Optimal Controlled Variables for Polynomial Systems

Johannes Jäschke, Sigurd Skogestad*

Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

Abstract

We present a method for finding optimal controlled variables, which are polynomial combinations of measurements. Controlling these variables gives optimal steady state operation. Our work extends the concept of self-optimizing control; starting from the first-order necessary optimality conditions, any unknown variables are eliminated using elimination theory for polynomial systems to obtain invariant variable combinations, which contain only known variables (measurements). If a disturbance causes the active constraints to change, the invariants may be used to identify, and switch to the right region. This makes the method applicable over a wide disturbance range with changing active sets. The procedure is applied to two case studies of continuous stirred tank reactors.

Keywords: Self-optimizing control, Optimal operation, Optimization, Polynomial systems, Sparse resultants, Elimination theory

1. Introduction

For continuous processes which are operated in steady-state most of the time, an established method to achieve optimal operation in spite of varying disturbances is real-time optimization (RTO) [1]. The real-time optimizer generally uses a nonlinear steady-state model, which is updated at intervals based on measurements. This updated model is used to on-line recompute optimal setpoints for the controlled variables in the control layer below. The

*Corresponding author. Tel +47-735-94154. Email address: skoge@chemeng.ntnu.no

concept has gained acceptance in industry and is increasingly used for improving plant performance. However, installing an RTO system and maintaining it generally entails large costs.

An alternative approach to optimizing plant performance is to use a process model off-line to find a "self-optimizing control" structure. The basic concept of self-optimizing control was conceived by [2], who wrote that we "want to find a function $\mathbf{c}$ of the process variables which when held constant leads automatically to the optimal adjustments of the manipulated variables". However, they did not provide any method for identifying this function. The idea is to use this function as a controlled variable and keep it at a constant setpoint by simple control structures, e.g. PID controllers, or by more complex model predictive controllers (MPC). Using this kind of controlled variables disburdens the real-time optimizer [3], or may even make it superfluous.

The term "self-optimizing control" was coined in the context of controlled variable selection with the purpose of describing the practical goal of finding "smart" controlled variables $\mathbf{c}$:

> "Self-optimizing control is achieved if a constant setpoint policy results in an acceptable loss L (without the need to re-optimize when disturbances occur)." [4]

Many industrial processes are operated using self-optimizing control, although this term may not be used. Optimally active constraints may be considered as self-optimizing variables, for example, the use of maximum cooling for a compressor. However, the more difficult problem is to identify self-optimizing control variables associated with unconstrained degrees of freedom. An example for the unconstrained case is controlling the air/fuel ratio to a combustion engine at a constant value.

The concept of self-optimizing control enables us to separate the two problems of optimizing the system and designing the controller. Thus, in a first step the controlled variables $\mathbf{c}$ are determined based on steady-state optimization of the process, and in a second step a suitable controller is designed. In most cases, a PI controller will be sufficient, but also more advanced controllers can be used to control the self-optimizing variable. The advantage of this two-step approach is that it makes it possible to focus completely on steady-state optimal behavior when designing the control structure, while all issues which arise when handling dynamic systems are considered separately when designing the actual controllers.

In the last decade, several contributions have been made to the systematic search of controlled variables which have self-optimizing properties [5, 6, 7, 8, 9], but only for cases with linear measurement models and a quadratic cost function. This results in linear measurement combinations $\mathbf{c} = \mathbf{H}\mathbf{y}$ as controlled variables. Here, $\mathbf{y}$ includes all available measurements, and the goal is to find a good selection or combination matrix $\mathbf{H}$. In cases where a higher-order curvature is present at the optimum, the loss imposed by using linear measurement combinations may not be acceptable, and the controlled variables are not self-optimizing.

It has been noted previously [10, 11, 5, 12, 13], that the gradient of the cost function with respect to the degrees of freedom $\mathbf{u}$ would be the ideal controlled variable, $\mathbf{c} = \mathbf{J_u}$. However, the gradient $\mathbf{J_u}$ is usually not directly measurable, and analytical expressions for the gradient generally contain variables which are unmeasured (unknown disturbances). The concepts from self-optimizing control theory can be thought of as methods for identifying a measurement or a measurement combination $\mathbf{c}(\mathbf{y})$, which approximates the gradient (in some "best" way).

