC setup such that stability of the closed-loop can be guaranteed independently of the plant and performance specifications. This is usually achieved by adding suitable equality or inequality constraints and suitable additional penalty terms to the standard setup. The additional terms are generally not motivated by physical restrictions or performance requirements but have the sole purpose to enforce stability. Therefore, they are usually called stability constraints.

One possibility to enforce stability with a finite prediction horizon is to add a so called zero terminal equality constraint at the end of the prediction horizon, *i.e.*,

$$\bar{x}\,(t+T_p)=0\,, \tag{12}$$

is added to Problem 1 (Chen and Shaw, 1982; Mayne and Michalska, 1990; Keerthi and Gilbert, 1988; Meadows *et al.*, 1995). This leads to stability of the closed-loop, if the optimal control problem has a solution at $t = 0$. Similar to the infinite horizon case the feasibility at one sampling instant does imply feasibility at the following sampling instants and a decrease in the value function. One disadvantage of a zero terminal constraint is that the predicted system state is forced to reach the origin in finite time. This leads to feasibility problems for short prediction/control horizon lengths, *i.e.*, to small regions of attraction. From a computational point of view, an exact satisfaction of a zero terminal equality constraint does require in general an infinite number of iterations in the optimization and is thus not desirable. The main advantages of a zero terminal constraint are the straightforward application and the conceptual simplicity.

Many schemes exist that try to overcome the use of a zero terminal constraint of Eq. (12). Most of them use a so called terminal region constraint

$$\bar{x}\,(t+T_p)\in\Omega\,, \tag{13}$$

and/or a terminal penalty term $E(\bar{x}(t+T_p))$ that are added to Problem 1. The terminal penalty E and the terminal region Ω in Eq. (13) are often determined off-line such that the modified cost function

$$J(x(t),\ \bar{u}(\cdot))=\int_t^{t+T_p}F(\bar{x}(\tau),\bar{u}(\tau))\,\mathrm{d}\tau+E(\bar{x}(t+T_p))\,, \tag{14}$$

gives an upper bound on the infinite horizon cost and guarantees a decrease in the value function. The resulting open-loop optimization problem takes the form:

Problem 2:

Find $\min\limits_{\bar{u}(\cdot)} J(x(t),\ \bar{u}(\cdot))\,,$

with J as in Eq. (14)

subject to: $\dot{\bar{x}}(\tau)=f(\bar{x}(\tau),\bar{u}(\tau)),\quad \bar{x}(t)=x(t)\,,$

$$\bar{u}(\tau)\in U,\quad \forall\,\tau\in[t,\ t+T_p]\ ,$$
$$\bar{x}(\tau)\in X,\quad \forall\,\tau\in[t,\ t+T_p]\ ,$$
$$\bar{x}(t+T_p)\in\Omega\,.$$

If the terminal penalty term E and the terminal region Ω are chosen suitably, it is possible to guarantee closed-loop stability. The following (simplified) stability theorem is based on the results given in (Fontes, 2000). However, in comparison to (Fontes, 2000) the formulation used is more along the lines of the ideas presented in (Chen and Allgöwer, 1998) and (Chen, 1997):

Theorem 1

Assume that:

(1) $U\subset\mathbf{R}^m$ is compact, $X\subseteq\mathbf{R}^n$ is connected and the origin is contained in the interior of $U\times X$.

(2) The vector field $f:\mathbf{R}^n\times\mathbf{R}^m\to\mathbf{R}^n$ is continuous in u and locally Lipschitz in x and satisfies $f(0, 0) = 0$.

(3) $F:\mathbf{R}^n\times U\to\mathbf{R}$ is continuous in all arguments with $F(0, 0) = 0$ and $F(x, u) > 0\ \forall(x, u)\in\mathbf{R}^n\times U\backslash\{0, 0\}$.

(4) The terminal penalty E: $\Omega\to\mathbf{R}$ is continuous with $E(0) = 0$ and that the terminal region Ω is given by $\Omega:=\{x\in X\mid E(x)\le e_1\}$ for some $e_1>0$ such that $\Omega\subset X$.

(5) There exists a continuous local control law $u = k(x)$ such that $k(x)\in U$ for all $x\in\Omega$ and

$$\frac{\partial E}{\partial x}f(x,\ k(x))+F(x,\ k(x))\le 0,\quad \forall x\in\Omega\,.$$

(6) The NMPC open-loop optimal control problem 2 has a feasible solution for $t = 0$. Then for any sampling time $0<\delta\le T_p$ the nominal closed-loop system given by Problem 2 and (10) is asymptotically stable and the region of attraction $\mathbf{R}$ is given by the set of states for which the open-loop optimal control problem has a feasible solution.

Loosely speaking E is a F conform local control Lyapunov function in Ω. Equation (15) allows to upper bound the optimal infinite horizon cost inside Ω by E if a locally stabilizing feedback $k(x)$ is known. The terminal region constraint enforces feasibility at the next sampling instants and allows, similar to the infinite horizon case, to show that the value function is strictly decreasing. Thus stability can be established. Note that this result is nonlocal in nature, *i.e.*, there exists a region of attraction R which is of at least the size of Ω. In general the region of attraction resulting from this scheme is much larger than Ω and is given by all states for which the open-loop optimal control problem has a feasible solution.

Many NMPC schemes guaranteeing stability do fit into this setup (Chen and Allgöwer, 1998; Chen *et* *al*., 2002; de Nicolao *et al*., 1997; Jadbabaie *et al*., 2001; Mayne *et al*., 2000). They vary in how the terminal region and terminal penalty terms are obtained and/or if they appear at all. For example, if a global control Lyapunov function for the system is known and no constraints are present, then no terminal region constraint is needed and as terminal penalty term the control Lyapunov function can be used (Jadbabaie *et al*., 2001). Another possibility is based on a known local control law that stabilizes the system and satisfies the constraints. This local control law can be utilized to obtain a terminal penalty and terminal region satisfying the assumptions of Theorem 0.1. One example for such an approach is the so called quasi-infinite horizon NMPC scheme (Chen and Allgöwer, 1998), which is based in its simplest version on a linearization of the system to obtain a suitable quadratic terminal penalty E and elliptic terminal region Ω.

The use of a terminal inequality constraint leads to computational and feasibility advantages compared to the infinite horizon and zero terminal constraint approach. No zero terminal constraint must be met in finite time. The solution time necessary for solving the open-loop optimal control problem is decreased, since no "boundary-value" problem stemming from the zero terminal constraint must be solved. Furthermore, note that in NMPC it is in general not necessary to find always an optimal solutions of Problem 1 in order to guarantee stability (Michalska and Mayne, 1993; Chen and Allgöwer, 1998; Scokaert *et al*., 1999; Jadbabaie *et al*., 2001). Only a feasible solution leading to a decrease in the value function is necessary. This can be utilized to decrease the necessary on-line solution time and makes the practical application more robust.

Summarizing, the nominal stability question of NMPC is well understood. Various NMPC approaches that guarantee stability exist.

Robustness

The NMPC schemes presented up to now are based on the assumption that the actual system is identical to the model used for prediction, *i.e*., that no model/plant mismatch or unknown disturbances are present. Clearly, this is very unrealistic for practical applications and the development of an NMPC framework to address robustness issues is of paramount importance. In general one distinguishes between the inherent robustness properties of NMPC and NMPC designs taking the uncertainty/disturbances directly into account.

The inherent robustness of NMPC corresponds to the fact that nominal NMPC can cope with uncertainties and disturbances without taking them directly into account. The inherent robustness of NMPC property stems from the close relation of NMPC to optimal control. Results on the inherent robustness of instantaneous NMPC can for example be found in Magni and Sepulchre (1997); Chen and Shaw (1982); Mayne *et al*. (2000). Discrete time results are given in Scokaert *et al*. (1997), and results for sampled-data NMPC are given in Michalska and Mayne (1993); Findeisen *et al*. (2003c). We do not give further details here.

