2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Self-optimizing Invariants in Dynamic Optimization

Johannes Jäschke, Miroslav Fikar and Sigurd Skogestad

Abstract— **In optimal control, the input trajectories are often solved numerically or analytically. This requires that all variables which enter the optimality conditions are known or measured. We use techniques from polynomial elimination theory to eliminate variables which are not known from the optimality conditions. The result is an expression of the optimality conditions in known variables only, which can easily be evaluated and controlled by feedback.**

I. Introduction

Dynamic optimization problems are ubiquitous in science and engineering. In process control, they are found in the optimization of batch reactors or grade transitions in continuous processes. Most approaches in literature deal with optimization based on a model.

One of the oldest approaches is to find the optimal inputs using the Pontryagin minimum principle [3]. This is basically an open-loop approach and requires a simple model, where all parameters and variables are known (measured).

A second approach, which is very popular today, is nonlinear model predictive control. Here, the dynamic optimization problem is converted to a nonlinear optimization problem, and solved repeatedly at given sample times [1], [8], [14], [17]. Measurements are used to update the process model states and parameters by e.g. moving horizon estimation [15]. Although this is conceptually an open-loop approach, feedback is introduced by repeated optimization.

A third approach is to use the model off-line and exploit the solution structure to find variables, which give optimal or near-optimal operation, when kept at constant setpoints using a feedback policy. This approach is followed in NCO tracking [20] and self-optimizing control [18].

Whenever a model is used, handling uncertainty is a major challenge. Uncertainty may arise from different sources, such as incomplete information (unmeasured states), parametric disturbances, and model structure error. There are several approaches to handle the uncertainty:

1) Estimate the unknown variables using a filter or moving horizon estimation [15], as done in model predictive control. This approach is used frequently; however, it can be difficult to obtain converging estimates within reasonable time.
2) Use a robust control approach [23] or stochastic optimization approach [2]. Here we attempt to find an open-loop control policy, which gives the best performance over a range of disturbances. Generally it has to compromise performance to gain robustness.
3) Neighboring extremal control [3], where the optimization problem does not have to be re-solved completely when a disturbance occurs. Instead, corrections to the nominal input trajectory are found by solving a linear approximation to the nonlinear problem.
4) The approach presented in this paper, where we use model equations to eliminate the unknown or uncertain variables from the optimality conditions.

Our work contributes to handling parametric uncertainty for dynamic optimization problems. The main contribution is the extension of concepts from steady state self-optimizing control [18] to a class of polynomial dynamic optimization problems.

The idea is to formulate the optimality conditions ($H_u = 0$) which include unknown parameters, and then use tools from elimination theory [10], [7] for eliminating the unknown parameters to obtain optimal invariants in known (measured) variables which can be controlled using feedback. Controlling these invariants and the optimality conditions is equivalent.

In Section II we present the optimal control problem and state the optimality conditions. Section III describes how to eliminate the adjoint variables from the optimality conditions. In Section IV we introduce concepts from toric elimination theory, and apply them in Section V to eliminate unknown parameters from the optimality conditions. Section VI gives a case study of a fed batch reactor, and Section VII closes the paper with a short discussion and conclusion.

II. Optimal control

A. Problem Formulation

We consider a class of dynamic optimization problems, which can be written in following form:

$$\min_{\mathbf{u}(t)} \Phi(t_f) = J(\mathbf{x}(t_f)) \tag{1a}$$

$$\text{s.t. } \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}(t)) + \mathbf{G}(\mathbf{x}(t))\mathbf{u}(t); \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{1b}$$

$$\mathbf{u}^L \leq \mathbf{u}(t) \leq \mathbf{u}^U. \tag{1c}$$

The scalar function J denotes the terminal cost, and the functions $\mathbf{u} : [0, t_f] \to \mathbb{R}^{n_\mathbf{u}}$ and $\mathbf{x} : [0, t_f] \to \mathbb{R}^{n_\mathbf{x}}$ denote the input and state functions, respectively. $\mathbf{F}(\mathbf{x})$ is a vector valued function of dimension $n_\mathbf{x}$, and $\mathbf{G}(\mathbf{x})$ is a matrix of dimension $n_\mathbf{x} \times n_\mathbf{u}$. The elements of $\mathbf{F}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ are polynomials in the ring $\mathbb{R}[\mathbf{x}]$, that is, every row in $\mathbf{F}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ contains polynomials in the variables $\mathbf{x}$ and coefficients in $\mathbb{R}$. The

This work was partially supported by the Project NIL-I-007-d.

J. Jäschke and S. Skogestad are with the Department of Chemical Engineering, Norwegian institute of Science and Technology, (NTNU) <jaschke><skoge>@chemeng.ntnu.no

M. Fikar is with the Institute of Information Engineering, Automation and Mathematics, STU in Bratislava, Slovakia, miroslav.fikar@stuba.sk

variables $\mathbf{u}^L$ and $\mathbf{u}^U$ denote the time invariant lower and upper bounds for the inputs $\mathbf{u}$. Note that the system is input affine and we consider only input constraints. All functions are assumed to be sufficiently smooth and differentiable.