The main contribution of this work is to use polynomial elimination theory to extend the ideas of self-optimizing control, in particular the concept of the null-space method [6], to constrained systems described by multivariable polynomials. This results in controlled variables which are polynomials in the measurements, $\mathbf{c}(\mathbf{y})$.

A summary of the proposed procedure for achieving steady state optimal operation is given in Table 1. In steps 1 and 2 we formulate the optimization problem and determine regions of constant active constraints, also called critical regions. This is done by offline calculations, for example, by gridding the disturbance space with a sufficiently fine grid and optimizing the process for each grid point. In step 3, for each critical region, (a) the optimality conditions are formulated, and (b) the Lagrangian multipliers are eliminated. Then (c) the unknown variables, i.e. the disturbances and the internal state variables are eliminated from the optimality conditions to obtain an invariant variable combination $\mathbf{c}(\mathbf{y})$ which contains only measured variables and known parameters. Optimal operation is achieved in each critical region by controlling the active constraints and the invariant measurement combinations, step 4. Finally, we monitor the active constraints and the invariants of the neighboring regions to determine when to switch to a new region.

The rest of this paper is structured as follows: The next section contains the problem formulation and derives an expression for the optimality condi-

1. Formulate optimization problem
2. For the expected set of disturbances, find all regions with different sets of active constraints $\mathcal{A}_i$
3. For each region of active constraints $\mathcal{A}_i$
   a Formulate optimality conditions
   b Eliminate Lagrangian multipliers $\boldsymbol{\lambda}$ from optimality conditions to obtain invariants $\mathbf{J}_{\mathbf{z},red}$ (reduced gradient)
   c Obtain measurement invariants $\mathbf{c}(\mathbf{y})$ by eliminating unknowns, such that
   $$\mathbf{c}(\mathbf{y}) = 0 \Longleftrightarrow \mathbf{J}_{\mathbf{z},red} = 0$$
4. In each region $\mathcal{A}_i$
   a Control active constraints
   b Control the invariants $\mathbf{c}(\mathbf{y})$

Use controlled variables and measured constraints for changing regions

Table 1: Procedure for finding nonlinear invariants as controlled variables

tions which does not contain Lagrangian multipliers. Sections 3 and 4 show how the unknown states and disturbances can be eliminated from the optimality conditions without explicitly solving for them. In Section 5 we give an example, followed by a discussion on changing active constraints (Section 6). Section 7 presents a case study with changing active constraints, and our paper is closed with a discussion and conclusions in Sections 8 and 9.

2. Optimal operation using the optimality conditions

2.1. Problem formulation

Steady state optimal operation is defined as minimizing a scalar cost index $J(\mathbf{u}, \mathbf{x}, \mathbf{d})$ subject to satisfying the model equations, $\mathbf{g} = 0$, and operational constraints, $\mathbf{h} \leq 0$:

$$\min_{\mathbf{u},\mathbf{x}} J(\mathbf{u}, \mathbf{x}, \mathbf{d}) \quad \text{s.t} \quad \begin{cases} \mathbf{g}(\mathbf{u}, \mathbf{x}, \mathbf{d}) & = 0 \qquad \text{(model)} \\ \mathbf{h}(\mathbf{u}, \mathbf{x}, \mathbf{d}) & \leq 0 \qquad \text{(constraint).} \end{cases} \tag{1}$$

Here $\mathbf{u} \in \mathbb{R}^{n_\mathbf{u}}$, $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$, $\mathbf{d} \in \mathbb{R}^{n_\mathbf{d}}$ denote the manipulated input variables, the internal state variables, and the unmeasured disturbance variables, respectively. In this paper, the J is assumed to be a polynomial in the polynomial

ring $\mathbb{R}[\mathbf{u}, \mathbf{x}, \mathbf{d}]$, that is, a polynomial in the variables $\mathbf{x}, \mathbf{u}$ and $\mathbf{d}$ with coefficients in $\mathbb{R}$. Similarly, the functions $\mathbf{g}$ and $\mathbf{h}$ are assumed to be vectors with elements in the polynomial ring $\mathbb{R}[\mathbf{u}, \mathbf{x}, \mathbf{d}]$.