Most robust NMPC schemes that take the uncertainty/disturbance directly into account are based on a min-max formulation. At least three main formulations can be distinguished:

Robust NMPC solving an open-loop min-max problem (*Lall and Glover*, 1994; *Chen et al*., 1997; *Blauwkamp and Basar*, 1999)

In this formulation the standard NMPC setup is kept. However, the cost function takes the worst case uncertainty (or disturbance) out of a set D into account. Thus, the following min-max problem is solved on-line

$$\min_{\bar{u}(\cdot)} \max_{\Delta \in D} \int_t^{t+T_p} F(\bar{x}(\tau), \bar{u}(\tau))\, \mathrm{d}\tau + E(\bar{x}(t+T_p)) ,$$

subject to: $\dot{\bar{x}}(\tau) = f_\Delta(\bar{x}(\tau), \bar{u}(\tau)), \bar{x}(t) = x(t)$.

Here D is the set of uncertainties and f_Δ is the system realization including the uncertainty. The resulting open-loop optimization is a min-max problem. Adding stability constraints similar to the nominal case, is difficult since no feasible solution might be found at all, as all possible uncertainty/disturbance scenarios have to be considered. One open-loop input signal must lead to stability for a whole class of systems "spanned" by the uncertainty while guaranteeing satisfaction of the stability constraints.

H_∞ based NMPC (*Magni et al*., 2001*b*, 2001*c*; *Chen et al*., 1997)

Another possibility is to consider the standard H_∞ problem in a receding horizon framework. The key obstacle is that an infinite horizon min-max problem must be solved (solution of the nonlinear Hamilton-Jacobi-Bellman-Isaacs equation). The main obstacle is the prohibitive computation time and the fact that the global optimum must be found in order to guarantee robust stability.

Robust NMPC design using multiobjective optimization (*Darlington et al*., 2000; *Rustem*, 1994)

The problem with most of the aforementioned min-max approaches is that they usually handle only bounded uncertainties. Additionally, the min-max (or worst-case) formulation of the objective, provides good performance for the worst-case scenario, which has a low probability of occurring, while poor results are typically obtained for more representative situations (such as the nominal case). Different approaches have been proposed to avoid the disadvantages of the min-max techniques. In these approaches usually an objective is optimized that accounts for both nominal and robust performance:

$$J = (1 - w)\, P + w\, R\,, \tag{15}$$

where P and R are the performance and robustness terms, respectively. These two terms are conflicting objectives and the final solution is the result of a compromise between the two, quantified by the parameter w. Many techniques have been proposed which mainly differ in the form and computation of the P and R terms. The sensitivity robustness approach, for example, uses a deterministic framework in which the objective function is the weighted sum of the performance index computed with the nominal parameters, and a penalty term that minimizes the sensitivity of the performance index to variations of the parameters due to uncertainties. Another technique is the mean-variance approach, which uses a general stochastic framework to incorporate uncertainty in the optimization. In this approach the performance term is the expected value of the performance objective, and the robustness term is the variance of the performance index due to the effects of uncertainty. An elegant generalization of the aforementioned multiobjective approaches can be formulated in the differential game framework (Terwiesch *et al*., 1994; Chen *et al*., 1997). This approach simulates a game of "engineer" versus "nature" where the engineer represents the nominal performance and nature represents the uncertainty. Although this technique can be considered a unifying framework for all other multiobjective approaches, it involves highly complex numerical problems and large computing effort, with very few practical applications. Although these approaches do enhance robust performance (Nagy and Braatz, 2003) stability results are not really available in context of these techniques.

Robust NMPC via optimizing a feedback controller used in between the sampling times (*Kothare et al*., 1996; *Magni et al*., 2001*b*)

The open-loop formulation of the robust stabilization problem can be seen as very conservative, since only open-loop control is used during the sampling times, *i.e*., the disturbances are not directly rejected in between the sampling instants. Instead of optimizing the open-loop input signal directly, one can search for an optimal feedback controller that is applied in between the sampling instants, thus introducing instantaneous feedback. In this approach the optimization variables are the design parameter of a "sequence" of control laws $u_i = k_i(x)$ applied in between the sampling instants, *i.e*., the optimization problem has as optimization variables the parameters of the feedback controllers $\{k_1, \cdots, k_N\}$. This formulation overcomes the conservatism of the first approach, since not one single input signal must overcome all possible disturbances. However, often the solution is prohibitively complex.

Summarizing, a deficiency of the standard NMPC setup is that the control action consists of a sequence of open-loop controls, *i.e*., no feedback is used in between the sampling instants. Thus, the input signal in the open-loop min-max formulation has to be chosen such that it can cope with all possible uncertainties/disturbances. Due to this, the solution is in general very conservative or no solution exists at all. One possibility to overcome this problem is to alter the MPC formulation such that an optimization over feedback control laws is undertaken at each step instead of optimizing an open-loop control. This in combination with a min-max formulation can achieve satisfying robustness properties.

Output feedback

One of the key obstacles for the application of NMPC is that at every sampling instant t_i the system state is required for prediction. However, often not all system states are directly accessible, *i.e*., only an output y is directly available for feedback:

$$y = h(x, u)\,, \tag{16}$$

where $y(t) \in \mathbf{R}^p$ are the measured outputs and where h: $\mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}^p$ maps the state and input to the output. To overcome this problem one typically employs a state observer for the reconstruction of the states. In principle, instead of the optimal feedback the following feedback, based on the certainty equivalence principle, is applied:

$$u(t;\hat{x}(t_j)) = \overline{u}^*(t;\ \hat{x}(t_j)), \quad t \in [t_i,\ t_{j+1})\,. \tag{17}$$

Yet, due to the lack of a general nonlinear separation principle, stability is not guaranteed, even if the state observer and the NMPC controller are both stable.

Several researchers have addressed this problem (see for example for an review (Findeisen *et al*., 2003b)). The approach in de Oliveira Kothare and Morari (2000) derives local uniform asymptotic stability of contractive NMPC in combination with a "sampled" state estimator. In Magni *et al*. (2001a), see also Scokaert *et al*. (1997), asymptotic stability results for observer based discrete-time NMPC for "weakly detectable" systems are given. The results allow, in principle, to estimate a (local) region of attraction of the output feedback controller from Lipschitz constants. In Michalska and Mayne (1995) an optimization based moving horizon observer combined with a certain NMPC scheme is shown to lead

lead to (semi-global) closed-loop stability. In Findeisen *et al*. (2003a), and Imsland *et al*. (2003), where semi-global stability results for output-feedback NMPC using high-gain observers are derived. Furthermore, in Findeisen *et al*. (2003b), for a broad class of state feedback nonlinear model predictive controllers, conditions, on the observer that guarantee that the closed-loop is semi-global practically stable.

Even so that a series of output feedback results for NMPC using observers for state recovery exist, most of these approaches are fare away from being implementable. Thus, further research has to address this important question to allow for a practical application of NMPC.

COMPUTATIONAL ASPECTS OF NMPC

NMPC requires the repeated on-line solution of a nonlinear optimal control problem. In the case of linear MPC the solution of the optimal control problem can be cast as the solution of a quadratic program which can be solved efficiently on-line. This is one of the reasons why linear MPC is widely used in industry. For the application of NMPC one has to solve a nonlinear program, which is in general computationally expensive. This was and still is one of the key limiting factors for a successful practical application of NMPC. The available applications of NMPC often consider rather slow processes, allowing sufficient solution time for the optimization. However, the nonlinear program that must be solved in NMPC shows special structure that can be exploited to achieve a real-time feasible solution to the NMPC optimization problem.

