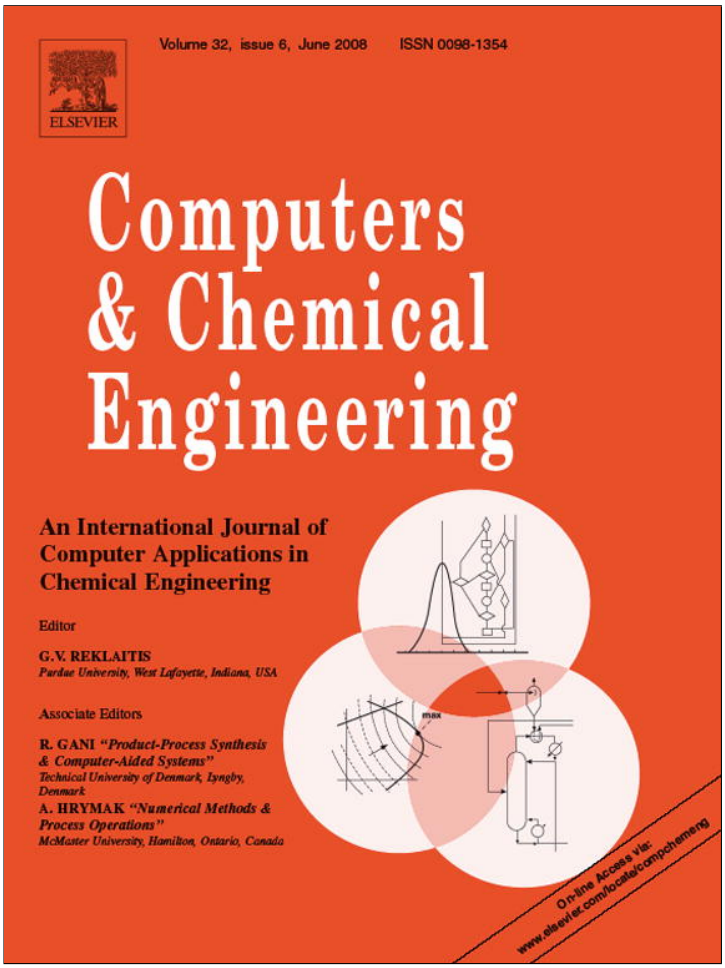
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
ScienceDirect
Computers and Chemical Engineering 32 (2008) 1287–1300
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

Numerical solution approaches for robust nonlinear optimal control problems

Moritz Diehl [a], Johannes Gerhard [b], Wolfgang Marquardt [b,*], Martin Mönnigmann [c]

[a] *Optimization in Engineering Center (OPTEC), Electrical Engineering Department (ESAT), K.U. Leuven, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium*
[b] *Lehrstuhl für Prozesstechnik, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany*
[c] *Institut für Wärme- und Brennstofftechnik, Technische Universität Braunschweig, Postfach 3329, D-38023 Braunschweig, Germany*

Received 24 February 2006; received in revised form 5 June 2007; accepted 5 June 2007
Available online 14 June 2007

Abstract

Nonlinear equality and inequality constrained optimization problems with uncertain parameters can be addressed by a robust worst-case formulation that leads to a bi-level min–max optimization problem. We propose and investigate a numerical method to solve this min–max optimization problem exactly in the case that the underlying maximization problem always has its solution on the boundary of the uncertainty set. This is an adoption of the local reduction approach used to solve generalized semi-infinite programs. The approach formulates an equilibrium constraint employing first order derivatives of both the uncertainty set and the user defined constraints. We propose two different ways for computation of these derivatives, one similar to the forward mode, the other similar to the reverse mode of automatic differentiation. We show the equivalence of the proposed approach to a method based on geometric considerations that was recently developed by some of the authors. We show how to generalize the techniques to optimal control problems. The robust dynamic optimization of a batch distillation illustrates that both techniques are numerically efficient and able to overcome the inexactness of another recently proposed numerical approach to address uncertainty in optimal control problems.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Robust optimization; Optimal control; Parametric uncertainty; Semi-infinite program; Local reduction; Batch distillation

1. Introduction

We consider uncertain nonlinear programming problems (NLP) of the form

$$\min_{x\in\mathbb{R}^{n_x},u\in\mathbb{R}^{n_u}} f_0(x,u,p), \quad \text{s.t.} \begin{cases} f(x,u,p)\le 0 \\ g(x,u,p)=0 \end{cases} \tag{1}$$

with uncertain parameters $p\in\mathbb{R}^{n_p}$. The optimization variables are partitioned into states $x\in\mathbb{R}^{n_x}$ and controls $u\in\mathbb{R}^{n_u}$. The objective function f_0, inequality constraints f, and equality constraints g are smooth functions which map from $\mathbb{R}^{n_p}\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}$ into $\mathbb{R}$, $\mathbb{R}^{n_f}$, and $\mathbb{R}^{n_x}$, respectively. We assume the Jacobian $\partial g/\partial x$ to be invertible everywhere, so that we can regard the state variables x as an implicit function of u and p, which we will denote by $x(u,p)$. This division into states x and controls u arises naturally in model based optimization, where the equalities $g(x,u,p)=0$ often contain discretized ordinary or partial differential equations, such that we are in particular interested in case where $n_x\gg 1$.

We assume we have some knowledge about the uncertain parameters p such that they are restricted to the compact set

$$\mathbb{P}(u) := \{p\in\mathbb{R}^{n_p} | \exists x\in\mathbb{R}^{n_x} : g(x,u,p)=0, h(x,u,p)\le 0\}. \tag{2}$$

The nonlinear program (1) together with the uncertainty set (2) is a generalized semi-infinite program (GSIP) (Hettich & Kortanek, 1993) with differentiable functions h that map from $\mathbb{R}^{n_p}\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}$ into $\mathbb{R}^q$. In this work we consider a subclass of GSIP with the assumption that h is a differentiable scalar function. This definition of the uncertainty set in particular includes the case of a confidence ellipsoid $\mathbb{P}_{\text{ellpsd.}}$ for a Gaussian random variable p with expectation value $\bar{p}$, variance–covariance matrix

* Corresponding author.
E-mail address: marquardt@lpt.rwth-aachen.de (W. Marquardt).

0098-1354/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2007.06.002

Author's personal copy

Σ, and a scalar $\gamma > 0$ depending on the desired confidence level (that is independent of u):

$$\mathbb{P}_{\text{ellpsd.}} = \{p \in \mathbb{R}^{n_p} | (p - \bar{p})^{\mathrm{T}} \Sigma^{-1} (p - \bar{p}) - \gamma \leq 0\}.$$

Other types of smooth uncertainty sets used in robust optimization also include confidence regions derived from the likelihood ratio test (Rooney & Biegler, 2001).

In order to incorporate the uncertainty in the optimization problem formulation, we choose the worst-case min–max formulation of (1). For this aim we assume that the optimizer, that chooses u first, has "nature" as an adverse player that chooses afterwards p and x. Whatever u the optimizer chooses, for each of the functions $f_i(x, u, p), i = 0, \ldots, n_f$, the worst case, $\phi_i(u)$, is chosen by the adverse player by selecting a suitable $p \in \mathbb{P}$:

$$\begin{aligned} &\phi_i(u) := \max_{p \in \mathbb{P}(u)} f_i(x(u,p), u, p) \\ (\text{WC}) \qquad &= \max_{x \in \mathbb{R}^{n_x}, p \in \mathbb{R}^{n_p}} f_i(x, u, p) \quad \text{s.t.} \quad \begin{cases} g(x, u, p) = 0, \\ h(x, u, p) \leq 0. \end{cases} \end{aligned} \tag{3}$$

Note that the adverse player "nature" is restricted by both the model equations $g(x, u, p) = 0$ and the scalar inequality constraint $h(x, u, p) \leq 0$. Employing the functions $\phi_i(u)$ we arrive at the following worst-case formulation that is often referred to as the "robust counterpart" of (1) (cf. Ben-Tal & Nemirovskii, 2001):

$$(\text{RC}) \quad \min_{u \in \mathbb{R}^{nu}} \phi_0(u) \quad \text{s.t.}\, \phi_i(u) \leq 0; \quad \text{for} \quad i = 1, \ldots, n_f. \tag{4}$$

Due to the bi-level structure and the semi-infinite character, both formulations (1) and (RC) pose challenges for their efficient numerical solution. Different approaches to tackle these problems have been presented in a large number of articles. Methods to solve GSIP include discretization of the uncertainty set $\mathbb{P}(u)$ and the so called local reduction approach (Hettich & Kortanek, 1993). In both approaches the GSIP is approximated by a NLP with a finite number of constraints. In discretization methods problem (1) is solved on a finite grid of points $\bar{\mathbb{P}}(u) \subset \mathbb{P}(u)$ within the uncertainty set. The local reduction approach is based on the worst-case min–max formulation (RC). In this approach the constraints are reduced to a finite number of restrictions by considering only the (local) solutions of the inner maximization problem (WC), which are implicitly defined by the necessary conditions of optimality of (WC). In the locally reduced problem of (RC) then a finite number of worst cases within $\mathbb{P}(u)$ is tracked. For a series of comprehensive review articles on different solution strategies of GSIP-problems, including discretization and local reduction methods we refer to the recent monograph (Reemtsen & Rückmann, 1998).

Approaches to problem (1) have also been addressed by a number of articles on flexible and feasible process design. Measures of feasibility and flexibility have been introduced by Halemane and Grossmann (1983) and Swaney and Grossmann (1985). In this approach the optimization variables are partitioned into design and control variables. The design variables are specified by the outer minimization problem 1. Flexibility and feasibility in the presence of parametric uncertainty is addressed by formulating a nested max–min–max constraint for the feasibility constraints f, where the control variables are chosen by the minimization problem and the uncertain parameters by the outer maximization problem. The flexibility and feasibility measures were the basis for the development of various robust optimization methods in a series of papers, e.g., a two level optimization approach (Bahri, Bandoni, & Romagnoli, 1996) or an active set strategy (Mohideen, Perkins, & Pistikopoulos, 1996), which also considers worst-case dynamic disturbances. A different approach to robust optimization not based on the feasibility and flexibility measures was developed in a series of papers by Mönnigmann and Marquardt (2002, 2003, 2005). In this approach parametric uncertainty is taken into account by backing off the optimal design from critical boundaries in the space of the uncertain parameters. It furthermore allows the simultaneous treatment of feasibility and stability.

In this work we adopt the local reduction approach and present a numerical approach for the solution of (RC). We reformulate (RC) by introducing the necessary conditions of optimality of the inner maximization problem (WC). Due to the local character of the necessary conditions of optimality several worst-case points may exist for each function f_i, $i = 0, \ldots, n_f$. Afterwards we simplify the resulting constraints with the assumption that the worst-case point is always located on the *boundary* of the uncertainty set $\mathbb{P}(u)$. Solving the resulting NLP is then equivalent to track worst-case solutions situated on the boundary of the uncertainty set $\mathbb{P}(u)$. The local reduction approach is an efficient approach to rigorously solve (RC) also in case of a robust optimal control problem by tracking worst-case trajectories.