B. First order optimality conditions

Assumption 1: The optimal control problem (1) is feasible and has a unique solution $\mathbf{u}^*(t)$.
We define the Hamiltonian

$$H(\mathbf{x},\mathbf{u},\lambda,\mu^L,\mu^U) = \lambda^{\mathbf{T}}(\mathbf{F}(\mathbf{x})+\mathbf{G}(\mathbf{x})\mathbf{u}) + {\mu^L}^{\mathbf{T}}(\mathbf{u}^L-\mathbf{u}) + {\mu^U}^{\mathbf{T}}(\mathbf{u}-\mathbf{u}^U), \tag{2}$$

where λ, μ^L and μ^U are adjoint variables corresponding to the model, lower and upper input constraints, respectively.

Theorem 1 (Pontryagin Minimum Principle [3], [11]): If the control $\mathbf{u}$ is optimal, then there exist nontrivial vectors of adjoint variables λ and μ, such that the following conditions are satisfied:

1)

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \lambda}, \qquad \mathbf{x}(0)=\mathbf{x}_0 \tag{3a}$$

$$\dot{\lambda}^{\mathbf{T}} = -\frac{\partial H}{\partial \mathbf{x}}, \qquad \lambda^{\mathbf{T}}(t_f) = \frac{\partial J}{\mathbf{x}(t_f)} \tag{3b}$$

$${\mu^L}^{\mathbf{T}}(\mathbf{u}^L-\mathbf{u})=0, \qquad {\mu^U}^{\mathbf{T}}(\mathbf{u}-\mathbf{u}^U)=0 \tag{3c}$$

2) For all $t\in[t_0,t_f]$, the Hamiltonian has a global minimum with respect to $\mathbf{u}$, i.e.

$$H(\mathbf{x}^*,\mathbf{u}^*,\lambda^*,{\mu^L}^*,{\mu^U}^*) \le H(\mathbf{x}^*,\mathbf{u},\lambda^*,\mu^*,{\mu^L}^*,{\mu^U}^*) \tag{4}$$

for all $\mathbf{u}^L\le\mathbf{u}\le\mathbf{u}^U$ and $t\in[t_0,t_f]$.

3) If the final time is free, we have the transversality condition

$$H(\mathbf{x}(t_f),\mathbf{u}(t_f)^*,\lambda(t_f),\mu^L(t_f),\mu^U(t_f))=0. \tag{5}$$

III. Eliminating Adjoint Variables

The optimal solution of problem (1) consists of a sequence of arcs (regions) which are defined in certain intervals. The arcs are defined by the set of active constraints, and are continuous and differentiable within each interval [3]. We distinguish two types:

1) Constrained arcs (boundary arcs): One or more inputs are at a constraint.
2) Unconstrained arcs: The inputs are all unconstrained.

In the constrained arcs we simply keep the inputs at the active constraint. If there are unconstrained degrees of freedom left, the remaining problem can be reformulated as an unconstrained problem by redefining the input set. Therefore, in the following, we consider only the case where no constraint is active.

At the minimum of the Hamiltonian (4), we must have $H_{\mathbf{u}}=\frac{\partial H}{\partial \mathbf{u}}=0$. Considering one input at a time, the condition reads:

$$H_{\mathbf{u}_i}=\frac{\partial H}{\partial \mathbf{u}_i}=0, \quad i=1\dots n_{\mathbf{u}} \tag{6}$$

Unfortunately, we cannot control $H_{\mathbf{u}_i}$ to zero, because it generally contains unknown variables, including the adjoint variables λ. To eliminate the adjoint variables, we perform successive time differentiations.

Definition 1 (Lie bracket, [16]): Given two vector fields $f,g:\mathbb{R}^n\to\mathbb{R}^n$. The Lie bracket $[f,g]$ is the vector field defined by

$$[f,g]=\frac{\partial g}{\partial x}f-\frac{\partial f}{\partial x}g. \tag{7}$$

Recursive bracketing is defined as $ad_f^k g=[f,ad_f^{k-1}g]$, with $ad_f^0 g=g$.

It can be shown [20], [13], that the k-th time derivative of $H_{\mathbf{u}_i}$ can be written as

$$H_{\mathbf{u}_i}^{(k)}=\frac{\mathrm{d}^{(k)}H_{\mathbf{u}_i}}{\mathrm{d}t^{(k)}}=\lambda^{\mathbf{T}}\left(ad^k_{\mathbf{F}(\mathbf{x})}\mathbf{G}_i(\mathbf{x})\right)=\lambda^{\mathbf{T}}A_k^i, \tag{8}$$

where $\mathbf{G}_i(\mathbf{x})$ denotes the i-th column in $\mathbf{G}(\mathbf{x})$. Since for the optimal solution $H_{\mathbf{u}_i}=0$ holds at all times, its optimal time derivatives must be zero at all times, too. We write the time derivatives up to the $n_{\mathbf{x}}-1$-th derivative as