As shown, a computational feasible solution of the NMPC problem involves besides the efficient solution of the optimization problem also an efficient formulation of the NMPC strategy.

Possible solution methods

Ideally one would desire to find a closed solution to Problem 2 in feedback form, since this would allow to implement the control without the necessity to perform significant on-line computations. However, finding the closed solution is often not possible. Even in the simple case considering no constraints on the states and inputs finding a solution of the corresponding Hamilton-Jacobi Bellman equation is often not possible or computationally intractable. Thus typically direct optimization methods using a finite parameterization of the input (and possibly the state) are used to find a open-loop solution to Problem 2 that is then implemented in receding horizon fashion to achieve feedback. The basic idea behind the direct methods is to transform the infinite dimensional problem into a finite dimensional nonlinear programming problem. Basically this is done by parameterizing the input (and possibly the states) by a finite number of parameters and to solve/approximate the differential equations during the optimization. In principle any parameterization of the input can be chosen, *i.e.*, the parametrized input is given by

$$\bar{u}(\tau;\ q), \quad \tau \in [t, t+t_p], \tag{18}$$

where the q is the vector of parameterization parameters. The parametrized $\bar{u}(\tau; q)$ might for example be given by a sum of basis functions such as a Fourier series or the input is parameterize as piecewise constant.

Note that while the space of free parameters is finite dimensional, the constraints on the inputs and states do lead to a semi-infinite optimization problem. While the input constraints can often be rewritten into constraints on the input parameterization parameters leading to a finite number of input constraints, the state constraints are more difficult to capture. They are either enforced by adding an exact penalty term to the cost function or are approximately enforced at a finite number of time points over the prediction horizon. The resulting finite dimensional optimization problem takes the form:

$$\min_{q}\ J(\bar{u}(\cdot\,;\ q), x(t);\ t, t+T_p), \tag{19}$$

subject to the state and input constraints and the system dynamics.

Mainly three strategies for the solution of the NMPC optimal control problem using mathematical programming can be distinguished (Pytlak, 1999; Binder *et al*., 2001; Biegler and Rawlings, 1991).

Sequential approach/feasible path approach

In the sequential approach (de Oliveira and Biegler, 1994; Hicks and Ray, 1971; Kraft, 1985) the control is finitely parametrized in the form $\bar{u}(\tau; q)$ and the state trajectories are eliminated by numerically integrating the differential equation and cost. Only the control parameterization parameters remain as degree of freedom in a standard mathematical program given by Eq. (19). For each evaluation of the cost $J(\{\bar{u}_1, \cdots, \bar{u}_N, x(t);\ t, t+T_p\}$ in the solution of the mathematical program the differential equation and the cost function are numerically integrated using the current guess of the input parameterization parameters of the optimizer. Thus the name sequential or feasible path approach, since the optimization steps and the simulation are performed sequentially leading to a valid/feasible state trajectory. The sequential solution method is depicted in Fig. 2.

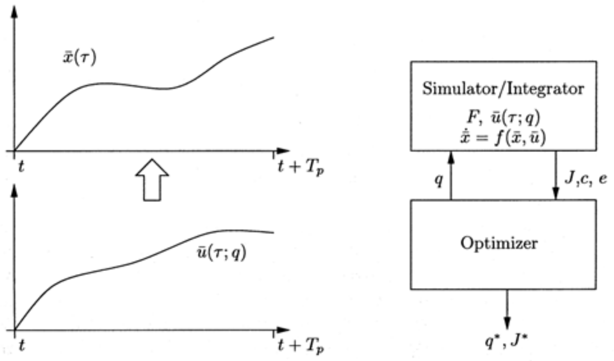

Fig. 2. Sequential solution methods.

Simultaneous approach

In the simultaneous approach the solution to the differential equation and the optimization is obtained simultaneously. For this purpose the differential equations are discretized and enter the optimization problem as additional constraints. Typical simultaneous approaches use collocation methods to parameterize/discretized the differential equations. In the collocation methods (Tsang *et al*., 1975; Biegler, 2000; Cuthrell and Biegler, 1989) collocation is applied to the differential equations. The resulting nonlinear programming problem is very large but also very sparse. The sparsity can be exploited to achieve an efficient solution.

Direct multiple shooting approach

In the direct multiple shooting approach (Bock and Plitt, 1984; Tanartkit and Biegler, 1996; Leineweber, 1998; Bock *et al*., 2000a) the optimization horizon of interest is divided into a number of subintervals with local control parameterizations. The differential equations and cost on these intervals are integrated independently during each optimization iteration based on the current guess of the control. The continuity/consistency of the final state trajectory at the end of the optimization is enforced by adding consistency constraints to the nonlinear programming problem. The resulting nonlinear program takes a special sparse structure which can be utilized for an efficient solution.

Efficient solution via direct multiple shooting

In the following we only outline for the direct multiple shooting approach, how the special NMPC problem structure can be taken into account. Similar statements can be made for the other solution strategies.

To allow a sufficiently fast solution, the following factors are of key importance:

Use of fast integration algorithms

Since for every multiple shooting interval the dynamic system equations must be solve, the use of an efficient integration approach is very important.

Initial value embedding strategy (*Bock et al.*, 2000*b*)

Optimization problems at subsequent sampling instances differ only by different initial values, that are imposed via the initial value constraint. Accepting an initial violation of this constraint, the solution trajectory of the previous optimization problem can be used as an initial guess for the current problem. Furthermore, all problem functions, derivatives as well as an approximation of the Hessian matrix are already available for this trajectory and can be used in the new problem, so that the first QP solution can be performed without any additional ODE solution.

Efficient treatment of least squares cost functions

An efficient approach to obtain a cheap Hessian approximation – the constrained Gauss-Newton method – is recommended in the special case of a least squares type cost function. In NMPC, the involved least squares terms arise in integral form $\int_{t_j}^{t_{j+1}} \| l(x,u) \|_2^2 \, \mathrm{d}t$. Specially adapted integrators that are able to compute a numerical approximation of the Gauss-Newton Hessian for this type of least squares term have been developed (Diehl *et al*., 2001).

The consideration of these factors does improve robustness and speed of the optimization algorithm significantly. An efficient solution strategy taking the afore mentioned points into account has been implemented in a special version of the dynamic optimization software MUSCOD-II (Diehl *et al*., 2002a, 2002b; Diehl, 2002).

Efficient NMPC formulations

Besides an efficient solution strategy of the open-loop optimal control problem the NMPC problem should be also formulated efficiently. Different possibilities for an efficient NMPC formulation exist:

Use of short horizon length without loss of performance and stability

As was outlined in section describing the general properties of NMPC, short horizons are desirable

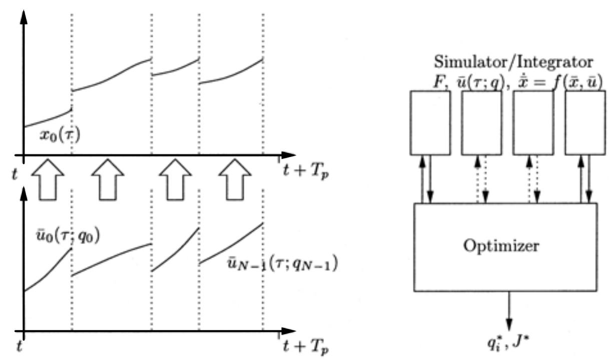

Fig. 3. Simultaneous solution with multiple shooting.

from a computational point of view, but long horizons are desirable for closed-loop stability and to achieve good closed-loop performance. The general NMPC framework outlined previously in the section about NMPC approaches with guaranteed stability, offers a way out of this dilemma. The terminal penalty term can be used as a good approximation of the infinite horizon cost. The often necessary terminal region constraint is in general not very restrictive, *i.e.*, it does not complicate the dynamic optimization problem in an unnecessary manner, as for example a zero terminal constraint does. In some cases, *e.g.*, stable systems, feedback linearizable systems or systems for which a globally valid control Lyapunov function is known, it is not even necessary to consider a terminal region constraint for stability. Thus, such an approach offers the possibility to formulate a computationally efficient NMPC scheme with a short horizon while not sacrificing stability and performance.