The aim in robust optimal control is to find an optimal profile of the manipulated variables such that none of the specified constraints are violated despite model uncertainties. Robust optimal control is also addressed by robust nonlinear model predictive control (NMPC). NMPC solves on-line repeatedly a dynamic optimization problem on a shrinking time horizon (Diehl, Bock, & Schlöder, 2005; Nagy & Braatz, 2003; Terwiesch, Agarwal, & Rippin, 1994) or on a moving horizon (Biegler, 2000; Binder et al., 2001; Diehl et al., 2002). For an assessment of the state of the art in this very active field we refer to Findeisen, Allgöwer, and Biegler (2006). Robust optimal control and robust NMPC is approached by Ma and Braatz (2001), Nagy and Braatz (2003), and Diehl, Bock, and Kostina (2006). For robust optimization, a linear or higher order approximation of the uncertain objective and constraints is used to faciliate the solution of the robust counterpart NLP (RC), involving some approximation errors. The algorithms of Ma and Braatz (2001) and Nagy and Braatz (2003) contain comparisons with full nonlinear uncertainty simulations. However, there it is not indicated how to numerically solve the exact robust counterpart NLP efficiently.

We have to mention here that the minimax approach for robust open-loop optimization is often considered to be too conservative to be useful in practice, as discussed, e.g., by Morari (1983). On the other hand the minimax approach allows to rigorously quantify the profit loss that has to be paid for the robustification

Author's personal copy

of the open-loop optimal solution. If the impact of the uncertainty on the profit is small, the open-loop operation of a batch process represents an alternative to closed-loop on-line optimization schemes, which require additional effort and cost to tackle the higher computational burden. Finally, in the case that no measurement data is available for a batch process on-line schemes cannot be employed, and there are no alternatives to robust open-loop optimization in the presence of uncertainty.

The paper is organized as follows. In Section 2 we introduce the necessary conditions of optimality of the underlying maximization problem and elaborate on the assumption that the worst case is always located on the boundary of the uncertainty set $\mathbb{P}(u)$. We show that the resulting NLP is closely related to the robust optimization approach presented by Mönnigmann and Marquardt (2002, 2003, 2005). In Section 3 the NLP is generalized to an optimal control problem. In Section 4 we apply the presented approach to the robust dynamic optimization of a batch distillation column.

2. Solution by tracking of worst-case solutions

In this section we reformulate the NLP (1) by substituting the inner maximization problem (WC) by its necessary conditions of optimality. We then introduce the simplifying assumption that the worst case will always be located on the boundary of the uncertainty set $\mathbb{P}(u)$. Two different formulations of the resulting NLP are presented.

2.1. Formulation of necessary optimality conditions

The maximizer of the worst-case problem (WC), repeated here for convenience

$$\max_{x\in\mathbb{R}^{n_x},\, p\in\mathbb{R}^{n_p}} f_i(x,u,p), \quad \text{s.t.} \begin{cases} g(x,u,p)=0, \\ h(x,u,p)\le 0, \end{cases}$$

is achieved in a point x_i, p_i that can be characterized by the necessary optimality conditions of first order. These state that there exists a multiplier vector $\lambda_i \in \mathbb{R}^{n_x}$ and a scalar ν_i such that[1]

$$\nabla_p f_i(x_i,u,p_i) + \nabla_p g(x_i,u,p_i)\lambda_i = \nabla_p h(x_i,u,p_i)\nu_i \tag{5a}$$

$$\nabla_x f_i(x_i,u,p_i) + \nabla_x g(x_i,u,p_i)\lambda_i = \nabla_x h(x_i,u,p_i)\nu_i \tag{5b}$$

$$g(x_i,u,p_i)=0, \tag{5c}$$

$$h(x_i,u,p_i)\nu_i = 0, \tag{5d}$$

$$\nu_i \ge 0, \tag{5e}$$

$$h(x_i,u,p_i)\le 0. \tag{5f}$$

The $n_p + 2n_x + 1$ equality constraints (5a)–(5d) determine the $n_p + 2n_x + 1$ unknowns p_i, λ_i, x_i, and ν_i, such that $p_i = p_i(u)$ by virtue of the implicit function theorem. Note that the necessary conditions of optimality (5) are only local, i.e., there may exist more than one local maximum $p_{i,j}$, $\lambda_{i,j}$, $x_{i,j}$, $\nu_{i,j}$ with $j = 1,\ldots,J_i$, where we assume that $J_i < \infty$, for each f_i within $\mathbb{P}(u)$. We will discuss and illustrate this issue in more detail below (cf. Fig. 3(b), Section 2.5).

We now want to use Eqs. (5) as constraints of the outer minimization problem (1), transforming the GSIP into a locally reduced finite NLP. The optimality conditions (5), however, include the complementarity constraint (5d)–(5f). It is well known that NLPs with complementarity constraints are difficult to solve (Scheel & Scholtes, 2000). To eliminate the complementary constraint we therefore introduce the following assumption:

Assumption 1. For any control $u \in \mathbb{R}^{n_u}$ and function f_i, the maximum in (WC) is attained in a point $p^*(u,i)$ on the boundary of $\mathbb{P}(u)$, i.e., it satisfies

$$h(x(u,p^*(u,i)),u,p^*(u,i)) = 0.$$

Because of this assumption we can drop the factor ν_i in Eq. (5d) and omit Eq. (5f). Assumption 1 is in general impossible to check but we claim that it is likely to be satisfied in applications of interest. It has also been used by Yu and Ishii (1998) where the worst case is assumed to be on the boundary of a joint confidence ellipsoid. Please note that we implicitly assume that the boundary of the uncertainty set $\mathbb{P}(u)$ is smooth, such that, e.g., box uncertainties – that would lead to an NLP with complementarity constraints (Scheel & Scholtes, 2000) – cannot be treated. In that case, however, a similar but more restrictive assumption would be that the worst case is found on one of the corners of the hyperrectangular uncertainty set (e.g., Halemane & Grossmann, 1983).

Our assumption is justified if we focus our interest to smooth compact convex sets $\mathbb{P}(u)$ (the confidence ellipsoid being the most important application we have in mind). In practice, the assumption is true for an active constraint $f_i^* = 0$. In this case the maximum must be located on the border of $\mathbb{P}(u)$ to ensure that the constraint f_i is fulfilled for all values within the region of parametric uncertainty. If the constraint f_i has a local maximum $f_i^* < 0$, however, Assumption 1 might force this local maximum to the boundary of the uncertainty set even though the local maxima could be located inside $\mathbb{P}(u)$ without introducing infeasible points within $\mathbb{P}(u)$. This problem can, however, easily be avoided by tracking only active constraints $f_i^* = 0$ on the boundary of $\mathbb{P}(u)$, neglecting constraints with $f_i^* < 0$.

2.2. Adjoint based robust counterpart formulation

As mentioned above we can omit Eq. (5f) and the factor ν_i in Eq. (5d). This allows us to formulate a large, but structured nonlinear programming problem, that is equivalent to the robust counterpart (RC) and comprises, besides the control degrees of freedom u, also

$$n_c = \sum_{i=0}^{n_f} J_i \ge n_f + 1$$

[1] Throughout this report we use the convention that for any function $f : \mathbb{R}^n \to \mathbb{R}^m$ the matrix $\nabla_x f(x)$ equals the transposed Jacobian: $\nabla_x f(x) := \partial f/\partial x(x)^{\mathrm{T}}$.

Author's personal copy

worst-case values $x_{i,j} \in \mathbb{R}^{n_x}$, $p_{i,j} \in \mathbb{R}^{n_p}$ with adjoint variables $\lambda_{i,j} \in \mathbb{R}^{n_x}$ and $\nu_{i,j}$:

$$\min_{u,x_0,p_0,x_{i,j},p_{i,j},\lambda_{i,j},\nu_{i,j}} \quad \max_{j=1,\ldots,J_0} f_0(x_{0,j}, u, p_{0,j}) \tag{6a}$$

$$\text{s.t.} \quad f_i(x_{i,j}, u, p_{i,j}) \le 0, \quad i=1,\ldots,n_f, \quad j=1,\ldots,J_i, \tag{6b}$$

$$g(x_{i,j}, u, p_{i,j}) = 0, \quad i = 0,\ldots,n_f, \quad j = 1,\ldots,J_i, \tag{6c}$$

$$h(x_{i,j}, u, p_{i,j}) = 0, \tag{6d}$$

$$\nabla_p f_i(x_{i,j}, u, p_{i,j}) + \nabla_p g(x_{i,j}, u, p_{i,j})\lambda_{i,j} = \nabla_p h(x_{i,j}, u, p_{i,j})\nu_{i,j}, \tag{6e}$$

$$\nabla_x f_i(x_{i,j}, u, p_{i,j}) + \nabla_x g(x_{i,j}, u, p_{i,j})\lambda_{i,j} = \nabla_x h(x_{i,j}, u, p_{i,j})\nu_{i,j}, \tag{6f}$$

$$\nu_{i,j} \ge 0. \tag{6g}$$

2.3. Interpretation of gradients as normal vectors

By inverting $\nabla_x g(x_{i,j}, u, p_{i,j})$ in Eq. (6f) to eliminate $\lambda_{i,j}$ in Eq. (6e), we could write Eqs. (6e) and (6f) equivalently as

$$\nabla_p f_i(x_{i,j}, u, p_{i,j}) - \nabla_p g(x_{i,j}, u, p_{i,j})\nabla_x g(x_{i,j}, u, p_{i,j})^{-1} \times \nabla_x f_i(x_{i,j}, u, p_{i,j}) = (\nabla_p h(x_{i,j}, u, p_{i,j}) - \nabla_p g(x_{i,j}, u, p_{i,j}) \times \nabla_x g(x_{i,j}, u, p_{i,j})^{-1}\nabla_x h(x_{i,j}, u, p_{i,j}))\nu_{i,j}. \tag{7}$$

Both sides of this equality can be interpreted as normal vectors to two $(n_p - 1)$ dimensional manifolds in $\mathbb{R}^{n_p}$, the space of parameters. The left hand side corresponds to a normal vector of the set $M^{\mathrm{c}}_{i,j}$ of constant values of f_i:

$$M^{\mathrm{c}}_{i,j}(u) := \{p | \exists x : g(x, u, p) = 0,\ f_i(x, u, p) = f_i(x_{i,j}, u, p_{i,j})\}.$$

The right hand side vector corresponds to a normal vector of the set $M^{\mathrm{r}}_{i,j}$ of constant values of h:

$$M^{\mathrm{r}}_{i,j}(u) := \{p | \exists x : g(x, u, p) = 0, h(x, u, p) = h(x_{i,j}, u, p_{i,j})\}.$$

Because we require $h(x_{i,j}, u, p_{i,j}) = 0$ in the optimization, this second manifold is independent of the indices i, j, and we can simply define

$$M^{\mathrm{r}}_{i,j}(u) = M^{\mathrm{r}}(u) := \{p | \exists x : g(x, u, p) = 0, h(x, u, p) = 0\}.$$

Note that this is the surface of the set $\mathbb{P}(u)$ of possible parameter variations.