$$\lambda^{\mathbf{T}}\left[A_0^i,A_1^i,A_2^i,\dots A_{n_{\mathbf{x}}-1}^i\right]=\lambda^{\mathbf{T}}\mathbf{A}^i=0, \tag{9}$$

where all terms A_0^i, A_1^i, A_2^i,..., $A_{n_{\mathbf{x}}-1}^i$ are collected in the matrix $\mathbf{A}^i$. Eq. (9) has a nontrivial solution for λ only if $\det(\mathbf{A}^i)=0$. Therefore, controlling

$$c_i=\det(\mathbf{A}^i) \tag{10}$$

to zero gives optimal operation. If we have several inputs, we may collect all c_i into a vector $\mathbf{c}=[c_1,\dots,c_i,\dots,c_{n_u}]^{\mathbf{T}}$. The vector $\mathbf{c}$ generally contains unknown variables, such as unmeasured states or unmeasured disturbances $\mathbf{d}$. Therefore it cannot be used for control directly.

Since the optimal control system (1) is defined in polynomial equations, and all calculations above preserve the polynomial structure, we use results from elimination theory to eliminate unknowns in each $c_i=\det(\mathbf{A}^i)$, to obtain variables suitable for control.

IV. Toric elimination theory

We give a very short introduction to toric elimination theory, for more detailed information we refer to [6], [7], [9], [12], [22]. More specifically, we present the sparse resultant from algebraic geometry [7], [10] to eliminate the unknowns. Casually speaking, the resultant is a condition for an overdetermined system of polynomials to have a common root.

We consider a system of $n+1$ polynomials,

$$f_0=\dots=f_n=0, \tag{11}$$

in n variables $\mathbf{x}=[x_1,\dots x_n]^{\mathbf{T}}$, and let $\mathbb{C}^*$ denote the complex numbers without zero, $\mathbb{C}^*=\mathbb{C}\setminus 0$. Toric elimination theory considers solutions of the polynomials (11) in $(\mathbb{C}^*)^n$. Since none of the variables is allowed to be zero, the theory is valid for Laurent polynomials in $\mathbb{R}[\mathbf{x},\mathbf{x}^{-1},\mathbf{u},\mathbf{u}^{-1}]$, that is, polynomials with positive and negative integer exponents.

Definition 2 (Monomial): We define a monomial $\mathbf{x}^a$ as the power product $\mathbf{x}^a=x_1^{a_1}x_2^{a_2}\dots x_n^{a_n}$, where $(a_1,a_2,\dots,a_n)\in\mathbb{Z}^n$.

Definition 3 (Support): Let the support $\mathscr{E}_i = \{a_{i,1},\dots,a_{i,m_i}\}$ denote the set of exponent vectors corresponding to monomials in

$$f_i = \sum_{j=1}^{m_i} c_{i,j}\mathbf{x}^{a_{i,j}}, \quad c_{i,j} \neq 0. \tag{12}$$

We denote as $Q_i = conv(\mathscr{E}_i)$ the convex hull of the support of the polynomial f_i.

Definition 4 (Affine variety): Consider $f_1,\dots,f_m$ polynomials in $\mathbb{C}[x_1,\dots x_n]$. The affine variety $V(f_1,..,f_m)$ is defined by the set

$$V(f_1,\dots,f_m) = \{(x_1,..,x_n) \in \mathbb{C}^s : f_i(x_1,..,x_n) = 0, i = 1\dots m\}. \tag{13}$$

Definition 5 (Zariski closure): Given a subset $S \subset \mathbb{C}^m$, the smallest affine variety containing S is called the Zariski closure of S and is denoted as $\bar{S}$.

Let $L(\mathscr{E}_i)$ be the set of all polynomials that have exponents in the support $\mathscr{E}_i$

$$L(\mathscr{E}_i) = \left\{c_{i,1}\mathbf{x}^{a_{i,1}} + \dots + c_{i,m_i}\mathbf{x}^{a_{i,m_i}} : \quad c_{i,j} \in \mathbb{C}^*\right\}, \tag{14}$$

Then the coefficients of a polynomial define a point in $\mathbb{C}^{m_i}$. Now let

$$Z(\mathscr{E}_0,\dots,\mathscr{E}_n) \subset L(\mathscr{E}_0) \times \dots \times L(\mathscr{E}_n) \tag{15}$$

be the Zariski closure of the set of all $(f_0,\dots f_n)$, for which (12) has a solution in $(\mathbb{C}^*)^n$. For an overdetermined system of polynomials we then have this result.

Theorem 2 (Sparse resultant [12], [7]): Assume that $Q_i = \text{conv}(\mathscr{E}_i)$ is a n dimensional polytope for $i = 0,\dots,n$. Then there is an irreducible polynomial $\mathscr{R}$ in the coefficients of the f_i such that

$$(f_0,..,f_n) \in Z(\mathscr{E}_0,\dots,\mathscr{E}_n) \Leftrightarrow \mathscr{R}(f_0,..f_n) = 0. \tag{16}$$

In particular, if the system

$$f_0 = f_1 = \dots = f_n \tag{17}$$

has a solution $(x_1,\dots,x_n) \in (\mathbb{C}^*)^n$, then

$$\mathscr{R}(f_1,\dots,f_n) = 0. \tag{18}$$

We call $\mathscr{R}$ the sparse resultant.