Use of suboptimal NMPC strategies, feasibility implies stability

To achieve stability it is often not necessary to find the global minima of the open-loop optimization problem. It is sufficient to achieve a decrease in the value function at every time to guarantee stability (Chen and Allgöwer, 1998; Scokaert *et al*., 1999; Jadbabaie *et al*., 2001). Thus, stability is implied by feasibility. If one uses an optimization strategy that delivers feasible solutions at every subiteration and a decrease in the cost function, the optimization can be stopped if no more time is available and still stability can be guaranteed.

The mentioned points have been taken into account in the implementation of the specially to NMPC tailored version of MUSCOD-II (Diehl *et al*., 2002a, 2002b; Diehl, 2002). Considering the outlined points it is possible to apply NMPC to realistically sized applications even with nowadays computational power (Bartlett *et al*., 2000; Nagy *et al*., 2000; Findeisen *et al*., 2000; Diehl, 2002; Tenny and Rawlings, 2001).

PRACTICAL NMPC APPLICATIONS: FEATURES AND NEEDS

While linear MPC has been classified as the only control technique with a substantial impact on industrial control (Maciejowski, 2002; Qin and Badgwell, 2003), NMPC imposes, besides theoretical issues, practical problems which are considerably more challenging than those associated with linear MPC. This is well reflected in the large difference between the number of industrial applications of linear and nonlinear MPC. Recently more than 4500 industrial LMPC applications were reported, while the number of industrial NMPC applications barely reaches 100 (Qin and Badgwell, 2003). Thus, the development of practical NMPC techniques with quantifiable plant performance improvements that can be supported in an industrial environment is a very challenging task. Important practical issues such as model validation, state estimation and the treatment of model plant mismatch have to be discussed before an NMPC approach can be tested in practice, and ultimately implemented in industry. The experimental results are generally in agreement with those predicted by simulation, but there are certain features of the experimental implementation, which are not visible in the more simplified simulation studies. These results reinforce the need for experimental testing of NMPC algorithms. Several key problems which lead to the major difficulties in the practical implementation of NMPC approaches are discussed below.

Efficient development and identification of control-relevant model

The importance of modelling in NMPC applications is straightforward. Unlike traditional control where modelling and analysis represent a small part of the effort in the application development, a significant amount of time and expense is attributed to the modelling and system identification step in the design and installation of a NMPC. The key aspects which need to be taken into consideration in the development of the control-relevant model are summarized next.

Manageable complexity

Although a detailed model may lead to better performance in principle, often, practical situations require a compromise since the effort for model building and the computational burden involved in the resulting optimization may be prohibitively large. This is one of the reasons why most of the industrial NMPC products use empirical models identified through plant tests, although the attractive features of using first-principles (FP) models, which give the most inside about the process, are widely accepted. First principles models are globally valid and therefore well suited for optimization that can require extrapolation beyond the range of data used to fit the model. However, often the control-relevant model does not necessarily have to be the most accurate, and finding the proper balance between the accuracy of the model and computational burden is usually very challenging.

Measurements based NMPC

One of the most important questions which need to be discussed in the initial phase of the design of any practical NMPC application, is related to what measurements are available versus what it is needed for the model. Often the answer to this question determines the NMPC approach used finally. Besides the modelling difficulty and computational complexity attached to large-scale models, including many details into the model can lead to large number of states resulting in unobservable models based on available measurements. The determination of the control-relevant model always has to be done in conjunctions with the observer design. We believe that the major bottleneck in the increase of FP based NMPC applications is in the lack of proper sensor technology or the availability of state-of-the-art sensors in industrial plants. Due to observability problems with the available sensors, often in industrial applications empirical input-output models are preferred or simplified FP models need to be used so that computational burden is not a major pitfall. Fortunately recently sensors technology has witnessed a significant development. New software and hardware sensors can provide comprehensive information about the processes in a fast and reliable way. These developments have brought many new NMPC applications (*e.g.*, distribution control) within the realm of possibility.

Model identification

Since modelling is a time consuming part of the NMPC design the choice of proper modelling environment is crucial. Almost all industrial NMPC approaches in chemical industry try to beneficiate from the power of modelling software (ASPEN, gPROMS, Hysys, *etc*.) rather then building and solving the models from scratch. Although using existing models developed in these high level programming environments initially might seem to be an alluring approach, often these models are not appropriate for control purposes. However they can serve as the starting point for the identification of the control relevant model. The model identification is important whether empirical or first principle models are used. In the latter case usually offline parameter identification has to be performed. It is very important to keep in mind (however often overlooked) that models (whether empirical or FP) are imperfect (both in term of structure and parameters). Therefore robustness is a key issue in NMPC applications. Robustness has also been identified as one of the major deficiencies of current industrial NMPC products, despite some available theoretical results (see previous section). Choosing the proper plant tests to identify the best model is an important question (Gopaluni *et al*., 2003; Shouche *et al*., 2002). For example many vendors believe that the plant test is the single most important phase in the implementation of linear model predictive controllers. Optimal experimental design approaches which refers to the proper input sequence design (length, excitation, plant-friendly) can provide a recipe to the model identification question. Additionally, robust NMPC design can lead to significant performance improvement (Nagy and Braatz, 2003; Wang and Romagnoli, 2003). Model identification refers not only to the derivation of model parameters but also to the identification of the proper model structure. Ideally for NMPC applications one should identify model structures that capture the process behavior, are amenable to optimization-based control, and admit analysis. To achieve this, model reduction techniques relevant to NMPC need to be used in connection with low order physical modelling approaches. Most likely for successful practical NMPC applications hybrid models need to be used, which have to be developed using methodology that efficiently combines fundamental and empirical modeling techniques.

Reliable and fast solution of the on-line optimization problem

The use of efficient optimization approach is crucial in the real-time feasibility of NMPC applications. Although the need for the development of improved, customizable optimization techniques to handle the rigorous demands imposed by increasingly complex nonlinear models is incontestable, computational performance of NMPC algorithms can be enhanced using other algorithm engineering approaches to the tailored solution of typical problems that arise in NMPC. For example hierarchical solution approaches that exploit problem structure can be applied. CPU time can be used more efficiently, *e.g*., when convergence is achieved before the end of the sampling time the surplus time can be used to pre-compute some parts of the next step. Understanding how various model attributes contribute to particular features of the corresponding optimization problem (*e.g*., non-smooth error surface) can help in the more efficient solution of the optimization. An important aspect in the efficient solution of the optimization problem can be related to finding a more efficient way to represent complex physical systems so that they can be handled more efficiently in the optimization algorithm.

Managing the on-line computational load is not the only concern related to practical NMPC applications. Another important problem, which need to be addressed in practical NMPC applications, is related to the robustness of the optimization. Proper strategies in case of failures of the main NMPC controller due to numerical errors in the optimization or model integration need to be incorporated in practical NMPC approaches. The most straightforward but suboptimal approach is to use the last control input and then to let the lower level controller to act until the system is reset. Another more efficient approach is to run two NMPCs in parallel, when one uses a less demanding but probably slower optimization approach (*e.g.*, based on random search, genetic algorithm, *etc*.).