It is interesting to compare the NLP (6) with the robust optimization approach developed by Mönnigmann and Marquardt (2002, 2003, 2005). While here we are tracking, for each local maximum of the problem function f_i, only *one* point $p_{i,j}$ on the surface $M^{\mathrm{r}}(u)$ of the parameter disturbance set $\mathbb{P}(u)$, Mönnigmann and Marquardt track *two* points $p^{\mathrm{r}}_{i,j}$, $p^{\mathrm{c}}_{i,j}$. For a visualization, see Fig. 1. The first point, $p^{\mathrm{r}}_{i,j}$, is also on the surface $M^{\mathrm{r}}(u)$, but the second one, $p^{\mathrm{c}}_{i,j}$, is the point on the zero level set surface

$$M^{\mathrm{c},0}_i(u) := \{p | \exists x : g(x, u, p) = 0,\ f_i(x, u, p) = 0\}$$

which is closest to the surface $M^{\mathrm{r}}(u)$. Note that this zero level set surface may also be called the "critical manifold" (Mönnigmann & Marquardt, 2002). The surface M^{r} of the uncertainty set may be denoted accordingly as the "robustness manifold". The two points are determined by the condition that the normal vector direction of the critical manifold $n^{\mathrm{c}}_{i,j}$ and the normal vector direction of the robustness manifold $n^{\mathrm{r}}_{i,j}$ point in the same direction. Robustness is ensured by requiring that the manifolds keep a non-negative distance $l_{i,j}$

$$p^{\mathrm{c}}_{i,j} - p^{\mathrm{r}}_{i,j} = l_{i,j} n_{i,j}; \quad l_{i,j} n_{i,j} \ge 0,\ i = 1,\ldots,n_f,\ j = 1,\ldots,J_i. \tag{8}$$

Because of the nonlinear character of the critical manifold more than one closest connection may exist for one constraint f_i, i.e., $J_i \ge 1$ (Mönnigmann & Marquardt, 2003). The objective is not included in (8) as f_0 is not bounded and there is no critical manifold associated with f_0.

In order to show the close relationship of NLP (6) to the normal vector approach we briefly introduce the normal vector constraints. The normal space of the manifold $M^{\mathrm{c},0}_i$ is spanned by the gradient of the defining equations

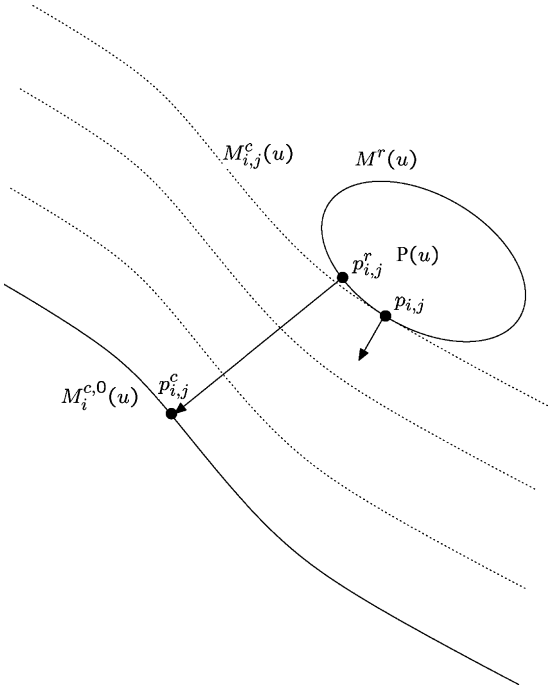


Fig. 1. Sketch of the critical manifold $M^{\mathrm{c},0}_i$ (left bottom) in the parameter space $\mathbb{R}^{n_p}$, along with some other (decreasing) level sets of f_i (dotted), and the boundary $M^{\mathrm{r}}(u)$ of the parameter set $\mathbb{P}(u)$ (ellipsoid). The large arrow connects the two closest points $p^{\mathrm{r}}_{i,1} \in M^{\mathrm{r}}$ and $p^{\mathrm{c}}_{i,1} \in M^{\mathrm{c},0}_i$ of the approach by Mönnigmann and Marquardt (2002, 2003, 2005). It also equals the respective normal vector on both manifolds. The small arrow is the normal vector at the worst-case point $p_{i,1} \in M^{\mathrm{r}}$ that is tracked in our approach. It is normal both to M^{r} and to the f_i level set $M^{\mathrm{c}}_{i,j}$.

Author's personal copy

$g = 0$, $f_i = 0, i = 1, \ldots, n_f$ with respect to the variables $(x^c_{i,j}, p^c_{i,j})$ for a given control variable vector u:

$$B^c = \begin{bmatrix} \nabla_x g(x^c_{i,j}, u, p^c_{i,j}) & \nabla_x f_i(x^c_{i,j}, u, p^c_{i,j}) \\ \nabla_p g(x^c_{i,j}, u, p^c_{i,j}) & \nabla_p f_i(x^c_{i,j}, u, p^c_{i,j}) \end{bmatrix}.$$

As we are interested in the normal direction in the subspace of parameters p we chose $\kappa^c_{i,j} \in \mathbb{R}^{n_x+1}$ such that the linear combination of the first n_x rows of B^c with $\kappa^c_{i,j}$ is 0

$$[\nabla_x g(x^c_{i,j}, u, p^c_{i,j}) \quad \nabla_x f_i(x^c_{i,j}, u, p^c_{i,j})]\kappa^c_{i,j} = 0,$$
$$\kappa^{cT}_{i,j} z - 1 = 0,$$

with $z \in \mathbb{R}^{n_x+1}$ not normal to $\kappa^c_{i,j}$. This means that $\kappa^c_{i,j}$ is in the kernel of $[\nabla_x g, \nabla_x f_i]$, the $n_x \times n_{x+1}$ submatrix of B^c. The set of $n_x + 1$ equations defines the $n_x + 1$ elements of $\kappa^c_{i,j}$. The normal vector $n^c_{i,j}$ in the parameter subspace is then given by

$$[\nabla_p g(x^c_{i,j}, u, p^c_{i,j}) \quad \nabla_p f_i(x^c_{i,j}, u, p^c_{i,j})]\kappa^c_{i,j} = n^c_{i,j}.$$

The normal vector direction for the robustness manifold n^r is defined analogously:

$$[\nabla_x g(x^r_{i,j}, u, p^r_{i,j}) \quad \nabla_x h(x^r_{i,j}, u, p^r_{i,j})]\kappa^r_{i,j} = 0,$$
$$[\nabla_p g(x^r_{i,j}, u, p^r_{i,j}) \quad \nabla_p h(x^r_{i,j}, u, p^r_{i,j})]\kappa^r_{i,j} = n^r_{i,j},$$
$$\kappa^{rT}_{i,j}\tilde{z} - 1 = 0,$$

with $\kappa^r_{i,j} \in \mathbb{R}^{n_x+1}$ and $\tilde{z} \in \mathbb{R}^{n_x+1}$ not normal to $\kappa^r_{i,j}$. As both normal vector directions must point in the same direction it must hold that

$$n^r_{i,j}\rho_{i,j} = n^c_{i,j}, \quad \rho_{i,j} \in \mathbb{R}.$$

The complete nonlinear program including the normal vector constraints then reads

$$\min_{u, x^r_{i,j}, p^r_{i,j}, x^c_{i,j}, p^c_{i,j}, \kappa^c_{i,j}, \kappa^r_{i,j}, \rho_{i,j}, l_{i,j}, n^c_{i,j}, \nu_{0,j}, \lambda_{0,j}} \max_{j=1,\ldots,J_0} f_0(x_{0,j}, u, p_{0,j})$$
$$\text{s.t.} \quad g(x^r_{1,j}, u, p^r_{1,j}) = 0,$$
$$h(x^r_{1,j}, u, p^r_{1,j}) = 0, \tag{9a}$$

$$\nabla_p f_0(x_{0,j}, u, p_{0,j}) + \nabla_p g(x_{0,j}, u, p_{0,j})\lambda_{0,j} = \nabla_p h(x_{0,j}, u, p_{0,j})\nu_{0,j},$$
$$\nabla_x f_0(x_{0,j}, u, p_{0,j}) + \nabla_x g(x_{0,j}, u, p_{0,j})\lambda_{0,j} = \nabla_x h(x_{0,j}, u, p_{0,j})\nu_{0,j}, \tag{9b}$$
$$\nu_{0,j} \geq 0,$$

$$\rho_{i,j}(\nabla_x g(x^r_{i,j}, u, p^r_{i,j})\hat{\kappa}^r_{i,j} + \nabla_x h(x^r_{i,j}, u, p^r_{i,j})\tilde{\kappa}^r_{i,j}) = 0,$$
$$\rho_{i,j}(\nabla_p g(x^r_{i,j}, u, p^r_{i,j})\hat{\kappa}^r_{i,j} + \nabla_p h(x^r_{i,j}, u, p^r_{i,j})\tilde{\kappa}^r_{i,j}) = n^c_{i,j}, \tag{9c}$$
$$[\hat{\kappa}^{rT}_{i,j}, \tilde{\kappa}^r_{i,j}]\tilde{z} - 1 = 0,$$

$$g(x^c_{i,j}, u, p^c_{i,j}) = 0,$$
$$f_i(x^c_{i,j}, u, p^c_{i,j}) = 0,$$
$$\nabla_x g(x^c_{i,j}, u, p^c_{i,j})\hat{\kappa}^c_{i,j} + \nabla_x f_i(x^c_{i,j}, u, p^c_{i,j})\tilde{\kappa}^c_{i,j} = 0, \tag{9d}$$
$$\nabla_p g(x^c_{i,j}, u, p^c_{i,j})\hat{\kappa}^c_{i,j} + \nabla_p f_i(x^c_{i,j}, u, p^c_{i,j})\tilde{\kappa}^c_{i,j} = n^c_{i,j},$$
$$[\hat{\kappa}^{cT}_{i,j}, \tilde{\kappa}^c_{i,j}]z - 1 = 0,$$

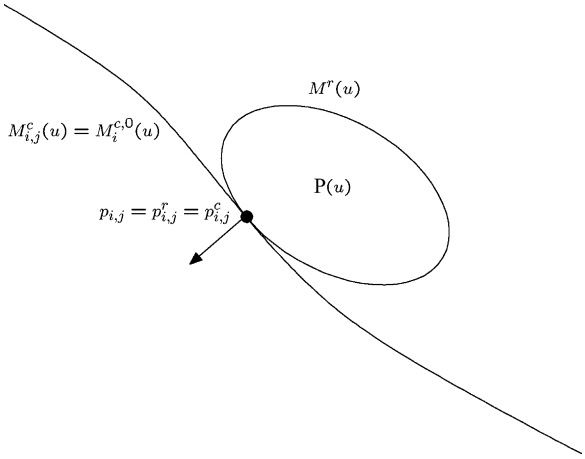


Fig. 2. If the manifolds $M^{c,0}_i(u)$ and $M^r(u)$ touch, the closest connecting points $p^r_{i,j}$, $p^c_{i,j}$ of Mönnigmann and Marquardt (2002, 2003, 2005) coincide with the worst-case point $p_{i,j}$ tracked in our approach (cf. Fig. 1). The normal vectors at $p_{i,j}$ with respect to both manifolds point into the same direction.