In addition, we assume that we have measurements $\mathbf{y} \in \mathbb{R}^{n_\mathbf{y}}$, which are polynomial functions of $\mathbf{u}$, $\mathbf{x}$ and $\mathbf{d}$, which provide information about internal states, inputs, and disturbances. To handle the measurements in a consistent way when dealing with polynomials, we write the measurement relations implicitly as

$$\mathbf{m}(\mathbf{u}, \mathbf{x}, \mathbf{d}, \mathbf{y}) = 0, \tag{2}$$

with $\mathbf{m}(\mathbf{u}, \mathbf{x}, \mathbf{d}, \mathbf{y}) \in \mathbb{R}[\mathbf{u}, \mathbf{x}, \mathbf{d}, \mathbf{y}]$. To simplify notation, we combine the state and input variables in a vector $\mathbf{z} \in \mathbb{R}^{n_\mathbf{z}}$,

$$\mathbf{z} = \begin{bmatrix} \mathbf{u} \\ \mathbf{x} \end{bmatrix}. \tag{3}$$

Problem (1) is similar to the one solved on-line at given sample times when using real-time optimization (RTO). In this work, however, we do not solve the optimization problem on-line; instead, we analyze the problem using offline calculations in order to find good controlled variables $\mathbf{c}(\mathbf{y})$, which yield optimal operation when controlled at fixed setpoints, even for a change in the disturbance $\mathbf{d}$.

2.1.1. Optimality conditions

Let $\mathbf{z}^*$ be a feasible point of optimization problem (1), and assume that all gradient vectors $\nabla_\mathbf{z} g_i(\mathbf{z}^*, \mathbf{d})$ and $\nabla_\mathbf{z} h_i(\mathbf{z}^*, \mathbf{d})$ associated with $g_i(\mathbf{z}^*, \mathbf{d}) = 0$ (model) and $h_i(\mathbf{z}^*, \mathbf{d}) = 0$ (active constraints), are linearly independent (linear independence constraint qualification, (LICQ)).

If $\mathbf{z}^*$ is locally optimal, then there exist Lagrangian multiplier vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$, such that the following conditions, known as the KKT conditions, are satisfied [14, 15]:

$$\begin{aligned} \nabla_\mathbf{z} J(\mathbf{z}^*, \mathbf{d}) + [\nabla_\mathbf{z} \mathbf{g}(\mathbf{z}^*, \mathbf{d})]^\mathbf{T} \boldsymbol{\lambda} + [\nabla_\mathbf{z} \mathbf{h}(\mathbf{z}^*, \mathbf{d})]^\mathbf{T} \boldsymbol{\nu} &= 0 \\ \mathbf{g}(\mathbf{z}^*, \mathbf{d}) &= 0 \\ \mathbf{h}(\mathbf{z}^*, \mathbf{d}) &\leq 0 \\ [\mathbf{h}(\mathbf{z}^*, \mathbf{d})]^\mathbf{T} \boldsymbol{\nu} &= 0 \\ \boldsymbol{\nu} &\geq 0. \end{aligned} \tag{4}$$

When optimizing nonlinear systems, such as polynomial systems, there are several complications which may arise. The optimality conditions (4) will in general not have a unique solution. There may be multiple maxima, minima and saddle points, so finding the global minimum is not an easy task in itself. When a solution to (4) is found, it has to be checked whether it indeed is the desired solution (minimum). In addition, there may be solutions which are not physical (complex). Before controlling a controlled variable which is based on the first-order optimality condition, it has to be assured that the process actually is at the desired optimum.

These and other issues from nonlinear and polynomial optimization are not addressed in this work. The focus of this paper is rather to present a method which gives a controlled variable $\mathbf{c}(\mathbf{y})$ which is a function of measurements $\mathbf{y}$, and which is zero at all points that satisfy the KKT conditions, while it is nonzero whenever the KKT conditions are not satisfied.