Design as an integrated system

In industrial NMPC optimization is involved in several steps of the controller (steady-state optimization, dynamic optimization, state estimation, and model identification). An efficient approach to design model, estimator (of model parameters and states), and optimization algorithm as an integrated system (that are simultaneously optimized) rather then independent components could be the capstone in NMPC design.

Long-term maintenance and support of the NMPC algorithm in an industrial application

It is clear the implementation complexity of NMPC approaches is high. But it is important to asses how, long term maintenance can be performed and what the limits of the approaches in face of changing process and operating conditions are. The development of suitable support tools for long term on-site maintenance should be done in parallel with the controller design and implementation. Such a tool should be able to perform *e.g.*, performance assessment, model identification, and controller tuning. The complexity of the models also plays a role in the support and maintenance issues. A very complex model will probably require important effort in the development phase but can provide more flexibility during the long-term maintenance. In contrary, reduced or empirical models might be rigid in face of changing conditions, hence requiring more significant maintenance effort.

Analysis tool for justification of NMPC applications

The advantages of NMPC are straightforward. However, NMPC applications are clearly characterized by a series of practical problems (see above). Usually the economical advantages of NMPC over other simpler approaches are not always clear, and often it can turn out that the application of a NMPC is not recommended due to the excessive overhead required for the design, implementation, and maintenance. For example, answering the question when would NMPC yield significant advantages over MPC in an industrial application, is not always straightforward. Although we believe that there is no "silver bullet" algorithm that would be able to indicate whether the implementation of NMPC to a certain process is beneficial, clearly there is a need for an analysis tool to determine the appropriate technology to use in a particular control application. Such a tool could use for example as inputs: a process description, a performance objective, and the definition of the operating region, could provide the appropriate input set to characterize the process nonlinearity, and finally would supply as an output the answer to the aforementioned question.

APPLICATION EXAMPLE — CONTROL OF A DISTILLATION COLUMN

In the following we outline an example application of NMPC for the control of a high purity binary distillation column with 40 trays for the separation of Methanol and *n*-Propanol.

The binary mixture is fed into the column (compare Fig. 4) with the flow rate F and the molar concentration x_F . The products are removed at the top and bottom of the column with the concentrations x_B and x_D. The manipulated variables are the reflux flow L and the vapor flow V. The control problem is to maintain the specifications on the product concentrations x_B and x_D despite occurring disturbances. Different models for the distillation column are available. The basis model is a 42nd order ODE model describing the concentrations on the trays. Based on this model a reduced 5th order index-one DAE model with 2 differential variables describing the wave positions is available. Furthermore a 164th order model with 44 differential states (concentrations on the trays) and 122 algebraic states (liquid flows, vapor flows and temperatures on each tray) is available. For more details see (Nagy *et al.*, 2000; Diehl *et al.*, 2001). For all simulations the sampling time is δ = 30 s.

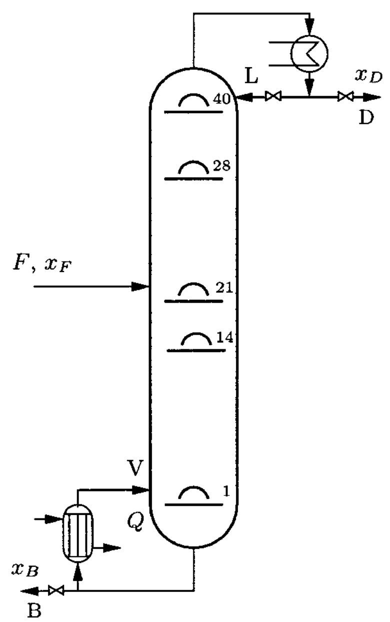

Fig. 4. High purity distillation column.

The concentrations x_B and x_D are not controlled directly. Instead, an inferential control scheme controlling the concentrations (simulations) or the temperatures (experiments) on the 14th tray and on the 28th tray is used. In the cost function F the quadratic deviation of these variables and the inputs from their steady state values are weighed. All results have been obtained using the efficient NMPC implementation in MUSCOD-II as outlined in the Computational Aspects section.

Comparison of different NMPC methods and necessary solution time

In this section we compare infinite horizon NMPC with quasi-infinite horizon NMPC (QIH-NMPC) (Chen and Allgöwer, 1998) and simulation approximated NMPC (SAIH-NMPC). In SAIH-NMPC (de Nicolao *et al.*, 1996) the predicted state at the end of the horizon is forced to lie in the terminal set Ω by a terminal constraint as discussed above. $E(\bar{x}(t+T_p))$ is obtained by an on-line "integration" of the system up to "infinity" using a local control law that stabilizes the system in Ω. In QIH-NMPC an explicit upper bound of the infinite horizon cost inside of Ω that is obtained off-line is used.

The plant is simulated for the comparison by the 164th order model, which is also used in the controllers for the prediction (no model plant mismatch). Furthermore, we assume that the full state information is available by measurements. An LQR controller is used as local controller to calculate the terminal region and terminal penalty term for the QIH-NMPC approach and the SAIH-NMPC approach. It is derived on the basis of the linearization of the system around the considered steady state. Since this controller has a rather large region of attraction, no terminal region is considered in the SAIH-NMPC approach. The terminal region and the quadratic upper bound for the QIH-NMPC approach was found by direct optimization. The infinite horizon NMPC scheme was approximated by an NMPC scheme with 40 control intervals plus a prediction interval of 10,000 s at the end, where the input was fixed to the steady state value. Simulation experiments showed that the performance does not change much, if more than 40 intervals are used. For the SAIH-NMPC approach 5 control intervals were used. The end penalty term E was calculated by simulating the system with the local linear controller for another 10,000 s at the end of the control horizon. In the QIH-NMPC scheme also 5 control intervals were used. To allow the upper bounding of the infinite horizon cost by the (quadratic) penalty term, the final predicted state was constrained to lie in the quadratic terminal region.

We exemplary show the simulation results for a rather drastic scenario, see Fig. 5. At t = 270 s the reflux is decreased from 4 L/h to 0.5 L/h (reflux breakdown), while the vapor flow V is kept constant, *i.e.*, the manipulated variables are fixed to these values. After 180 s the controller is switched on again. This leads to a rather large offset from the steady state. All three controllers nicely return the system to the steady state. Not surprisingly the best performance is achieved by the "infinite" horizon controller. However, this controller can not be implemented in real-time, since the solution of one open-loop optimal control problem in general requires more than the 30 sampling interval, see Fig. 5. In comparison to this the SAIH-NMPC approach as well as the QIH-NMPC controller are able to meet the real-time requirement. As expected, the QIH-NMPC approach does lead to a little bit degraded performance. However, the security margin to the maximum allowed CPU time is significantly higher as for the SAIH-NMPC approach.

Influence of model size and state estimation

In this part we shortly outline the influence of a state observer and different model sizes to the solution time of the optimal control problem on the closed loop. For this purpose we consider again the QIH-NMPC as described above. The simulation setup is similar to the experimental setup presented in the next section. So far we assumed that all states can be accessed directly by measurements. In reality this is however not possible. In the case of the pilot scale distillation column we assume that only the feed tray temperature and the temperature on the 14th and 28th tray are directly measurable. The remaining states are recovered by an Extended Kalman Filter. Also the unknown "disturbances" x_F and F are estimated by the EKF by augmenting the system model by two additional integrators.