$$p^c_{i,j} - p^r_{i,j} + l_{i,j}n^c_{i,j} = 0,$$
$$l_{i,j} \geq 0, \tag{9e}$$

with indices $\hat{i} = 0, \ldots, n_f$, $i = 1, \ldots, n_f$ and $j = 1, \ldots, J_{\hat{i}}$, $j = 1, \ldots, J_i$, respectively. For the objective f_0 the necessary conditions of optimality (9b) are used, as there is no critical manifold associated with f_0. The normal vector direction $n^r_{i,j}$ of the robustness manifold is defined by Eq. (9c) and the normal vector direction $n^c_{i,j}$ of the critical manifold by Eq. (9d). Both $\kappa^r_{i,j}$ and $\kappa^c_{i,j}$ have been split in $\hat{\kappa}_{i,j} \in \mathbb{R}^{n_x}$ and $\tilde{\kappa}_{i,j} \in \mathbb{R}$. To simplify comparison between NLP (6) and NLP (9) the first equation of (9c) has been multiplied by $\rho_{i,j}$. For constraints that are active at the solution $l_{i,j} = 0$ and correspondingly $p^c_{i,j} = p^r_{i,j} = p_{i,j}$ and $x^c_{i,j} = x^r_{i,j} = x_{i,j}$, both formulations (9) and (6) are equivalent, as illustrated in Fig. 2. This can easily be checked by subtracting the first and second equation of (9c) from the third and fourth equation of (9d). The resulting terms are equivalent to Eqs. (6e) and (6f) with

$$\lambda_{i,j} = \frac{\hat{\kappa}^c_{i,j} - \rho_{i,j}\hat{\kappa}^r_{i,j}}{\tilde{\kappa}^c_{i,j}}, \qquad \nu_{i,j} = \frac{\rho_{i,j}\tilde{\kappa}^r_{i,j}}{\tilde{\kappa}^c_{i,j}}.$$

On the other hand, when $p^c_{i,j} \neq p^r_{i,j}$, our point $p_{i,j}$ would be at a similar, but not identical position compared to $p^r_{i,j}$, as seen in Fig. 1. However, in both cases it is guaranteed that the critical manifold $M^{c,0}_i(u)$ and $\mathbb{P}(u)$ do not intersect, and both formulations lead to the same robust optimization results.

2.4. An equivalent formulation based on direct sensitivities

Eq. (7) can be used to formulate an alternative NLP based on direct sensitivities by substituting Eqs. (6e) and (6f) of the NLP (6). Introducing the matrix $D_{i,j} \in \mathbb{R}^{n_x \times n_p}$

$$D_{i,j} = -(\nabla_p g(x_{i,j}, u, p_{i,j})\nabla_x g(x_{i,j}, u, p_{i,j})^{-1})^T, \tag{10}$$

Author's personal copy

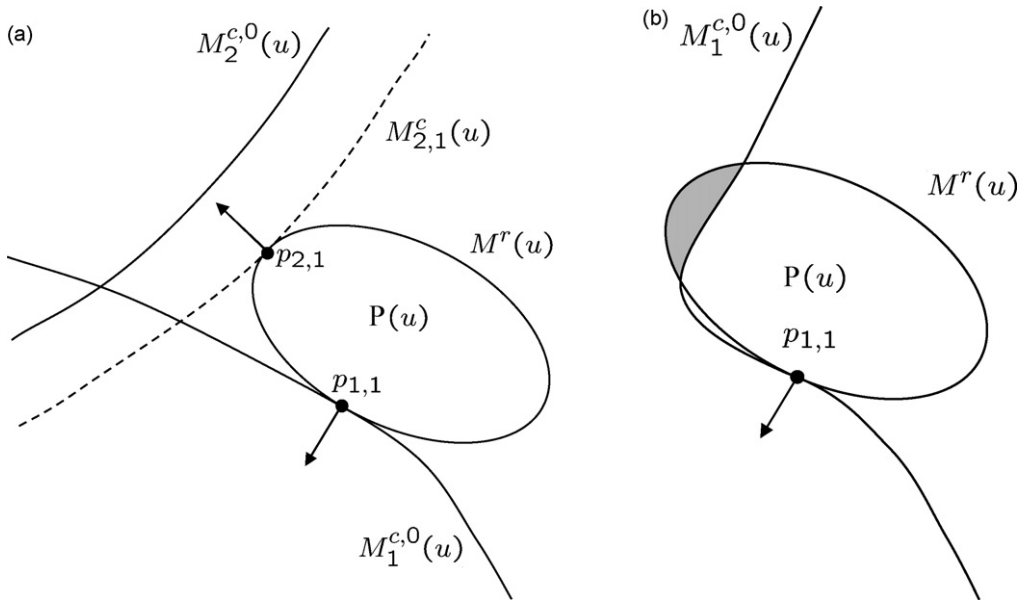


Fig. 3. (a) Robustness constraints (6b)–(6g) or (11b)–(11g) must be considered for each constraint $f_i \le 0,\ i = 1, \ldots, n_f$. (b) In case of large curvature of $M_i^{c,0}(u)$ more than one critical point must be tracked on the boundary $M^r(u)$ of the uncertainty set $\mathbb{P}(u)$ to guarantee robustness of the optimal solution.

we can simplify Eq. (7) and then rewrite NLP (6) according to

$$\min_{u, x_0, p_0, x_{i,j}, p_{i,j}, \lambda_{i,j}, \nu_{i,j}} \max_{j=1,\ldots,J_0} f_0(x_{0,j}, u, p_{0,j}) \tag{11a}$$

$$\text{s.t.} \quad f_i(x_{i,j}, u, p_{i,j}) \le 0, \quad i = 1, \ldots, n_f,\ j = 1, \ldots, J_i \tag{11b}$$

$$g(x_{i,j}, u, p_{i,j}) = 0, i = 0, \ldots, n_f,\ j = 1, \ldots, J_i, \tag{11c}$$

$$h(x_{i,j}, u, p_{i,j}) = 0, \tag{11d}$$

$$\nabla_p f_i + D_{i,j}^{\mathrm{T}} \nabla_x f_i|_{(x_{i,j}, u, p_{i,j})} = \nu_{i,j} (\nabla_p h + D_{i,j}^{\mathrm{T}} \nabla_x h)|_{(x_{i,j}, u, p_{i,j})}, \tag{11e}$$

$$\nabla_x g(x_{i,j}, u, p_{i,j})^{\mathrm{T}} D_{i,j} + \nabla_p g(x_{i,j}, u, p_{i,j})^{\mathrm{T}} = 0, \tag{11f}$$

$$\nu_{i,j} \ge 0. \tag{11g}$$

2.5. Steps required for solving NLP (6) or (11)

In this section we summarize the steps required to solve NLPs of the form (6) or (11). The steps necessary to solve these problems were first proposed in the algorithm presented by Mönnigmann and Marquardt (2005). Among the necessary steps are the initialization of the robust counterpart constraints (6b)–(6g) or (11b)–(11g) of the NLP and the detection of critical points. Note that at least one set of robust counterpart constraints for each $f_i, i = 0, \ldots, n_f$ has to be included in the NLP. In Fig. 3 (a) the situation is shown for two constraints, f_1 and f_2, with $f_1(p_{1,1}) = 0$ being active and $f_2(p_{2,1}) < 0$.

Generally, a feasible initialization of the NLPs (6) or (11) satisfying the robust counterpart constraints (6b)–(6g) or (11b)–(11g) is not known beforehand. A feasible point is found by solving the inner maximization problem (WC) for all constraints f_i with $h = 0$ by Assumption 1 and fixed u. At the initialization step it is also usually not known if more than one maximum exists on surface $M^r(u)$ for f_i, in general therefore, $J_i = 1$ for $i = 0, \ldots, n_f$.

After convergence of the optimization problem it is necessary to check that there are no points within the uncertainty set $\mathbb{P}(u)$ that violate the constraints f_i. According to the example sketched in Fig. 3(b) infeasible points may be located within $\mathbb{P}(u)$ if critical manifolds $M_i^{c,0} = 0$ are strongly nonlinear and bend into $\mathbb{P}(u)$, or if $f_i = 0$ is fulfilled on a second insular type of manifold within the uncertainty set (not shown in Fig. 3(b)). In this case two or more critical points have to be tracked for one constraint, i.e., $J_i > 1$. The NLP (6) or (11) then has to be solved again with updated J_i. Mönnigmann and Marquardt (2005) showed that a search for critical points can be implemented by checking a grid of points within the uncertainty set $\mathbb{P}(u)$ for constraint violations $f_i > 0, i = 1, \ldots, n_f$. These authors further showed that this type of search is feasible for problems with a number of uncertain parameters beyond ten. A rigorous search for critical points within the uncertainty set can be implemented with interval arithmetics (Mönnigmann et al., 2007). Such a rigorous search is, however, computationally expensive and therefore only feasible for problems with a few uncertain parameters.

Summarizing, our robust optimization approach proceeds along the same iterative algorithm presented by Mönnigmann and Marquardt (2005) that consists of three steps:

Algorithm 1.

(i) initialize the NLP for all f_i, usually starting with $J_i = 1$,
(ii) solve NLP (6) or (11),
(iii) check for critical points within the robustness set $\mathbb{P}(u)$.

Author's personal copy

If new critical points are detected, J_i has to be updated and step (i) has to be repeated to initialize the detected critical points. If no further critical points are found in step (iii), an optimum is found at which constraints $f_i \leq 0$ robustly hold despite uncertainties in p. The main difference between the algorithm presented here and the algorithm presented by Mönnigmann and Marquardt (2005) for the normal vector approach is the first initialization step. In the normal vector approach generally no points on the critical manifolds $M_i^{c,0}$ are known beforehand and the first iteration starts with $J_i = 0, i = 1, \ldots, n_f$, whereas initialization for the new approach can be realized by evaluating the inner maximization problem (WC) for every f_i such that $J_i = 1, i = 0, \ldots, n_f$.

Note that for a large number of constraints f_i it is unlikely that all constraints are active at the optimum. A switch-off threshold $f_i < \underline{f}_i$ can therefore be introduced to remove those robustness constraints from the NLP which are far away from the critical boundary $M_i^{c,0}$. Checking this threshold between steps (iii) and (i) helps to reduce the number of constraints of the NLP.

3. Generalization to optimal control problems

We will briefly describe how the presented approach can be applied in the context of optimal control problems. We regard a simplified optimal control problem in ordinary differential equations on the (variable) time horizon $[0, T]$ of the following form:

$$\min_{T,u(\cdot),x(\cdot)} f_0(x(T), T) \quad \text{subject to} \tag{12a}$$

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = G(x(t), u(t)), \quad \forall t \in [0, T], \tag{12b}$$

$$x(0) = x_0(p), \tag{12c}$$

$$f_i(x(T), T) \leq 0, \quad i = 1, \ldots, n_f, \tag{12d}$$

with parameterized initial values $x_0(p) \in \mathbb{R}^{n_x}$. We consider the horizon length T and the control profile $u : [0, T] \to \mathbb{R}^{n_u}$ as decision variables and the trajectory $x : [0, T] \to \mathbb{R}^{n_x}$ as the dependent variables. The latter are chosen by "nature", which is only restricted by the model equations. (12b), the initial value constraint (12c), and the simple scalar bound $h(p) \leq 0$ on the parameters $p \in \mathbb{R}^{n_p}$, i.e., $p \in \mathbb{P} = \{p | h(p) \leq 0\}$. Note that problem (12) includes ODEs with uncertainty $(\mathrm{d}x/\mathrm{d}t)(t) = G(x(t), u, \tilde{p})$ by adding trivial differential equations $\mathrm{d}\tilde{p}/\mathrm{d}t = 0$ with initial values defined by uncertain parameters $\tilde{p}(0) = p$.