Remark 1: There exist more general versions of Theorem 2, which do not require the convex hull of the supports to be n-dimensional. [21]. However, for simplicity we chose to present this simplified version here.

Example 1 (One variable): Consider the system

$$\begin{aligned} f_0 &= a_{11} + a_{12}x \\ f_1 &= a_{21} + a_{22}x + a_{2,3}x^2. \end{aligned} \tag{19}$$

The supports of this system are $\mathscr{E}_0 = \{(0),(1)\}$, and $\mathscr{E}_1 = \{(0),(1),(2)\}$. Clearly, the convex hulls of the supports are the line segments $[0,1]$ and $[0,2]$, which have dimension $n = 1$. For arbitrary a_{ij} (19) does not have a solution in $\mathbb{C}^*$. The sparse resultant for this system is calculated as the determinant of the Sylvester matrix

$$\begin{aligned} \mathscr{R}(f_0,f_1) &= \det\left(\begin{bmatrix} a_{12} & a_{11} & 0 \\ 0 & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \end{bmatrix}\right) \\ &= a_{12}^2 a_{21} - a_{12}a_{11}a_{22} + a_{23}a_{11}^2. \end{aligned} \tag{20}$$

Note that we have eliminated x from (19), and the statement $\mathscr{R}(f_0,f_1) = 0$ is identical to stating that there exist some x such that $f_0 = f_1 = 0$.

The calculation of the sparse resultant for multivariate polynomials is more involved. An algorithm is given in [5]. In this work, we use the software `multires` [4].

V. Using resultants in optimal control

After introducing the sparse resultant, we can apply it to our optimal control problem. We collect all unknown (unmeasured) variables in a vector $\mathbf{d}$, so we have $c_i = c_i(\mathbf{d})$, and we write the model equations in the form

$$m(\mathbf{d}) = \mathbf{0}, \tag{21}$$

where we have omitted to explicitly state the dependency on the known variables.

Assumption 2: The model equations are polynomials in the polynomial ring $\mathbb{R}[\mathbf{d}]$.

Assumption 3: The variety $V(\mathbf{m}(\mathbf{d}))$ is zero-dimensional, that is, $\mathbf{m}(\mathbf{d}) = \mathbf{0}$ has a finite number of solutions.

Theorem 3 (Invariants for Control): If the number of unknown variables $n_\mathbf{d}$ is equal to the number of model equations $n_\mathbf{m}$, and Assumptions 2 and 3 hold, controlling

$$\mathscr{R}(c_i(\mathbf{d}),\mathbf{m}(\mathbf{d})) = 0 \tag{22}$$

is equivalent to controlling (10).

Proof: By assumption, the model equations $\mathbf{m}(\mathbf{d}) = \mathbf{0}$ have a finite number of solutions. $c_i(\mathbf{d}) = \det(\mathbf{A}^i)$ is a polynomial in the variables $\mathbf{d}$ whose coefficients are functions of $\mathbf{u}$, and thus can be manipulated. Arbitrary input values $\mathbf{u}$ will cause that $c_i(\mathbf{d}) = 0$ does not have any solution. The sparse resultant $\mathscr{R}(c_i(\mathbf{d}),\mathbf{m}(\mathbf{d}))$ gives the necessary and sufficient condition for the combined system

$$\begin{aligned} \mathbf{m}(\mathbf{d}) &= \mathbf{0} \\ c_i(\mathbf{d}) &= 0 \end{aligned} \tag{23}$$

to have a solution in $(\mathbb{C}^*)^{n_\mathbf{d}}$. By Theorem 2, we have

$$c_i(\mathbf{d}) = \det(\mathbf{A}^i) = 0 \Leftrightarrow \mathscr{R}(c_i(\mathbf{d}),\mathbf{m}) = 0. \tag{24}$$

■

Controlling $\mathscr{R}(c_i(\mathbf{d}),\mathbf{m}_i) = 0$ is equivalent to controlling the optimality conditions $c_i(\mathbf{d}) = 0$, as long as the model is satisfied. However, whereas the $c_i(\mathbf{d})$ contains unmeasured variables, they have all been eliminated from $\mathscr{R}(c_i(\mathbf{d}),\mathbf{m}_i)$.

Remark 2: Note that it is not necessary to be able to solve the model equations $\mathbf{m}(\mathbf{d}) = \mathbf{0}$ uniquely for $\mathbf{d}$. The only condition is that the model equations have a finite number of solutions.