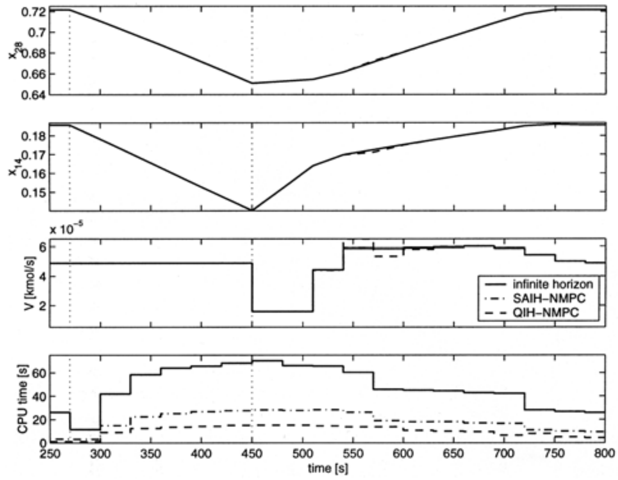

Fig. 5. Behavior for different NMPC schemes.

Figure 6 shows the behavior of the closed loop for different model sizes with respect to feed concentration disturbances. The concentrations are nicely kept in a narrow band around the desired steady state. One can see that the solution is easily possible even for the 164th order model. The reduced maximum computational time in comparison to the state feedback case examined in the section before is mainly due to the "smoothing" effect of the EKF. Since the disturbance and the system state are estimated by the observer and thus do not change instantaneously, the optimization problems do also change only little from sampling time to sampling time. Thus often one SQP iteration is sufficient to obtain a new optimal solution.

Experimental validation

The output feedback scheme presented in simulations in the previous section utilizes an EKF for state estimation. This scheme was implemented and compared to a PI control scheme on a pilot scale distillation column situated at the University of Stuttgart. In contrast to the simulation experiments the heat input Q to the boiler is considered as second manipulated variable and a detailed model of 204th order including hydrodynamic effects is utilized for predictions. The PI control scheme usually employed to control the column consists of two decoupled SISO PI loops. One uses the heat input Q to control the temperature on the 14th tray, the other uses the reflux L to control the temperature T_{28}.

Figure 7 shows the temperature on the 28th as well as the heat input Q to the column. Starting from a steady state, the feed flow F is increased at t = 1,000 s by 20 percent. The NMPC controller is able to complete the transition into the new steady state in approximately 1,000 s with a maximum deviation of T_{28} of 0.3°C. Even though no extensive tuning of the

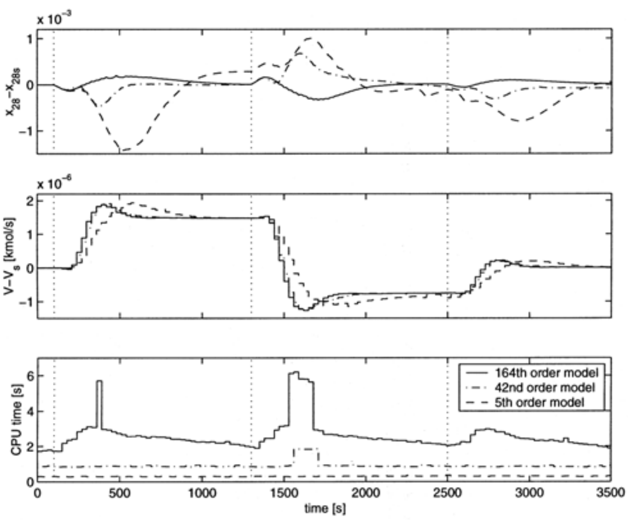

Fig. 6. Behavior of the closed loop EKF.

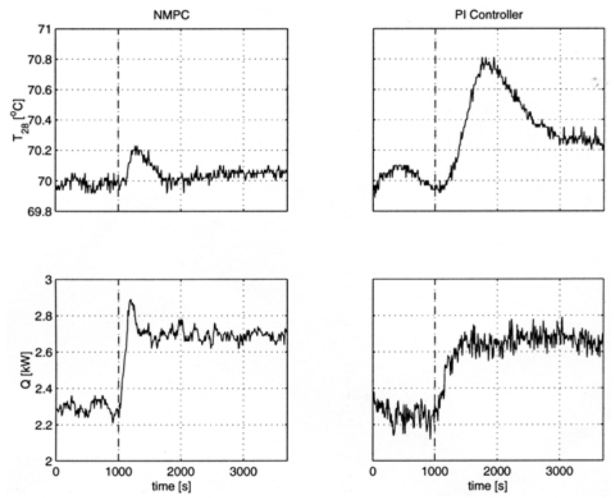

Fig. 7. Experimental results NMPC and PI controller.

NMPC controller was performed, which has a maximum deviation of 0.8°C and completes the transition to the new steady state with an inevitable offset of 0.25°C in T_{28} after 1,500 s.

The presented results underpin that an application of NMPC schemes with reduced computational demand is even possible nowadays if specially tailored optimization schemes and efficient NMPC formulations are used. Furthermore, the closed-loop shows good performance without the need to use a reduced order model or much tuning.

CONCLUSION

Model predictive control for linear constrained systems has been proven as a useful control solution for many practical applications. It is expected that the use of nonlinear models in the predictive control framework, leading to nonlinear model predictive control, results in improved control performance and allows the direct use of first principle based models. However, the consideration of nonlinear models also poses challenging theoretical, computational and implementational problems.

In this paper an overview of the theoretical and computational aspects of NMPC has been given. As shown, some of the challenges occurring in NMPC are already solvable. Nevertheless many unsolved questions remain. Here only a few are noted as a possible guide for future research:

(1) Output feedback NMPC: While some first results of a theoretically sound output feedback NMPC schemes exist, often none of them is applicable to real world processes at present. Especially the incorporation of suitable state estimation strategies in the NMPC formulation must be investigated further.

(2) Robust NMPC formulations: By now a few robust NMPC formulations exist. While the existing schemes increase the general understanding they are computationally intractable to be applied in practice. Further research is required to develop implementable robust NMPC strategies.
(3) Industrial applications of NMPC: The state of industrial application of NMPC is growing rapidly. However, none of the NMPC algorithms provided by vendors include stability constraints as required by control theory for nominal stability. Future developments in NMPC control theory should contribute to making the gap between academic and industrial developments smaller.

ACKNOWLEDGEMENT

The authors thankfully acknowledge the cooperation with M. Diehl, H. G. Bock and J. Schlöder from the Center for Scientific Computing (IWR) at the University of Heidelberg, Germany in the area of efficient solution of NMPC and I. Dislu and S. Schwarzkopf of the IST at the University of Stuttgart for the joint work applying these methods to the control of the distillation column.

NOMENCLATURE

D	set of uncertainties
E	terminal penalty term
f	differential model equations
f_Δ	system realization including uncertainty
h	measurements function
J	cost function
k_i	parameters of the feedback controller
P	performance term in the objective
Q	positive definite weighting matrix
q	vector of parametrization parameters
$\mathbf{R}$	positive definite weighting matrix
R	robustness term in the objective
T_c	control horizon, s
T_p	prediction horizon, s
t	time, s
U	input constraints
u_s	reference input trajectory
$u(t)$	vector of inputs
u_{max}	constant vector of upper bounds on control inputs
u_{min}	constant vector of lower bounds on control inputs
$\bar{u}$	internal variable for the control inputs used in the controller
$\bar{u}^*$	optimal solution of the open-loop control problem
V	value function
X	state constraints
x_{max}	constant vector of upper bounds on states
x_{min}	constant vector of lower bounds on states
x_s	reference state trajectory
$x(t)$	vector of states
$\bar{x}$	internal variable for the states used in the controller
$y(t)$	measured outputs

Greek symbols

δ	sampling time, s
Ω	terminal region

REFERENCES

Allgöwer, F., T. Badgwell, J. Qin, J. Rawlings, and S. Wright, “Nonlinear Predictive Control and Moving Horizon Estimation — An Introductory Overview,” P. Frank., Ed., *Advances in Control, Highlights of ECC’99*, Springer, p. 391 (1999).