We show two equivalent ways to formulate the robust counterpart of the above optimal control problem, first by using direct derivatives and second using adjoint sensitivities.

3.1. Direct robust counterpart formulation

Analogous to the second robust counterpart formulation (11), we obtain the following optimal control problem, where the n_c time dependent system states $x_{i,j}(t) \in \mathbb{R}^{n_x}$ are augmented by n_c time dependent sensitivity matrices $D_{i,j}(t) \in \mathbb{R}^{n_x \times n_p}$,

$$\min_{T,u(\cdot),p_{i,j},x_{i,j}(\cdot),D_{i,j}(\cdot),v_{i,j}} \max_{j=1,\ldots,J_0} f_0(x_{0,j}(T), T) \quad \text{subject to} \tag{13a}$$

$$f_i(x_{i,j}(T), T) \leq 0, \quad i = 1, \ldots, n_f, \ j = 1, \ldots, J_i, \tag{13b}$$

$$\frac{\mathrm{d}x_{i,j}}{\mathrm{d}t}(t) = G(x_{i,j}(t), u(t)), \quad t \in [0, T], i = 0, \ldots, n_f, \ j = 1, \ldots, J_i, \tag{13c}$$

$$x_{i,j}(0) = x_0(p_{i,j}), \tag{13d}$$

$$h(p_{i,j}) = 0, \tag{13e}$$

$$D_{i,j}(T)^{\mathrm{T}} \nabla_x f_i(x_{i,j}(T), T) = v_{i,j} \nabla_p h(p_{i,j}) \tag{13f}$$

$$\frac{\mathrm{d}D_{i,j}}{\mathrm{d}t}(t) = \frac{\partial G}{\partial x}(x_{i,j}(t), u(t)) D_{i,j}(t), \tag{13g}$$

$$D_{i,j}(0) = \frac{\partial x_0}{\partial p}(p_{i,j}), \tag{13h}$$

$$v_{i,j} \geq 0. \tag{13i}$$

Note that the matrix $D_{i,j}(t) \in \mathbb{R}^{n_x \times n_p}$ can be regarded as the derivative

$$D_{i,j}(t) = \frac{\mathrm{d}x_{i,j}(t)}{\mathrm{d}x_{i,j}(0)} \frac{\partial x_0}{\partial p}(p_{i,j}),$$

where we denote by $\mathrm{d}x_{i,j}(t)/\mathrm{d}x_{i,j}(0)$ the linearized dependence of the state $x_{i,j}(t)$ at time t on the initial state $x_{i,j}(0)$.

3.2. Adjoint robust counterpart formulation

Analogous to the adjoint sensitivity formulation (6) we can alternatively formulate an optimal control problem where the n_c time dependent system state trajectories $x_{i,j}(t) \in \mathbb{R}^{n_x}$ are augmented by n_c time dependent adjoint sensitivities $\lambda_{i,j}(t) \in \mathbb{R}^{n_x}$:

$$\min_{T,u(\cdot),p_{i,j},x_{i,j}(\cdot),\lambda_{i,j}(\cdot),v_{i,j}} \max_{j=1,\ldots,J_0} f_0(x_{0,j}(T), T) \quad \text{subject to} \tag{14a}$$

$$f_i(x_{i,j}(T), T) \leq 0, \quad i = 1, \ldots, n_f, \ j = 1, \ldots, J_i, \tag{14b}$$

$$\frac{\mathrm{d}x_{i,j}}{\mathrm{d}t}(t) = G(x_{i,j}(t), u(t)), \quad t \in [0, T], \ i = 0, \ldots, n_f, \ j = 1, \ldots, J_i, \tag{14c}$$

$$x_{i,j}(0) = x_0(p_{i,j}), \tag{14d}$$

$$h(p_{i,j}) = 0, \tag{14e}$$

$$\nabla_p x_0(p_{i,j}) \lambda_{i,j}(0) = v_{i,j} \nabla_p h(p_{i,j}) \tag{14f}$$

$$\frac{\mathrm{d}\lambda_{i,j}}{\mathrm{d}t}(t) = -\nabla_x G(x_{i,j}(t), u(t)) \lambda_{i,j}(t), \tag{14g}$$

$$\lambda_{i,j}(T) = \nabla_x f_i(x_{i,j}(T), T), \tag{14h}$$

$$v_{i,j} \geq 0. \tag{14i}$$

Author's personal copy

Here, the adjoint variable $\lambda_{i,j}(t) \in \mathbb{R}^{n_x}$ can be interpreted as

$$\lambda_{i,j}(t)^{\mathrm{T}} = \frac{\partial f_i}{\partial x}(x_{i,j}(T), T)\frac{\mathrm{d}x_{i,j}(T)}{\mathrm{d}x_{i,j}(t)},$$

where we define $(\mathrm{d}x_{i,j}(T)/\mathrm{d}x_{i,j})(t)$ to be the linearized dependence of the final state $x_{i,j}(T)$ on the intermediate state $x_{i,j}(t)$ at time t.

4. Batch distillation example

In order to illustrate the practical applicability of the proposed technique we consider a batch distillation process. Here the objective is to separate an ideal mixture of two components and to produce distillate with a purity of at least 99%. The optimization of the batch shall minimize time and maximize the produced quantity of distillate.

4.1. Process model

The batch distillation, that is sketched in Fig. 4, is modelled as follows: the reboiler content M_0 with composition x_0 (molar percentage of the lighter component) is heated, such that vapor with equilibrium composition $y(x_0)$ is produced, with a constant molar flux $V = 100$ kmol/h. The vapor flux is totally condensed at the top, and a liquid molar flux $L = R/(1+R)V$ is fed back into the column, where the reflux ratio R is controlled. The remainder of the condensed liquid flux, $V - L$ with composition x_{C}, is collected within the distillate container with molar holdup M_{D} and composition x_{D}. The model is based on the assumption of constant relative volatility α such that a simple expression for the vapor equilibrium in the form $y(x) := x\alpha/(x(\alpha-1)+1)$ holds, and on the assumption of constant molar overflow, i.e., the vapor flux V and the liquid flux L are constant throughout

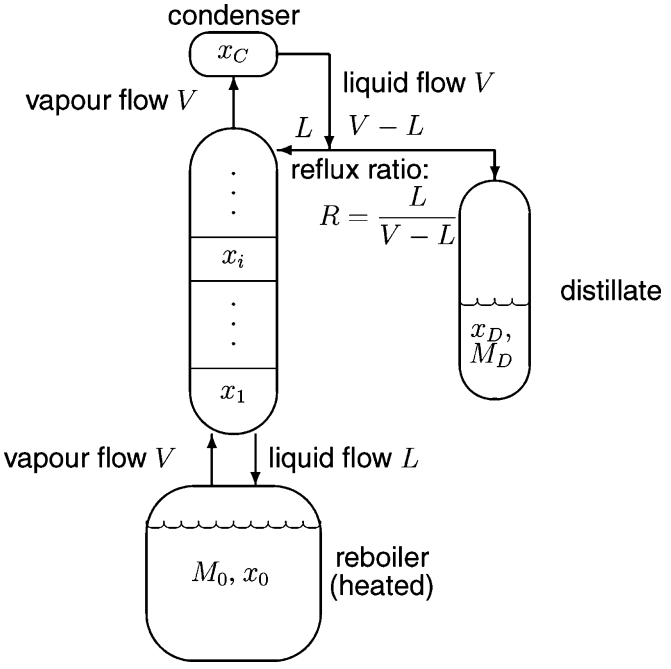


Fig. 4. Sketch of the batch distillation column.

all trays of the column. We then collect the following ordinary differential equations:

$$\dot{M}_0 = -V + L, \qquad \dot{x}_0 = \frac{1}{M_0}(Lx_1 - Vy(x_0) + (V-L)x_0) \tag{15}$$

for the reboiler, and

$$\dot{x}_i = \frac{1}{m}(Lx_{i+1} - Vy(x_i) + Vy(x_{i-1}) - Lx_i), \quad i = 1, \ldots, N, \tag{16}$$

for the tray concentrations (where $m = 0.1$ kmol is the molar holdup of each tray assumed to be constant). For the condenser concentration x_{C}, that we identify with x_{N+1}, we obtain

$$\dot{x}_{N+1} = \frac{V}{m_{\mathrm{C}}}(y(x_N) - x_{N+1}), \tag{17}$$

where $m_{\mathrm{C}} = 0.1$ kmol is the constant molar holdup of the condenser. For the distillate container, we finally find

$$\dot{M}_{\mathrm{D}} = V - L, \qquad \dot{x}_{\mathrm{D}} = \frac{V-L}{M_{\mathrm{D}}}(x_{N+1} - x_{\mathrm{D}}). \tag{18}$$

As we want to include the uncertainty in α in formally the same way as described in Section 3, we add the trivial equation $\dot{\alpha} = 0$. We choose a column with $N = 5$ trays. We stress that our distillation model with the assumptions of constant molar overflow and constant relative volatility does not represent the state of the art in distillation modeling. Especially for the separation of complex, real mixtures much more appropriate modeling techniques exist, see, e.g., the papers of Seppala and Luus (1972) and Skogestad (1997) or textbooks like Stichlmair and Fair (1998). Nevertheless, the results of the case study illustrate the relevant features of the robust dynamic optimization approach presented above.