Remark 3: Since the unknown variables $\mathbf{d}$ assume real values in the process, the existence of complex solutions for

TABLE I

PARAMETERS AND INITIAL CONDITIONS

Symbol	Value	Unit	Description
k_1	0.053	l/mol/min	parameter
k_2	0.128	l/mol/min	”
c_B^{in}	5	mol/l	”
t_f	250	min	”
u_{min}	0	l/min	input constraint
u_{max}	0.001	l/min	”
c_{A0}	0.72	mol/l	initial condition
c_{B0}	0.0614	mol/l	”
c_{C0}	0.0	mol/l	”
c_{D0}	0.0	mol/l	”
V_0	1	l	”

$\mathbf{m}(\mathbf{d}) = \mathbf{0}$ does not matter, because the Theorem 2 states that c_i becomes zero whenever the resultant is zero.

VI. CASE STUDY: FED BATCH REACTOR

A. Model

The case study is taken from [13]. We consider a fed batch reactor with two reactions,

$$A + B \longrightarrow C \text{ and } 2B \longrightarrow D, \tag{25}$$

where C is the desired product and D is the undesired side product. The objective is to maximize difference between the amount of C and the amount of D at the final batch time t_f. We use a simple dynamic model,

$$\begin{aligned} \dot{c}_A &= -k_1 c_A c_B - c_A u/V \\ \dot{c}_B &= -k_1 c_A c_B - 2k_2 c_B^2 - (c_B - c_B^{in})u/V \\ \dot{V} &= u, \end{aligned} \tag{26}$$

with the initial conditions: $c_A(0) = c_{A0}$, $c_B(0) = c_{B0}$, and $V(0) = V_0$. Initially the concentration of the products is zero, $c_{C0} = c_{D0} = 0$. All parameters and initial conditions are given in Table I. From the mass balance, we have ($c_{C0} = c_{D0} = 0$)

$$c_C(t) = \frac{1}{V}\left(c_{A0}V_0 - c_A(t)V\right) \tag{27}$$

and

$$c_D(t) = \frac{1}{2V}\left[\left(c_A + c_B^{in} - c_B\right)V - \left(c_{A0} + c_B^{in} - c_{B0}\right)V_0\right]. \tag{28}$$

B. Optimal control problem

The optimization problem is then formulated as

$$\min_{\mathbf{u}} J(t_f) \quad \text{s.t.} \quad \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}, \tag{29}$$

where the objective is

$$J(t_f) = (c_D(t_f) - c_C(t_f))V(t_f). \tag{30}$$

Further, we have the state and input vectors $\mathbf{x} = [c_A, c_B, V]^{\mathbf{T}}$ and $\mathbf{u} = u$, and

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} -k_1 c_A c_B \\ -k_1 c_A c_B - 2k_2 c_B^2 \\ 0 \end{bmatrix}, \quad \mathbf{G}(\mathbf{x}) = \frac{1}{V}\begin{bmatrix} -c_A \\ c_B^{in} - c_B \\ V \end{bmatrix}. \tag{31}$$

The constraints for the system are $u_{min} \le u \le u_{max}$.

TABLE II

SIMULATION CASES

	Case 1	Case 2
Unmeasured state:	c_A	c_B
Unmeasured disturbance:	none	k_1

C. Nominal optimal solution

For the given initial conditions the system is unconstrained, and the optimal trajectory consists of one interior arc. The Hamiltonian is

$$\begin{aligned} H = {} & \lambda_1\left(-k_1 c_A c_B - c_A u/V\right) \\ & + \lambda_2\left(-k_1 c_A c_B - 2k_2 c_B^2 + (c_B^{in} - c_B)u/V\right) \\ & + \lambda_3 u. \end{aligned} \tag{32}$$

Proceeding as in Section II we get $H_u = \boldsymbol{\lambda}^{\mathbf{T}} A_0 = 0$ with

$$A_0 = \begin{bmatrix} -c_A/V & (c_B^{in} - c_B)/V & 1 \end{bmatrix}^{\mathbf{T}}. \tag{33}$$

We continue with the first and second time derivatives $\boldsymbol{\lambda}^{\mathbf{T}} A_1 = 0$ and $\boldsymbol{\lambda}^{\mathbf{T}} A_2 = 0$. Here, $A_1 = [a_{11}\, a_{12}, a_{13}]^{\mathbf{T}}$ with

$$\begin{aligned} a_{11} &= [-k_1 c_A (c_B - c_B^{in})]/V \\ a_{12} &= [-k_1 c_A (c_B - c_B^{in}) - 2k_2 c_B (c_B - 2c_B^{in})]/V \\ a_{13} &= 0, \end{aligned} \tag{34}$$

and $A_2 = [a_{21}, a_{22}, a_{23}]^{\mathbf{T}}$, with

$$\begin{aligned} a_{21} &= \left[c_B^{in} k_1 c_A V (k_1 c_A + 4k_2 c_B) + 2k_1 c_A \left(c_B - c_B^{in}\right) u\right]/V^2 \\ a_{22} &= \left[c_B^{in} V \left(4k_1 c_A k_2 c_B + 8k_2^2 c_B^2 + k_1^2 c_A^2\right)\right. \\ &\quad \left. + 2\left(c_B - c_B^{in}\right)\left(k_1 c_A + 2k_2 (c_B - c_B^{in})u\right)\right]/V^2 \\ a_{23} &= 0. \end{aligned} \tag{35}$$

$c_i = \det(\mathbf{A})$ becomes zero when $c = 0$, with

$$c = 4k_2 c_B^2 c_B^{in} V + 2c_B c_B^{in} u - k_1 c_A c_B^2 V + 2k_1 c_A c_B^{in} V c_B - 2u(c_B^{in})^2. \tag{36}$$

In optimal control literature, e.g. [3], this expression is commonly solved for u, and implemented in the process. However, this is not always possible, because c generally contains unmeasured states and disturbances.