2.2. Partitioning into sets of active constraints

Generally, the set of inequality constraints $h_i(\mathbf{z}, \mathbf{d}) \leq 0$ that are active varies with the value of the elements in $\mathbf{d}$. The disturbance space can hence be partitioned into regions which are characterized by their individual set of active constraints. These regions will be called critical regions.

The concept of critical regions allows one to decompose the original optimization problem (1) into a set of equality constrained optimization problems, which are valid in the corresponding critical region. This idea is also applied in multi-parametric programming [16]. However, we do not search for an explicit expression for the inputs $\mathbf{u}^*$, as in multi-parametric programming. We rather use each subproblem to find good controlled variables $\mathbf{c}$ for the corresponding critical region.

In order to obtain a fully specified system in each region,

1. the active constraints need to be controlled, and
2. a controlled variable must be controlled for each unconstrained degree of freedom.

For independent constraints and model equations, the number of unconstrained degrees of freedom, $n_{\mathbf{c}} = n_{DOF}$, is calculated according to

$$n_{DOF} = n_{\mathbf{z}} - n_{\mathbf{g}} - n_{\mathbf{h},active}, \tag{5}$$

where $n_{\mathbf{z}}$, $n_{\mathbf{g}}$, $n_{\mathbf{h},active}$ denote the number of variables $\mathbf{z}$, the number of model equations, $\mathbf{g}$, and the number of constraints from $\mathbf{h}$ which are active ($h_i = 0$).

Remark 1. *The presented method for finding the degrees of freedom is valid when the polynomials are algebraically independent. A more rigorous way to determine the degrees of freedom would be to examine the dimension of the variety defined by* $\mathbf{g}$ *and* $\mathbf{h}_{active}$ *[17, 18].*

Remark 2. *When the optimization problem* (1) *is composed of polynomial equations, the critical regions are defined by semialgebraic sets in* $\mathbb{R}^{n_\mathbf{u}+n_\mathbf{x}+n_\mathbf{d}}$ *[19]. A semialgebraic set is defined as the finite union of sets defined by a finite number of polynomial equalities and inequalities,*

$$g_i(\mathbf{x}, \mathbf{u}, \mathbf{d}) = 0, \quad i = 1 \dots n_\mathbf{g} \tag{6}$$

and

$$h_j(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0, \quad j = 1 \dots n_\mathbf{h} \tag{7}$$

Where g_i *and* h_j *are polynomials in the variables* $\mathbf{x}, \mathbf{u}$ *and* $\mathbf{d}$*, with coefficients in* $\mathbb{R}$*. Loosely speaking, a semialgebraic set can be thought of a set defined by a finite number of polynomial inequalities. The interior of an ellipsoid, or the set of points on a curve in the* $\mathbb{R}^n$ *are examples of semialgebraic sets.*

In the rest of the paper, to simplify notation, all active constraints $h_i(\mathbf{z}, \mathbf{d}) = 0$ are included in the equality constraint vector $\mathbf{g}(\mathbf{z}, \mathbf{d}) = 0$. Then in every critical region, optimization problem (1) can be written as

$$\begin{aligned} &\min_\mathbf{z} J(\mathbf{z}, \mathbf{d}) \\ &\text{s.t.} \\ &\mathbf{g}(\mathbf{z}, \mathbf{d}) = 0. \end{aligned} \tag{8}$$

The KKT first-order optimality conditions (4) simplify for problem (8) to

$$\begin{aligned} \nabla_\mathbf{z} J(\mathbf{z}, \mathbf{d}) + [\nabla_\mathbf{z} \mathbf{g}(\mathbf{z}, \mathbf{d})]^\mathbf{T} \boldsymbol{\lambda} = 0, \\ \mathbf{g}(\mathbf{z}, \mathbf{d}) = 0. \end{aligned} \tag{9}$$

These expressions cannot be used for control, because they still contain unknown variables, $\mathbf{x}$ (in $\mathbf{z} = [\mathbf{u}, \mathbf{x}]$), $\mathbf{d}$, and $\boldsymbol{\lambda}$, which must be eliminated.

2.3. Eliminating the Lagrangian multipliers $\boldsymbol{\lambda}$

In every critical region, a control structure that gives optimal operation has to satisfy (9).