Bartlett, R., A. Wächter, and L. Biegler, “Active Set vs. Inerior Point Strategies for Model Predicitve Control,” *Proc. Amer. Contr. Conf.*, Chicago, Il, p. 4229 (2000).

Biegler, L. and J. B. Rawlings, “Optimization Approaches to Nonlinear Model Predictive Control,” in W. Ray and Y. Arkun, Eds., *Proc. 4th International Conference on Chemical Process Control-CPC IV*, AIChE, CACHE, p. 543 (1991).

Biegler, L., “Efficient Solution of Dynamic Optimization and NMPC Problems,” F. Allgöwer and A. Zheng, Eds., *Nonlinear Predictive Control*, Birkhäuser, p. 219 (2000).

Binder, T., L. Blank, H. G. Bock, R. Burlisch, W. Dahmen, M. Diehl, T. Kronseder, W. Marquardt, J. Schlöder, and O. von Stryk, “Introduction to Model Based Optimization of Chemical Processes on Moving Horizons,” in M. Groetschel, S. Krumke, and J. Rambau, Eds., *Online Optimization of Large Scale Systems: State of the Art*, Springer, p. 295 (2001).

Blauwkamp, R. and T. Basar, “A Receding-Horizon Approach to Robust Output Feedback Control for Nonlinear Systems,” *Proc. 38th IEEE Conf. Decision Contr.*, p. 4879, San Diego, U.S.A. (1999).

Bock, H. G. and K. Plitt, “A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems,” *Proc. 9th IFAC World Congress*, Budapest, Hungary (1984).

Bock, H. G., M. Diehl, D. Leineweber, and J. Schlöder, “A Direct Multiple Shooting Method for Real-Time Optimization of Nonlinear DAE Processes,” in F. Allgöwer and A. Zheng, Eds., *Nonlinear Predictive Control*, Birkhäuser, p. 245 (2000a).

Bock, H. G., M. Diehl, J. Schlöder, F. Allgöwer, R. Findeisen, and Z. Nagy, "Real-Time Optimization and Nonlinear Model Predictive Control of Processes Governed by Differential-Algebraic Equations," *Proc. Int. Symp. Adv. Control of Chemical Processes*, p. 695, ADCHEM, Pisa, Italy (2000b).

Chen, C. and L. Shaw, "On Receding Horizon Feedback Control," *Automatica*, **18**(3), 349 (1982).

Chen, H., *Stability and Robustness Considerations in Nonlinear Model Predictive Control*, Fortschr.-Ber. VDI Reihe 8 Nr. 674, VDI Verlag, Düsseldorf, Germany (1997).

Chen, H., C. W. Scherer, and F. Allgöwer, "A Game Theoretic Approach to Nonlinear Robust Receding Horizon Control of Constrained Systems," *Proc. Am. Contr. Conf.*, Albuquerque, p. 3073 (1997).

Chen, H. and F. Allgöwer, "A Quasi-Infinite Horizon Nonlinear Model Predictive Control Scheme with Guaranteed Stability," *Automatica*, **34**(10), 1205 (1998).

Chen, W., D. Balance, and J. O'Reilly, "Optimisation of Attraction Domains of Nonlinear MPC via LMI Methods," *Proc. Am. Contr. Conf.*, p. 3067, Arlington, U.S.A. (2002).

Cuthrell, J. and L. Biegler, "Simultaneous Optimization and Solution Methods for Batch Reactor Profiles," *Comput. Chem. Eng.*, **13**(1/2), 49 (1989).

Darlington, J., C. C. Pantelides, B. Rustem, and B. A. Tanyi, "Decreasing the Sensitivity of Open-Loop Optimal Solutions in Decision Making under Uncertainty," *Eur. J. Oper. Res.*, **121**, 343 (2000).

de Nicolao, G., L. Magni, and R. Scattolini, "Stabilizing Nonlinear Receding Horizon Control via a Nonquadratic Terminal State Penalty," *Symposium on Control, Optimization and Supervision, CESA'96 IMACS Multiconference*, Lille, p. 185 (1996).

de Nicolao, G., L. Magni, and R. Scattolini, "Stabilizing Receding-Horizon Control of Nonlinear Time-Varying Systems," *Proc. 4th European Control Conference ECC'97*, Brussels, Belgium (1997).

de Nicolao, G., L. Magni, and R. Scattolini, "Stability and Robustness of Nonlinear Receding Horizon Control," F. Allgöwer and A. Zheng, Eds., *Nonlinear Predictive Control*, Birkhäuser, p. 3 (2000).

de Oliveira, N. and L. Biegler, "Constraint Handling and Stability Properties of Model-Predictive Control," *AIChE J.*, **40**(2), 1138 (1994).

de Oliveira Kothare, S. and M. Morari, "Contractive Model Predictive Control for Constrained Nonlinear Systems," *IEEE Trans. Automat. Contr.*, **45**(6), 1053 (2000).

Diehl, M., I. Uslu, R. Findeisen, S. Schwarzkopf, F. Allgöwer, H. G. Bock, T. Bürner, E. Gilles, A. Kienle, J. Schlöder, and E. Stein, "Real-Time Optimization of Large Scale Process Models: Nonlinear Model Predictive Control of a High Purity Distillation Column," in M. Groetschel, S. Krumke, and J. Rambau, Eds., *Online Optimization of Large Scale Systems: State of the Art*, Springer, p. 363 (2001).

Diehl, M., "Real-Time Optimization for Large Scale Nonlinear Processes," Vol. 920 of *Fortschr.-Ber. VDI Reihe* 8, *Meß, Steuerungs- und Regelungstechnik*, VDI Verlag, Düsseldorf, Germany (2002).

Diehl, M., R. Findeisen, Z. Nagy, H. G. Bock, J. Schlöder, and F. Allgöwer, "Real-Time Optimization and Nonlinear Model Predictive Control of Processes Governed by Differential-Algebraic Equations," *J. Proc. Contr.*, **4**(12), 577 (2002a).

Diehl, M., R. Findeisen, S. Schwarzkopf, I. Uslu, F. Allgöwer, H. G. Bock, and J. Schlöder, "An Efficient Approach for Nonlinear Model Predictive Control of Large-Scale Systems, Part I, Description of the Methodology," *Automatisierungstechnik*, **12**, 557 (2002b).

Findeisen, R., F. Allgöwer, M. Diehl, H. G. Bock, J. Schlöder and Z. Nagy, "Efficient Nonlinear Model Predictive Control," *6th International Conference on Chemical Process Control – CPC VI*, p. 454 (2000).

Findeisen, R., L. Imsland, F. Allgöwer, and B. Foss, "Output Feedback Stabilization for Constrained Systems with Nonlinear Model Predictive Control," *Int. J. Robust and Nonlinear Control*, **13**(3-4), 211 (2003a).

Findeisen, R., L. Imsland, F. Allgöwer, and B. Foss, "State and Output Feedback Nonlinear Model Predictive Control: An Overview," *Eur. J. Contr.*, **9**(2-3), 190 (2003b).

Findeisen, R., L. Imsland, F. Allgöwer, and B. Foss, "Towards a Sampled-Data Theory for Nonlinear Model Predictive Control," in W. Kang, C. Borges, and M. Xiao, Eds., *New Trends in Nonlinear Dynamics and Control*, Vol. 295 of *Lecture Notes in Controland Information Sciences*, p. 295, Springer-Verlag, New York, U.S.A. (2003c).

Fontes, F., "A General Framework to Design Stabilizing Nonlinear Model Predictive Controllers," *Syst. Control Lett*, **42**(2), 127 (2000).

Froisy, J. B., "Model Predictive Control: Past, Present, and Future," *ISA Transactions*, **33**, 235 (1994).