D. Eliminating unknown variables

We consider two different scenarios as summarized in Table II.

1) Case 1: Unknown variables in algebraic equations: Assume that the concentration c_A is difficult or expensive to measure. Then we have one unmeasured state, c_A. All other variables in c from (36) are known. However, the unmeasured state is present in the algebraic relationship (27). This gives the measurement polynomial

$$m_1 = V c_C(t) - (c_{A0}V_0 - c_A(t)V) = 0, \tag{37}$$

and we calculate the resultant

$$\begin{aligned} \mathscr{R}(c, m_1) = {} & -V c_B^2 k_1 c_C + 2V c_B k_1 c_B^{in} c_C - 4V c_B^2 k2 c_B^{in} - 2c_B^{in} u c_B \\ & + 2{c_B^{in}}^2 u - 2c_B k_1 c_B^{in} c_{A0} V_0 + c_B^2 k_1 c_{A0} V_0. \end{aligned} \tag{38}$$

$\mathscr{R}(c,m_1)$ does not contain the unmeasured state, and controlling it to zero will by Theorem 3 result in optimal operation.

2) Case 2: Unknown variables in the differential equations: Now assume that we have an unknown disturbance k_1, and that the concentration c_B is unmeasured. Since the reaction rate enters through a differential equation, we need to eliminate k_1 from c (36) using a differential equation, and we need to use a change rate as a measured variable, too.

We assume that we can measure the concentration c_A together with an estimate of its time derivative, $\dot{c}_A$. If the measurement of c_A is good (little or no noise), then we may use its past values to estimate its time derivative by filtering or using finite differences,

$$\dot{c}_A = (c_A(t) - c_A(t-1\text{min}))/(1\text{min}). \tag{39}$$

This does not give the exact derivative, but the approximation is considered good enough for our purposes.

To eliminate the unknowns c_B and k_1 we use an additional mass balance for component B,

$$m_2 = -c_B V + c_{B0}V_0 + c_B^{in}(V - V_0) - c_C V - 2c_D V = 0, \tag{40}$$

together with the implicit component balance for c_A from (26),

$$m_3 = \dot{c}_A V + k_1 c_A c_B V + c_A u = 0 \tag{41}$$

and we eliminate the unknowns by calculating the resultant with respect to the unknown variables k_1 and c_B:

$$\begin{aligned}
&\mathscr{R}(c,m_2,m_3) = \\
&-16V^2c_Ak_2c_B^{in}c_D + V^2c_A\dot{c}_A + 4V^2k_2c_B^{in}{c_A}^2 + 8V^2c_Ak_2{c_B^{in}}^2 \\
&+ Vu{c_A}^2 - 16Vc_Dk_2c_B^{in}V_0c_{B0} + 16Vc_Dk_2{c_B^{in}}^2V_0 \\
&+ 8VV_0c_{B0}k_2{c_B^{in}}^2 - 8Vc_Ak_2{c_B^{in}}^2V_0 + 16Vc_{A0}V_0k_2c_B^{in}c_D \\
&+ 8Vc_Ak_2c_B^{in}V_0c_{B0} - 8Vc_Ac_{A0}V_0k_2c_B^{in} - 8Vc_{A0}V_0k_2{c_B^{in}}^2 \\
&- 8V_0^2c_{B0}k_2{c_B^{in}}^2 + 4V_0^2c_{B0}^2k_2c_B^{in} + 2V_0c_{B0}c_B^{in}u - c_AV_0c_B^{in}u \\
&+ c_AV_0c_{B0}u + 16V^2c_D^2k_2c_B^{in} - 16V^2c_Dk_2{c_B^{in}}^2 - 2Vc_Ac_Du \\
&+ Vc_Ac_B^{in}u - 4Vc_Dc_B^{in}u - 8V{c_B^{in}}^3k_2V_0 - VV_0c_B^{in}\dot{c}_A + VV_0c_{B0}\dot{c}_A \\
&- Vc_{A0}V_0\dot{c}_A - c_Ac_{A0}V_0u - 2c_{A0}V_0c_B^{in}u + 8c_{A0}V_0^2k_2{c_B^{in}}^2 \\
&+ 4c_{A0}^2V_0^2k_2c_B^{in} + 4V_0^2{c_B^{in}}^3k_2 - 2V_0{c_B^{in}}^2u + 4V^2{c_B^{in}}^3k_2 \\
&- V^2c_B^{in}\dot{c}_A - 2V^2c_D\dot{c}_A - 8c_{A0}V_0^2k_2c_B^{in}c_{B0}.
\end{aligned} \tag{42}$$

This expression does not contain any of the unknown variables, so it can be evaluated online and controlled to zero using a P or PI controller.