Theorem 1. *Given optimization problem* (8)*, where we assume that the LICQ hold, and let* $\mathbf{N}(\mathbf{z},\mathbf{d}) \in \mathbb{R}^{n_\mathbf{z}\times(n_\mathbf{z}-n_\mathbf{g})}$ *be a basis for the null space of* $\nabla_\mathbf{z}\mathbf{g}(\mathbf{z},\mathbf{d})$*. Controlling the active constraints* $\mathbf{g}(\mathbf{z},\mathbf{d}) = 0$*, and the variable combination* $\mathbf{J}_{\mathbf{z},red} = [\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d}) = 0$ *then results in optimal steady-state operation.*

Proof. Select $\mathbf{N}(\mathbf{z},\mathbf{d})$ such that $\nabla_\mathbf{z}\mathbf{g}(\mathbf{z},\mathbf{d})\mathbf{N}(\mathbf{z},\mathbf{d}) = 0$. Since the LICQ are satisfied, the constraint Jacobian $\nabla_\mathbf{z}\mathbf{g}(\mathbf{z},\mathbf{d})$ has full row rank and $\mathbf{N}(\mathbf{z},\mathbf{d})$ is well defined and does not change dimension within the region. The first equation from the optimality conditions (9) is premultiplied by $[\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}$ to yield

$$\begin{aligned}[\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\left(\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d}) + [\nabla_\mathbf{z}\mathbf{g}(\mathbf{z},\mathbf{d})]^\mathbf{T}\boldsymbol{\lambda}\right) &= [\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z}.\mathbf{d}) + \underline{0}\boldsymbol{\lambda} \\ &= [\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d}).\end{aligned} \tag{10}$$

Since $\mathbf{N}(\mathbf{z},\mathbf{d})$ has full rank, we have that (9) are satisfied whenever $\mathbf{g}(\mathbf{z},\mathbf{d}) = 0$ and $[\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d}) = 0$. □

We call $\mathbf{J}_{\mathbf{z},red} = [\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d})$ the reduced gradient. By construction, the reduced gradient has $n_{DOF} = n_\mathbf{z} - n_\mathbf{g}$ elements. Controlling

$$\mathbf{J}_{\mathbf{z},red} = [\mathbf{N}(\mathbf{z},\mathbf{d})]^\mathbf{T}\nabla_\mathbf{z}J(\mathbf{z},\mathbf{d}) = 0 \tag{11}$$

together with the active constraints, $\mathbf{g}(\mathbf{z},\mathbf{d}) = 0$, fully specifies the system at the optimum and is equivalent to controlling the first-order optimality conditions (9). However, $\mathbf{J}_{\mathbf{z},red}$ cannot generally be used for control directly because it depends on the variables $\mathbf{d}$ and $\mathbf{x}$, which are usually unknown. Thus, we would like to eliminate the unknown disturbances $\mathbf{d}$ and the internal states $\mathbf{x}$ from the expression (11). The simplest approach (Approach 1) is to solve the measurement equations $\mathbf{m}(\mathbf{x},\mathbf{u},\mathbf{d},\mathbf{y}) = 0$ and the active constraints $\mathbf{g}(\mathbf{z},\mathbf{d}) = 0$ for the unknowns in $\mathbf{z}$ and $\mathbf{d}$, and substitute the solution into $\mathbf{J}_{\mathbf{z},red}$. To do this, we need as many equations as unknowns. As we show next, this elimination method is straightforward in case of linear equations, but it becomes significantly more complicated when working with polynomials of higher degree.

Alternatively (Approach 2), we search for necessary and sufficient conditions which guarantee that the measurement model $\mathbf{m}(\mathbf{x},\mathbf{u},\mathbf{d},\mathbf{y}) = 0$, the active constraints and the model $\mathbf{g}(\mathbf{z},\mathbf{d}) = 0$, and the reduced gradient $\mathbf{J}_{\mathbf{z},red} = 0$ are satisfied at the same time. We require that the necessary

and sufficient condition is a function of measurements $\mathbf{y}$ and known parameters, only. This more general approach is discussed below for the linear quadratic case (Section 3) before it is generalized for the polynomial case (Section 4).