4.2. Uncertain optimization problem

Summarizing the states in $x = (M_0, x_0, x_1, \ldots, x_{N+1}, M_{\mathrm{D}}, x_{\mathrm{D}}, \alpha)^{\mathrm{T}}$ and setting $u = R$, we obtain equations $\dot{x} = G(x,u)$ in the form given in Section 3. The objective is given by $f_0(x(T), T) := T - M_{\mathrm{D}}(T)$, and the only inequality constraint is an endpoint constraint on the distillate quality given by

$$f_1(x(T), T) := 0.99 - x_{\mathrm{D}}(T) \le 0. \tag{19}$$

Nominal initial values are $M_0 = 100$ kmol, $x_0 = 0.5$, $x_1 = \ldots = x_{N+1} = x_{\mathrm{D}} = 1$, $M_{\mathrm{D}} = 0.1$ kmol, and $\alpha = 6.0$. Uncertainty is present in the initial feed composition $x_0(0)$ as well as in the parameter α, and with $p = (x_0(0), \alpha)^{\mathrm{T}}$ we assume

$$h(p) := \left(\frac{x_0(0) - 0.5}{0.05}\right)^2 + \left(\frac{\alpha - 6.0}{5/3}\right)^2 - 1 \le 0 \tag{20}$$

This corresponds to independent relative errors of a size of 10% for $x_0(0)$ and 27.8% for α. In addition to the setup presented above, there are bounds $0 \le R \le 15$ on the controlled reflux ratio $u = R$. Note that both time T and molar holdup of the distillate container M_{D} used in the formulation of the objective

Author's personal copy

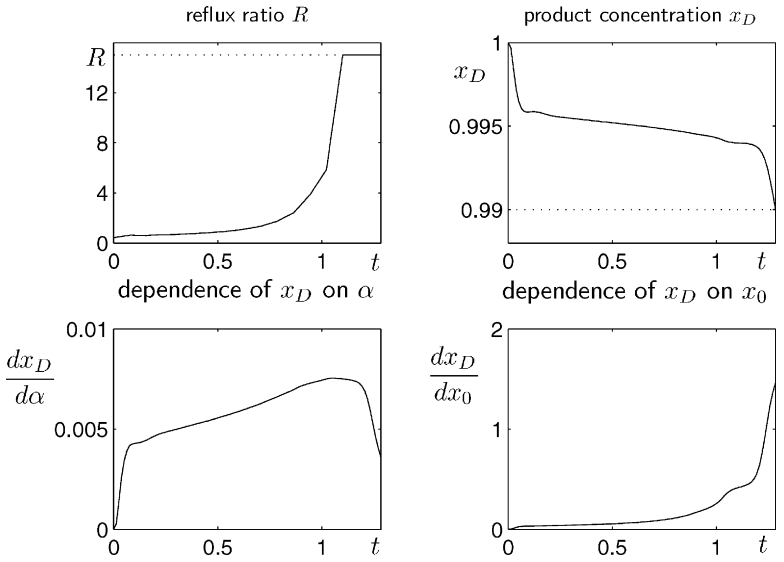


Fig. 5. The nominal (nonrobust) solution of the batch distillation problem with the profile of the reflux ratio R (top left), distillation concentration x_D (top right), and the sensitivities of x_D with respect to α (bottom left) and initial feed composition $x_0(0)$ (bottom right).

do not depend on the uncertain parameters $x_0(0)$, α. In our case study the constraints for f_0 can therefore be omitted, such that the index $i = 1$ in Eqs. (13c)–(13i) or Eqs. (14c)–(14i), respectively.

The same problem has been addressed in a recently presented work (Diehl et al., 2006). In that work robustness was approached by linear approximation of the constraints around the nominal point of operation $x_0(0) = 0.5$, $\alpha = 6.0$.

4.3. Optimization results

For the solution of the boundary value optimal control problems we employ the software package MUSCOD-II (Leineweber, Bauer, Bock, & Schlöder, 2003; Leineweber, Schäfer, Bock, & Schlöder, 2003) that is based on the direct multiple shooting method (Bock & Plitt, 1984). This method is particularly suitable for nonlinear optimal control problems with boundary and state constraints. The control profile is approximated by continuous piecewise linear interpolation using 15 equally spaced intervals on the variable time horizon $[0, T]$.

Since the ODE is highly stiff, the adjoint boundary value problem formulation (14) leads to strong instability in the adjoint equation (14g), and the problem cannot be treated by standard implementations of shooting techniques. Therefore, we employ the direct formulation (13), that carries along with each state trajectory the corresponding sensitivities with respect to the two parameters. Note, however, that the direct formulation (13) requires to integrate $n_c n_p n_x$ directional derivative states, instead of $n_c n_x$ for the adjoint approach (14).

4.3.1. Nominal solution

First, the nominal solution is computed and shown in Fig. 5. The employed time is $T = 1.28$ h while the produced distillate is $M_D(T) = 51.01$ kmol, leading to an objective of -49.73. In the plot of the distillate concentration (Fig. 5, top right) it can be seen that the lower bound of 99% on $x_D(T)$ is active. The plots in the lower row of Fig. 5 show the sensitivities of $x_D(t)$ with respect to α and $x_0(0)$, respectively, strongly indicating that the constraint may be violated due to parametric uncertainty.

This is confirmed by Fig. 6 showing the nominal point $x_0(0) = 0.5$, $\alpha = 6.0$ in the center of the uncertainty ellipse $\mathbb{P}(u)$ and the critical manifold $M^{c,0}$ consisting of the points in the $(x_0(0), \alpha)$-plane satisfying the endpoint constraint $x_D(T) = 0.99$ for the optimal control profile. For parameters on the right side of the critical boundary the constraint (19) is satisfied, whereas on the left side the constraint is violated. This critical boundary is evaluated by numerical parameter continuation with a predictor corrector algorithm (cf. Allgower & Georg, 2003). Because of the active constraint (19) the nominal point is located on the critical boundary. Fig. 6 shows that a large part of the uncertainty ellipse $\mathbb{P}(u)$ is located in the area where the constraint is violated.

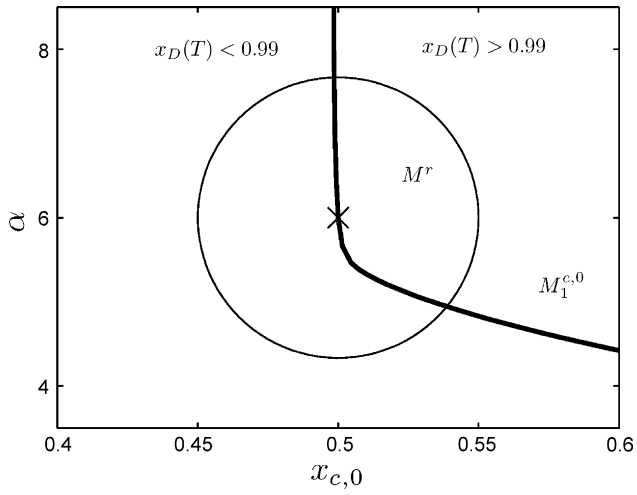


Fig. 6. Nominal point $x_0(0) = 0.5$, $\alpha = 6.0$ ($\times$) with boundary of the uncertainty set M^r and critical manifold $M_1^{c,0}$. Nominal point is located on the critical manifold.

Author's personal copy

4.3.2. Robust optimization

Following the necessary steps towards the robust optimum introduced with Algorithm 1 in Section 2.5, we first solve the inner maximization problem (WC) to initialize the NLP (13) for $J_1 = 1$. To find the worst-case point on the boundary M^r of the uncertainty ellipse $\mathbb{P}(u)$ for the nominal solution we solve the maximization problem (WC) with the fixed optimal control profile of R shown in Fig. 5. The distillate concentration $x_D(T) = 0.895$ of the worst case-point ($x_0 = 0.45$, $\alpha = 5.96266$) violates the end point constraint 19 by approximately 10%. Proceeding with the second step of Algorithm 1 we now introduce a single worst-case trajectory, instead of tracking the nominal case trajectory, i.e., we solve (13) with $J_1 = 1$. The number of ODE constraints is three times the state dimension, resulting from the state x itself and the two sensitivities with respect to $x_0(0)$ and α. Note that for illustration purposes these sensitivities have already been computed for the nominal solution in Fig. 5.

The solution obtained with $J_1 = 1$ is shown in Fig. 7. Tracking the worst-case point is achieved at the cost of an increased objective of -44.65, with slightly shorter distillation time $T = 1.27$ h but considerably less produced distillate $M_D(T) = 45.92$ kmol. Note that the worst-case parameter value is at $p_{1,1} = (x_0(0), \alpha)^T = (0.4502, 5.843)^T$. It is easily confirmed that the condition (13f), repeated here for convenience

$$D_{1,1}(T)^T \nabla_x f_1(x_{1,1}(T), T) = \nu_{1,1} \nabla_p h(p_{1,1}),$$

is satisfied, as the vector

$$\nabla_p h(p_{1,1}) = 2 \begin{bmatrix} \dfrac{0.4502 - 0.5}{(0.05)^2} \\ \dfrac{5.843 - 6.0}{(5/3)^2} \end{bmatrix} = \begin{bmatrix} -39.84 \\ -0.113 \end{bmatrix}$$

points into the same direction as the vector

$$D_{1,1}(T)^T \nabla_x f(x_{1,1}(T), T) = \begin{bmatrix} \dfrac{dx_D(T)}{dx_0(0)} \\ \dfrac{dx_D(T)}{d\alpha} \end{bmatrix} (-1) = -\begin{bmatrix} 1.537 \\ 4.37 \times 10^{-3} \end{bmatrix}$$

that consists of the final points of the two lower plots in Fig. 7. The profiles of the worst-case point in Fig. 7 differ only slightly from those of the nominal optimal point in Fig. 5. Before the maximum reflux ratio $R = 15$ is reached the reflux ratio of the robust solution is slightly higher than the reflux ratio of the nominal solution.

The third step of our robust optimization approach is to check the uncertainty set $\mathbb{P}(u)$ for points that violate the constraint (19). The boundary M^r of the uncertainty ellipse $\mathbb{P}(u)$ together with the the worst-case point and the critical manifold $M_1^{c,0}$ in the plane of the uncertain parameters $(x_0(0), \alpha)$ is shown in Fig. 8. The shape of the critical manifold is similar to the nominal case in Fig. 5. Roughly speaking the changes in the control profile shift the critical manifold to the left in the plane of the uncertain parameters. For the worst-case parameter value the end point constraint (19) is exactly fulfilled. The boundary touches the robustness ellipse tangentially at the critical point, as the conditions of optimality of the inner maximization problem are satisfied.

Fig. 8, however, shows that there are still infeasible points within $\mathbb{P}(u)$. The critical manifold has a strong curvature to the right below the worst-case point and crosses the uncertainty set. Obviously it is not sufficient to track a single worst-case point in our case study. Therefore, we have to add another critical point, i.e ., $J_1 = 2$, to the robust optimization problem (13) and return to the initialization step for the second critical point.

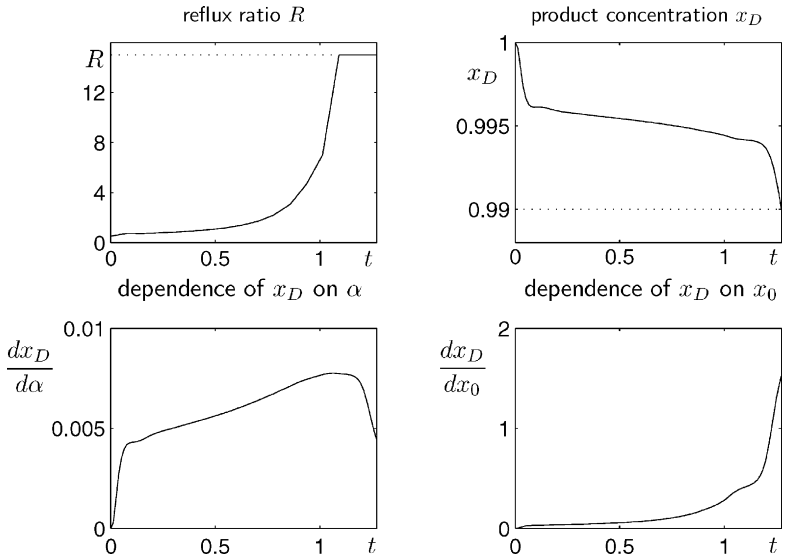


Fig. 7. Solution of the batch distillation problem for tracking one critical point, $J_1 = 1$. The plots are in the same order as in Fig. 5. The worst-case parameters are $x_0(0) = 0.4502$ and $\alpha = 5.843$.