E. Simulation Results

1) Nominal operation: The state and input trajectories for nominal optimal operation are given in Figure 1. These trajectories are generated by applying the optimal input. The final optimal cost is value is $J = 0.2717$.

2) Controlling the invariant:

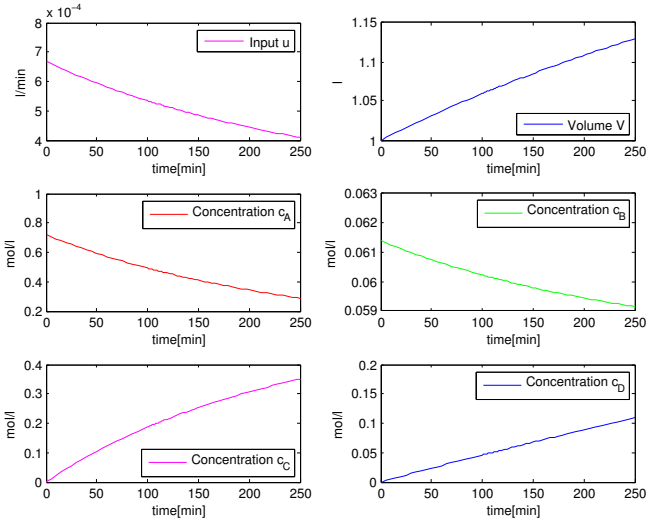

Fig. 1. Nominal optimal input, volume, and concentration trajectories

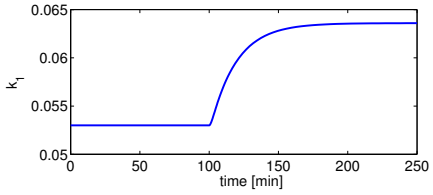

Fig. 2. Disturbance k_1

a) Case 1.: Variable c_A unmeasured - all other variables known: Here we cannot control $c = \det(A)$ to zero, because we cannot evaluate it since c_A is not known. Instead we control the resultant $\mathscr{R}(c,m_1)$ (38) to zero using a P-controller. The trajectories are identical to the optimal ones from the previous section, and the objective value is $J = 0.2717$. This is as expected, because by Theorem 3, controlling c and $\mathscr{R}(c,m_1)$ are equivalent. The suboptimality, which is introduced by the added P controller, does not become visible when considering the first seven digits of the objective function. However, whereas we need to know the value of c_A to control c, this is not necessary for controlling $\mathscr{R}(c,m_1)$ to zero.

b) Case 2.: Variables k_1, c_B unmeasured – $\dot{c}_A$ estimated and c_A measured: In this case, the state c_B and the parameter k_1 are not known (measured). Therefore we cannot control c directly. Instead we use a P-controller to control $\mathscr{R}(c,m_2,m_3)$, which contains neither k_1 nor c_B. This expression can be evaluated using the available measurements and controlled to zero. In the nominal case the trajectories look exactly the same as in Fig. 1.

Next, we consider a change in the reaction kinetics, where k_1 rises 20%, Figure 2. The input and the states are given in Figure 3. The final profit when controlling $\mathscr{R}(c,m_2,m_3)$ to zero is $J = 0.2970$, while the profit using the optimal input is $J_{opt} = 0.2971$. This difference comes from the approximation of $\dot{c}_A$ in (39). Using the exact derivative, we obtain $J =$

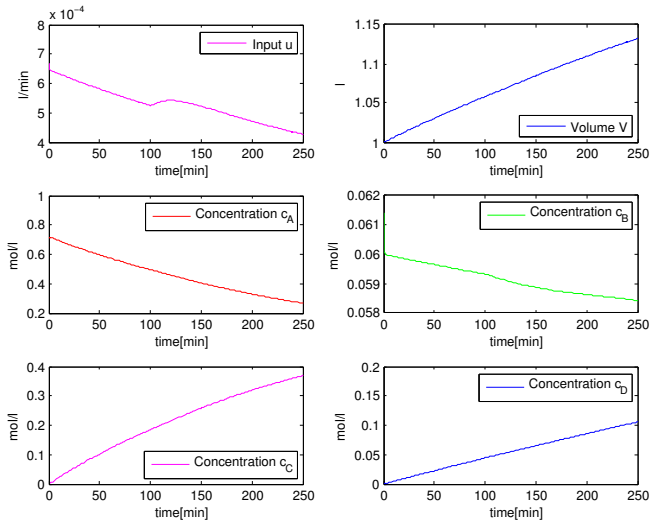

Fig. 3. Inputs and concentrations for unmeasured change of k_1 at time 100

0.2971, which is the same value as the optimal input gives.

If we had not eliminated k_1 in $\mathscr{R}(c, m_2, m_3)$, and just used the nominal value, the objective value would be $J = 0.2873$.