García, C., D. Prett, and M. Morari, "Model Predictive Control: Theory and Practice — A Survey," *Automatica*, **25**(3), 335 (1989).

Gopaluni, R. B., R. S. Patwardhan, and S. L. Shah, "The Nature of Data Pre-Filters in MPC Relevant Identification-Open- and Closed-Loop Issues," *Automatica*, in press (2003).

Hicks, G. and W. Ray, "Approximation Methods for Optimal Control Synthesis," *Can. J. Chem. Eng.*, **49**, 522 (1971).

Imsland, L., R. Findeisen, E. Bullinger, F. Allgöwer, and B. Foss, "A Note on Stability, Robustness, and Performance of Output Feedback Nonlinear Model Predictive Control," *J. Proc. Contr.*, **13**(7), 633 (2003).

Jadbabaie, A., J. Yu, and J. Hauser, "Unconstrained Receding Horizon Control of Nonlinear Systems," *IEEE Trans. Automat. Contr.*, **46**(5), 776 (2001).

Keerthi, S. and E. Gilbert, "An Existence Theorem for Discrete-Time Infinite-Horizon Optimal Control Problems," *IEEE Trans. Automat. Contr.*, **30**(9), 907 (1985).

Keerthi, S. and E. Gilbert, "Optimal Infinite-Horizon Feedback Laws for a General Class of Constrained Discrete-Time Systems: Stability and Moving-Horizon

Approximations," *J. Optim. Theory Appl.*, **57**(2), 265 (1988).

Kothare, M., V. Balakrishnan, and M. Morari, "Robust Constrained Model Predictive Control Using Linear Matrix Inequalities," *Automatica*, **32**(10), 1361 (1996).

Kraft, D., "On Converting Optimal Control Problems into Nonlinear Programming Problems," *Computational Mathematical Programming*, Kluwer Academic Publishers, Dodrecht, p. 261 (1985).

Lall, S. and K. Glover, "A Game Theoretic Approach to Moving Horizon Control," in D. Clarke, Ed., *Advances in Model-Based Predictive Control*, Oxford University Press (1994).

Lee, J. and B. Cooley, "Recent Advances in Model Predictive Control and Other Related Areas," in J. Kantor, C. Garcia, and B. Carnahan, Eds., *5th International Conference on Chemical Process Control — CPC V*, American Institute of Chemical Engineers, p. 201 (1996).

Leineweber, D., "Efficient Reduced SQP Methods for the Optimization of Chemical Processes Described by Large Sparse DAE Models," Ph.D. Dissertation, University of Heidelberg, Germany (1998).

Maciejowski, J. M., *Predictive Control with Constraints*, Prentice Hall, Pearson Education Limited, Harlow, England (2002).

Magni, L. and R. Sepulchre, "Stability Margins of Nonlinear Receding—Horizon Control via Inverse Optimality," *Syst. Contr. Lett.*, **32**(4), 241 (1997).

Magni, L., G. de Nicolao, and R. Scattolini, "Output Feedback and Tracking of Nonlinear Systems with Model Predictive Control," *Automatica*, **37**(10), 1601 (2001a).

Magni, L., G. de Nicolao, R. Scattolini, and F. Allgöwer, "Robust Receding Horizon Control for Nonlinear Discrete-Time Systems," *Proceedings 15th IFAC World Congress*, Barcelona, Spain (2001b).

Magni, L., H. Nijmeijer, and A. van der Schaft, "A Receding-Horizon Approach to the Nonlinear H_∞ Control Problem," *Automatica*, **37**(5), 429 (2001c).

Mayne, D. Q., and H. Michalska, "Receding Horizon Control of Nonlinear Systems," *IEEE Trans. Automat. Contr.*, **35**(7), 814 (1990).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and Optimality," *Automatica*, **26**(6), 789 (2000).

Meadows, E. S., M. Henson, J. W. Eaton, and J. B. Rawlings, "Receding Horizon Control and Discontinuous State Feedback Stabilization," *Int. J. Contr.*, **62**(5), 1217 (1995).

Michalska, H. and D. Q. Mayne, "Robust Receding Horizon Control of Constrained Nonlinear Systems," *IEEE Trans. Automat. Contr.*, **AC-38**(11), 1623 (1993).

Michalska, H. and D. Q. Mayne, "Moving Horizon Observers and Observer-Based Control," *IEEE Trans. Automat. Contr.*, **40**(6), 995 (1995).

Morari, M. and J. H. Lee, "Model Predicitve Control: Past, Present, and Future," *Comput. Chem. Eng.*, **23**(4/5), 667 (1999).

Nagy, Z., R. Findeisen, M. Diehl, F. Allgöwer, H. G. Bock, S. Agachi, J. Schlöder, and D. Leineweber, "Real-Time Feasibility of Nonlinear Predictive Control for Large Scale Processes — A Case Study," *Proc. Amer. Contr. Conf.*, Chicago, p. 4249 (2000).

Nagy, Z. K. and R. D. Braatz, "Robust Nonlinear Model Predictive Control of Batch Processes," *AIChE J.*, **49**(7), 1776 (2003).

Pytlak, R., *Numerical Methods for Optimal Control Problems with State Constraints*, Lecture Notes in Mathematics, Springer, Berlin, Germany (1999).

Qin, S. J. and T. Badgwell, "An Overview of Nonlinear Model Predictive Control Applications," F. Allgöwer and A. Zheng, Eds., *Nonlinear Predictive Control*, Birkhäuser, p. 369 (2000).

Qin, S. J. and T. Badgwell, "A Survey of Industrial Model Predictive Control Technology," Accepted for *Control Engineering Practice* (2003).

Rawlings, J. B., E. S. Meadows, and K. R. Muske, "Nonlinear Model Predictive Control: A Tutorial and Survey," *Proc. Int. Symp. Adv. Control of Chemical Processes*, p. 185, *ADCHEM*, Kyoto, Japan (1994).

Rawlings, J. B., "Tutorial Overview of Model Predictive Control," *IEEE Contr. Syst. Magazine*, **20**(3), 38 (2000).

Rustem, B., "Stochastic and Robust Control of Nonlinear Economic Systems," *Eur. J. Oper. Res.*, **73**, 304 (1994).

Scokaert, P., J. B. Rawlings, and E. S. Meadows, "Discrete-Time Stability with Perturbations: Application to Model Predictive Control," *Automatica*, **33**(3), 463 (1997).

Scokaert, P., D. Mayne, and J. B. Rawlings, "Suboptimal Model Predictive Control (Feasibility Implies Stability)," *IEEE Trans. Automat. Contr.*, **44**(3), 648 (1999).

Shouche, M. S., H. Genceli, and M. Nikolaou, "Effect of On-Line Optimization Techniques on Model Predictive Control and Identification (MPCI)," *Comput. Chem. Eng.*, **26**, 1241 (2002).

Tanartkit, P. and L. Biegler, "A Nested, Simultaneous Approach for Dynamic Optimization Problems – I," *Comput. Chem. Eng.*, **20**(4/5), 735 (1996).

Tenny, M. J. and J. B. Rawlings, "Feasible Real-Time Nonlinear Model Predictive Control," *6th International Conference on Chemical Process Control — CPC VI*, AIChE Symposium Series (2001).

Terwiesch, P., M. Agarwal, and D. W. T. Rippin, "Batch Unit Optimization with Imperfect Modeling: A Survey," *J. Proc. Contr.*, **4**, 238 (1994).

Tsang, T., D. Himmelblau, and T. Edgar, "Optimal Control via Collocation and Non-Linear Programming," *Int. J. Contr.*, p. 763 (1975).

Wang, D. and J. A. Romagnoli, "Robust Model Predictive Control Design using a Generalized Objective Function," *Comput. Chem. Eng.*, **27**, 965 (2003).