Author's personal copy

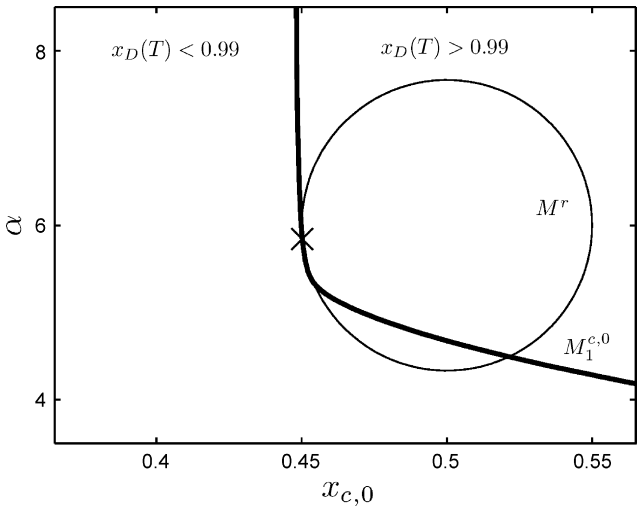


Fig. 8. Boundary of uncertainty ellipse M^r, critical point $p_{1,1}$ (×), and critical manifold $M_1^{c,0}$ for solution of the batch distillation problem for tracking single critical point $J_1 = 1$. Because of the strong curvature the critical manifold crosses the ellipse.

Initialization of the additional worst-case point is again realized by solving the inner maximization problem (WC), with the fixed control profile obtained by the optimization with a single worst-case point shown in Fig. 7. The maximum constraint violation is found at the point ($x_0(0) = 0.484$, $\alpha = 4.419$). At this point the distillate concentration $x_D(T) = 0.982$ is slightly lower than the concentration constraint (19). Compared to the nominal solution the severity of the maximum constraint violation is significantly reduced.

After initialization of the second worst-case point we proceed to the second step and solve NLP (13) for $J_1 = 2$. For tracking two worst-case points, the model equations along with the sensitivity equations have to be integrated twice. Compared to the problem of tracking a single critical point the number of states has therefore doubled to $6n_x$. The two worst-case points are located at $p_{1,1} : (x_0(0), \alpha)^T = (0.45, 5.909)^T$ and $p_{1,2} : (x_0(0), \alpha)^T = (0.484, 4.418)^T$. The quality constraint (19) is active at both points. Compared with the solution obtained by tracking a single critical point, the tracking of two worst-case points results in a longer distillation time $T = 1.36$ h. The amount of distillate $M_D(T) = 45.96$ kmol, however, has hardly changed, and the value of the objective −44.6 is only slightly higher.

The profiles for the first and second worst-case points $p_{1,1}$ and $p_{1,2}$ are shown in Figs. 9 and 10, respectively. Because of the longer distillation time T, the maximum of the reflux ratio $R = 15$ is reached later compared to the control profile of the solution with a single worst-case point. The profiles of the distillate concentration and sensitivities corresponding to the first worst-case point $p_{1,1}$ in Fig. 9 resemble those shown in Fig. 7. This is not surprising as the values of $x_0(0)$ and α are similar in both optimization results. Profiles corresponding to point $p_{1,2}$ in Fig. 10, however, show qualitative differences for the distillate concentration and the sensitivities, as the values for $x_0(0)$ and α of $p_{1,2}$ differ from the values of point $p_{1,1}$.

Again, one can easily check that condition (13f) is fulfilled for both critical points, $p_{1,1}$ and $p_{1,2}$, with the vectors of the robustness manifold M^r,

$$\nabla_p h(p_{1,1}) = \begin{bmatrix} -40.0 \\ -0.0655 \end{bmatrix}, \qquad \nabla_p h(p_{1,2}) = \begin{bmatrix} -12.8 \\ -1.139 \end{bmatrix},$$

having the same direction as the vectors

$$D_{1,1}(T)^T \nabla_x f_1(x_{1,1}(T), T) = \begin{bmatrix} 1.718 \\ 2.816 \times 10^{-3} \end{bmatrix},$$

$$D_{1,2}(T)^T \nabla_x f_1(x_{1,2}(T), T) = \begin{bmatrix} 0.172 \\ 0.0155 \end{bmatrix},$$

which can be inferred from the curves of the sensitivities shown in Figs. 9 and 10.

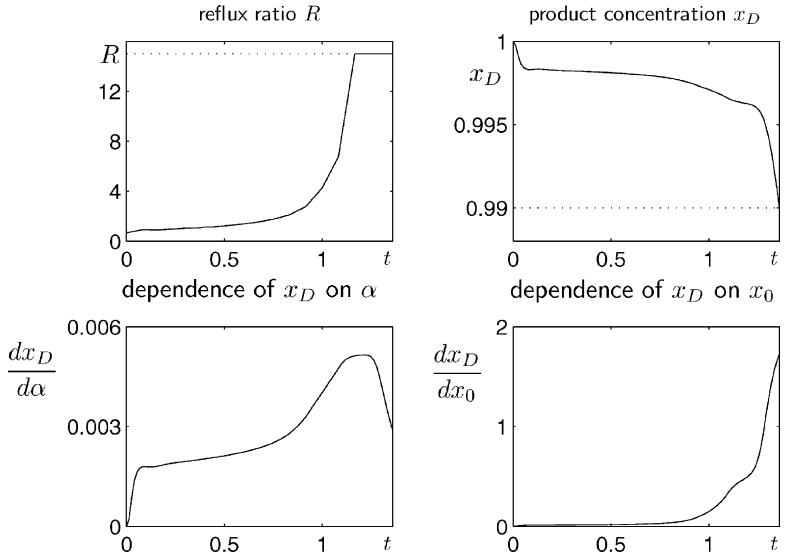


Fig. 9. Robust solution with two tracked worst cases. Trajectories for $p_{1,1}$ at $x_0(0) = 0.45$, $\alpha = 5.909$. The plots are in the same order as in Fig. 5.

Author's personal copy

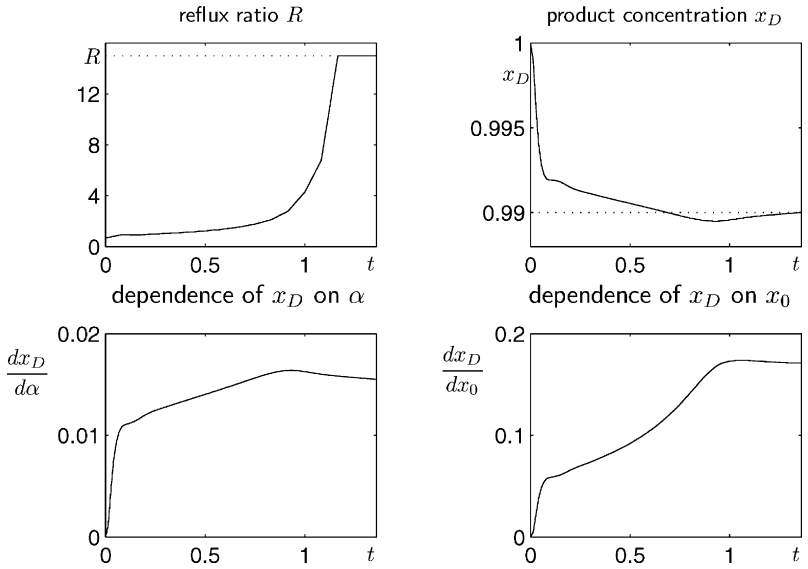


Fig. 10. Robust solution with two tracked worst cases. Trajectories for $p_{1,2}$ at $x_0(0) = 0.484$, $\alpha = 4.418$. The plots are in the same order as in Fig. 5.

To ensure that the obtained solution is robust against loss of feasibility in the presence of parametric uncertainty, we must verify, that there is no point within the uncertainty set $\mathbb{P}(u)$ that violates the quality constraint (19). Fig. 11 shows the critical manifold $M_1^{c,0}$ together with the uncertainty ellipse and the two worst-case points. The critical manifold passes through the two worst-case points as the constraint (19) is active at both points. The manifold touches the uncertainty ellipse tangentially at both critical points but does not cross the uncertainty ellipse $\mathbb{P}(u)$. Compared with Fig. 8 the critical manifold is shifted downwards. By tracking two critical points the critical manifold is pushed completely outside the uncertainty ellipse. For the distillation column there does not exist a second, insular type of critical manifold for the constraint (19). Therefore, it can be guaranteed that there are no points within $\mathbb{P}(u)$ where the constraint on the distillate concentration is violated. The optimization procedure described in Section 2.5 is terminated and a robust optimal control profile of the reflux ratio found.

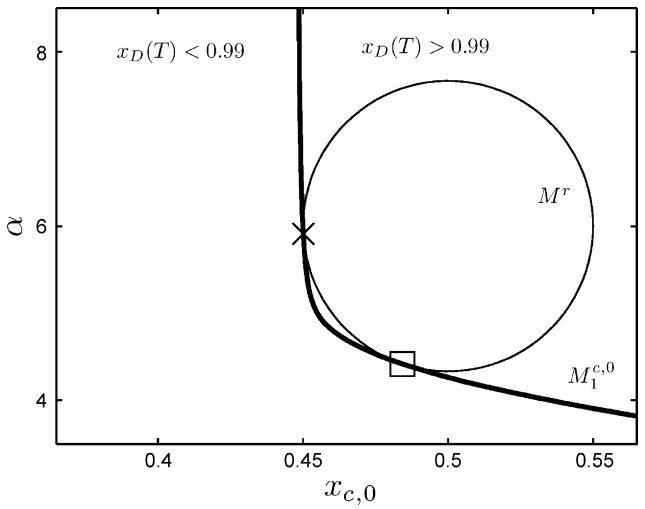


Fig. 11. Boundary of uncertainty ellipse $M^r(u)$ with two critical points $p_{1,1}$ (×) and $p_{1,2}$ (□) and critical manifold $M_1^{c,0}$ for solution of the batch distillation problem for tracking two critical points $J_1 = 2$. The constraint is fulfilled for all values within $\mathbb{P}(u)$ as the critical manifold does not cross the boundary $M^r(u)$.

4.3.3. Comparison to normal vector approach

Obviously the optimal solution corresponds to the situation sketched in Fig. 2 with the boundary M_1^c of the worst case corresponding to the zero level boundary $M_1^{c,0}$. Optimization with normal vector constraints according to NLP (9) leads therefore to the same solution with identical critical points as the presented approach. For the normal vector approach two critical points on $M_1^{c,0}$ and two points on the robustness manifold $M^r(u)$ have to be tracked that coincide at the robust optimum. As stated above, the main difference of both approaches is the initialization step. In our approach the initial values of the worst-case point is found by solving the inner maximization problem (WC). The NLP (9) of the normal vector approach is initialized by minimizing the distance between the uncertainty boundary M^r and the critical manifold $M_1^{c,0}$ (Mönnigmann & Marquardt, 2005). For the normal vector approach therefore at least one point on $M_1^{c,0}$ must be known to proceed with the initialization step. After initialization the computational costs are equivalent for both approaches as both NLPs, which have to be solved, have a very similar structure.