VII. Discussion and Conclusion

We have shown that the concepts of finding invariant variable combinations can be extended to dynamic systems, which are described by polynomial or rational equations. In addition, by not explicitly solving for the input $\mathbf{u}$, we do not have to be concerned whether the input appears in $\mathbf{c}$, because we use a P or PI controller to generate the optimal inputs. This is a simple alternative to analytically finding the optimal input by further differentiations.

Adding a controller to control $\mathbf{c}$ will often come at a negligible loss. This is confirmed in our example, where controlling the invariants using only a P controller gives virtually the same performance as when analytically solving for the optimal input.

In this work, we considered only parametric uncertainties and unmeasured states. The equally important issues of model error and measurement noise are beyond the scope of this work and have to be studied in future work.

We assumed that the uncertainty does not change the active constraints. This is valid for small disturbances. However, for larger disturbances, the new set of active constraints has to be determined.

Controlling the invariant can be used together with other NCO tracking methods to handle model mismatch or terminal constraints on a run-to-run basis, similar to [19]. Thus, we consider our method as a part of a larger procedure for implementing optimal batch performance.

Beside "normal measurements" we have also allowed measurements of their time derivatives. They may be estimated by finite differences as above, or by using some filtering. If a measurement is assumed to be reliable, then its change over time should also be possible to estimate reasonably well. Introducing measurements of the time derivatives, makes it possible to eliminate variables, for which we do not have a purely algebraic expression, and which enter through the differential equations only.

In the procedure for eliminating the adjoint variables, we have presented the common case of input affine systems. If the model is not input affine, elimination of the adjoint variables comes at the cost of introducing time derivatives of the input, which have to be measured.

We used the resultant to eliminate the unknown variables. Other techniques, such as Gröbner bases [6], could also have been applied. However, it is not easy to find appropriate monomial orderings which eliminate the unknown variables, while avoiding the trivial solution (the invariant is always zero when the model equations are satisfied).

References

[1] F. Allgöwer and A. Zheng, editors. *Nonlinear Model Predictive control*. Birkhäuser Verlag Basel, 2000.

[2] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press, New York, 1978.

[3] A. E. Bryson and Y.-C. Ho. *Applied Optimal Control, revised edition*. Taylor & Francis, 1975.

[4] L. Busé and B. Mourrain. Using the maple multires package, 2003.

[5] J. F. Canny and I. Z. Emiris. A subdivision-based algorithm for the sparse resultant. *Journal of the ACM*, 2000.

[6] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag, 1992.

[7] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer, 2005.

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

[9] I. Z. Emiris. Toric resultants and applications to geometric modelling. In M. Bronstein, A. M. Cohen, H. Cohen, D. Eisenbud, B. Sturmfels, A. Dickenstein, and I. Z. Emiris, editors, *Solving Polynomial Equations*, volume 14 of *Algorithms and Computation in Mathematics*, pages 269–300. Springer Berlin Heidelberg, 2005.

[10] I. Z. Emiris and B. Mourrain. Matrices in elimination theory,. *Journal of Symbolic Computation*, 28(1-2):3–43, 1999.

[11] H. P. Geering. *Optimal Control with Engineering Applications*. Springer, 2007.

[12] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants and Multidimensional Determinants*. Birkhäuser, Boston, MA, 1994.

[13] S. Gros, B. Srinivasan, B. Chachuat, and D. Bonvin. Neighbouring-extremal control for singular dynamic optimisation problems. part i: single-input systems. *International Journal of Control*, 82(6), 2009.

[14] M. Grötschel, S. O. Krumke, and J. Rambau, editors. *Online Optimization of Large Scale Systems*. Springer, Sept. 2001.

[15] P. Kühl, M. Diehl, T. Kraus, J. P. Schlöder, and H. G. Bock. A real-time algorithm for moving horizon state and parameter estimation. *Computers & Chemical Engineering*, 35(1):71 – 83, 2011.

[16] H. J. Marquez. *Nonlinear Control Systems*. John Wiley & Sons, 2003.

[17] M. Schlegel, K. Stockmann, T. Binder, and W. Marquardt. Dynamic optimization using adaptive control vector parameterization. *Computers & Chemical Engineering*, 29(8):1731 – 1751, 2005.

[18] S. Skogestad. Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control*, 10:487–507, 2000.

[19] B. Srinivasan and D. Bonvin. Real-time optimization of batch processes by tracking the necessary conditions of optimality. *Industrial & Engineering Chemistry Research*, 46(2):492–504, 2007.

[20] B. Srinivasan, S. Palanki, and D. Bonvin. Dynamic optimization of batch processes: I. characterization of the nominal solution. *Computers & Chemical Engineering*, 27(1):1 – 26, 2003.

[21] B. Sturmfels. On the newton polytope of the resultant. *Journal of Algebraic Combinatorics*, 3:207–236, 1994.

[22] B. Sturmfels. Solving systems of polynomial equations, 2002.

[23] P. Terwiesch, M. Agarwal, and D. W. T. Rippin. Batch unit optimization with imperfect modelling: a survey. *Journal of Process Control*, 4(4):238 – 258, 1994.