4.3.4. Comparison to linearization based approximation of robust optimization

Finally we want to compare the results achieved by our new approach with the results obtained with linearization of the constraint f_1 and the objective presented by Diehl et al. (2006) for the same batch distillation example.

The value of the objective function of −48.8 obtained by applying the linearization approach is only slightly smaller than the nominal solution. The distillation time $T = 1.36$ h is the same as in our approach, the produced distillate $M_D(T) = 50.2$ kmol, though, is significantly larger. Analysis of the results

Author's personal copy

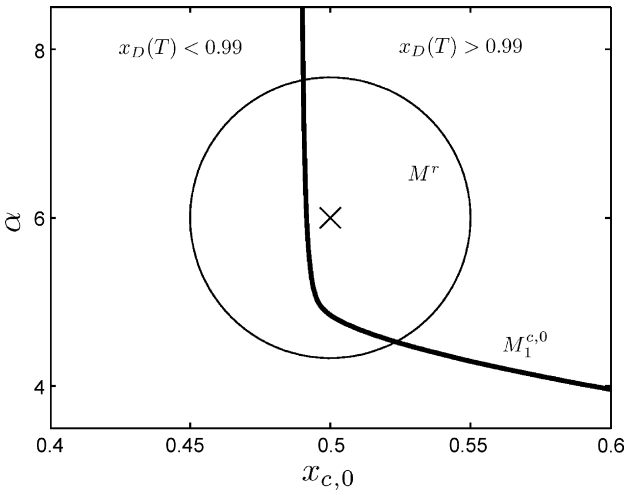


Fig. 12. Critical boundary $M_1^{c,0}$, nominal point (×) and boundary of the uncertainty ellipse $M^r(u)$ for approximated robust solution presented by Diehl et al. (2006). In large part of the uncertainty ellipse $\mathbb{P}(u)$ the endpoint constraint is violated.

obtained with the linearization approach shows, however, that the simplified NLP is only a rough approximation of the original NLP. The endpoint constraint is violated in nearly half of the area within the uncertainty ellipse $\mathbb{P}(u)$, as shown in Fig. 12. The gap between the nominal point and the critical boundary is much smaller than the specified uncertainties (20), especially for the initial value of the feed composition $x_0(0)$. The maximum violation of the quality constraint (19) on the boundary M^r of the uncertainty ellipse is at the point ($x_0(0) = 0.45, \alpha = 5.969$) with a distillate concentration of $x_D(T) = 0.91$. Obviously the batch distillation example exhibits strong nonlinearities that are neglected by the simplifying approach. The approximation by linearization underestimates the required back off of the robust solution.

This comparison shows that the NLP obtained with the linearization method may lead to too optimistic results. One has to keep in mind, however, that the linearization approach is simpler to handle than the rigorous approach presented in this work. It does not involve an iterative algorithm as described in Section 2.5 and does not require to track at least one worst-case point for every constraint.

5. Summary and conclusions

We have presented a technique for formulation of robust counterparts of nonlinear optimization problems. We assume that the optimizer chooses the controls u first, before an adverse player chooses the uncertain parameters p and states x, restricted by a bound $p \in \mathbb{P}(u)$ on disturbances in p and by the model equations $g(x, u, p) = 0$.

We adopt the well known local reduction approach for the solution of GSIP by assuming that the worst case is always located at the boundary $M^r(u)$ of the uncertainty set $\mathbb{P}(u)$. This assumption facilitates the solution of the resulting NLP by avoiding difficult to handle complementarity constraints. The approach formulates an equilibrium constraint employing first order derivatives of both the uncertainty set and the user defined objective and constraints. We showed that for active constraints $f_i = 0$ the formulation is equivalent to the normal vector based robust optimization method developed by Mönnigmann and Marquardt (2002, 2003, 2005). It has been pointed out that the main difference between both approaches is the initialization step of the robust optimization algorithm. It is important to note, however, that the strength of the normal vector approach is its versatility to treat both feasibility and stability constraints. With the present approach we have so far only considered robust optimization with feasibility constraints. Further investigations are necessary to check if the new approach can be extended to other problem formulations as well.

Two equivalent formulations for the numerical solution of the approach are proposed: they have similarity to the forward mode, respectively, to the backward mode of automatic differentiation, by computing direct, respectively, adjoint sensitivities. The adjoint formulation is in general preferable, in particular in the presence of many uncertain parameters. Both approaches keep track of at least $n_c n_x$ worst-case state solutions, but the direct approach requires $n_c n_x n_p$ derivative variables while the adjoint needs only $n_c n_x$ derivative variables. This means the problem size of the adjoint approach is independent of the number of uncertain parameters. However, the numerically stable solution of the adjoint equations may need specially tailored numerical techniques within the optimization procedure.

The method has been generalized to an optimal control problem and successfully applied for a case study with a batch distillation process. By comparison of our results to results obtained with the linearization approach of Diehl et al. (2006) we showed that the present method takes the nonlinear behavior of the model rigorously into account. The present approach therefore should be preferred in case of strongly nonlinear models and large parametric uncertainty as the linearization might give too optimistic results for such problems.

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft (DFG) within the project BO864/10-1 “Optimization Based Control of Chemical Processes” and within the project MA1188/22-1 is gratefully acknowledged. The first author gratefully acknowledges support by the Research Council of K.U. Leuven funding the Optimization in Engineering Center OPTEC (CoE EF/05/006).

References

Allgower, E., & Georg, K. (2003). *Introduction to numerical continuation methods*. Philadelphia: SIAM.

Bahri, P. A., Bandoni, J. A., & Romagnoli, J. A. (1996). Effects of disturbances in optimizing control: Steady-state open-loop backoff problem. *AIChE Journal*, *42*, 983–994.

Ben-Tal, A., & Nemirovskii, A. (2001). *Lectures on modern convex optimization: Analysis, algorithms, and engineering applications. MPS-SIAM Series on optimization*. Philadelphia: MPS-SIAM.

Author's personal copy

Biegler, L. (2000). Efficient solution of dynamic optimization and NMPC problems. In F. Allgöwer & A. Zheng (Eds.), *Nonlinear predictive control, Vol. 26 of progress in systems theory* (pp. 219–244). Basel: Birkhäuser.

Binder, T., Blank, L., Bock, H., Bulirsch, R., Dahmen, W., & Diehl, M. (2001). Introduction to model based optimization of chemical processes on moving horizons. In M. Grötschel, S. Krumke, & J. Rambau (Eds.), *Online optimization of large scale systems* (pp. 295–340). Springer.

Bock, H., & Plitt, K. (1984). A multiple shooting algorithm for direct solution of optimal control problems. In *Proc. 9th IFAC World Congress Budapest* (pp. 243–247).

Diehl, M., Bock, H., & Schlöder, J. (2005). A real-time iteration scheme for nonlinear optimization in optimal feedback control. *SIAM Journal of Control Optimization*, *43*(5), 1714–1736.

Diehl, M., Bock, H., Schlöder, J., Findeisen, R., Nagy, Z., & Allgöwer, F. (2002). Realtime optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control*, *12*(4), 577–585.

Diehl, M., Bock, H. G., & Kostina, E. (2006). An approximation technique for robust nonlinear optimization. *Mathematical Programming*, *107*, 213–230.

Findeisen, R., Allgöwer, F., & Biegler, L. (Eds.). (2006). *Assessment and future directions of nonlinear model predictive control. Lecture notes in control and information sciences*. Springer.

Halemane, K., & Grossmann, I. (1983). Optimal design under uncertainty. *AIChE Journal*, *29*, 425–433.

Hettich, R., & Kortanek, K. O. (1993). Semi-infinite programming theory, methods, and applications. *SIAM Review*, *35*(3), 380–429.

Leineweber, D., Bauer, I., Bock, H., & Schlöder, J. (2003). An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimisation. Part I: Theoretical aspects. *Computers and Chemical Engineering*, *27*, 157–166.

Leineweber, D., Schläfer, A., Bock, H., & Schlöder, J. (2003). An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimisation. Part II: Software aspects and applications. *Computers and Chemical Engineering*, *27*, 167–174.

Ma, D. L., & Braatz, R. D. (2001). Worst-case analysis of finite-time control policies. *IEEE Transactions On Control Systems Technology*, *9*(5), 766–774.

Mohideen, M., Perkins, J., & Pistikopoulos, E. (1996). Optimal design of dynamic systems under uncertainty. *AIChE Journal*, *42*(8), 2251–2272.

Mönnigmann, M., & Marquardt, W. (2002). Normal vectors on manifolds of critical points for parametric robustness of equilibrium solutions of ODE systems. *Journal of Nonlinear Science*, *12*, 85–112.

Mönnigmann, M., & Marquardt, W. (2003). Steady state process optimization with guaranteed robust stability and robust feasibility. *AIChE Journal*, *49*, 3110–3126.

Mönnigmann, M., & Marquardt, W. (2005). Steady state process optimization with guaranteed robust stability and flexibility: Application to HDA reaction section. *Industrial and Engineering Chemistry Research*, *44*, 2737–2753.

Mönnigmann, M., Marquardt, W., Bischof, C. H., Beelitz, T., Lang, B., & Willems, P. (2007). A hybrid approach for efficient robust design of dynamic systems. *SIAM Review*, *49*, 236–254.

Morari, M. (1983). Flexibility and resiliency of process systems. *Computers and Chemical Engineering*, *7*(4), 423–437.

Nagy, Z., & Braatz, R. (2003). Robust nonlinear model predictive control of batch processes. *AIChE Journal*, *49*(7), 1776–1786.

Reemtsen, R., & Rückmann, J.-J. (Eds.). (1998). *Semi-infinite programming*. Boston, Dordrecht, London: Kluwer Academic Publishers.

Rooney, W., & Biegler, L. (2001). Design for model parameter uncertainty using nonlinear confidence regions. *AIChE Journal*, *47*, 1794–1804.

Scheel, H., & Scholtes, S. (2000). Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, *25*, 1–22.

Seppala, R. E., & Luus, R. (1972). Evaluation and improvement of simulation procedures for multicomponent distillation. *Journal of the Franklin Institute*, *293*(5), 325–344.

Skogestad, S. (1997). Dynamics and control of distillation columns—a critical survey. *Modeling Identification and Control*, *18*(3), 177–217.

Stichlmair, J. G., & Fair, J. R. (1998). *Distillation*. Wiley-VCH.

Swaney, R., & Grossmann, I. (1985). An index for operational flexibility in chemical process design. *AIChE Journal*, *31*, 621–630.

Terwiesch, P., Agarwal, M., & Rippin, D. (1994). Batch unit optimization with imperfect modeling—a survey. *Journal of Process Control*, *4*, 238–258.

Yu, J. C., & Ishii, K. (1998). Design for robustness based on manufacturing variation patterns. *Journal of Mechanics Design*, *120*, 196–202.