3. Elimination for linear quadratic systems

The optimization problem we consider is

$$\begin{aligned} \min_{\mathbf{z}} J(\mathbf{z},\mathbf{d}) &= \left[\begin{array}{cc} \mathbf{z^T} & \mathbf{d^T} \end{array}\right] \left[\begin{array}{cc} \mathbf{J_{zz}} & \mathbf{J_{zd}} \\ \mathbf{J_{zd}}^{\mathbf{T}} & \mathbf{J_{dd}} \end{array}\right] \left[\begin{array}{c} \mathbf{z} \\ \mathbf{d} \end{array}\right] \\ &\text{s.t.} \\ \mathbf{g}(\mathbf{z}) &= \mathbf{Az} - \mathbf{b} = 0, \end{aligned} \tag{12}$$

and the linear measurement model is

$$\begin{aligned} \mathbf{m}(\mathbf{z},\mathbf{d},\mathbf{y}) &= \mathbf{y} - [\mathbf{G^y}\ \mathbf{G_d^y}] \left[\begin{array}{c} \mathbf{z} \\ \mathbf{d} \end{array}\right] = 0 \\ &= \mathbf{y} - \tilde{\mathbf{G}}^{\mathbf{y}} \left[\begin{array}{c} \mathbf{z} \\ \mathbf{d} \end{array}\right] = 0. \end{aligned} \tag{13}$$

We consider $[\mathbf{z},\mathbf{d}]^{\mathbf{T}}$ as unknown and we assume that (12) has a solution, $\mathbf{J_{zz}} > 0$, and $\mathbf{A}$ has full rank. If a variable in $\mathbf{z}$ is measured, we include it also in $\mathbf{y}$.

The null space of the constraint gradient, $\mathbf{N}$, is a constant matrix which is independent of $\mathbf{z}$, such that $\mathbf{AN} = 0$. The first-order necessary optimality conditions require that at the optimum

$$\mathbf{J}_{\mathbf{z},red} = \mathbf{N^T}\nabla_{\mathbf{z}} J(\mathbf{z},\mathbf{d}) = \mathbf{N^T} \left[\begin{array}{cc} \mathbf{J_{zz}} & \mathbf{J_{zd}} \end{array}\right] \left[\begin{array}{c} \mathbf{z} \\ \mathbf{d} \end{array}\right] = 0. \tag{14}$$

Approach 1. If the number of independent measurements ($n_{\mathbf{y}}$) is greater or equal to the number of unknown variables ($n_{\mathbf{z}} + n_{\mathbf{d}}$), the measurement relations (13) can be solved for the unknowns,

$$\left[\begin{array}{c} \mathbf{z} \\ \mathbf{d} \end{array}\right] = \left[\tilde{\mathbf{G}^{\mathbf{y}}}\right]^{\dagger} \mathbf{y}, \tag{15}$$

and substituted into the gradient expression (14) to obtain

$$\mathbf{c}(\mathbf{y}) = \mathbf{N^T} \left[\begin{array}{cc} \mathbf{J_{zz}} & \mathbf{J_{zd}} \end{array}\right] \left[\tilde{\mathbf{G}^{\mathbf{y}}}\right]^{\dagger} \mathbf{y}. \tag{16}$$

Here, $(\cdot)^\dagger$ denotes the pseudo-inverse of $(\cdot)$. Controlling $\mathbf{c}(\mathbf{y}) =$ and the active constraints $\mathbf{Az} - \mathbf{b}$ to zero, then results in optimal operation.

When there are no constraints, we have that $\mathbf{z} = \mathbf{u}$, and this method results in the null space method [6]. In this case, $\mathbf{N}$ may be set to any nonsingular matrix, for example the identity matrix $\mathbf{N} = \mathbf{I}$, and we get the same result as in [8],

$$\mathbf{c}(\mathbf{y}) = \begin{bmatrix} \mathbf{J_{uu}} & \mathbf{J_{ud}} \end{bmatrix} \left[\tilde{\mathbf{G}^{\mathbf{y}}}\right]^\dagger \mathbf{y}. \tag{17}$$

Approach 2. In the case of polynomial equations of higher degrees it is generally difficult to solve for the unknown variables, as done in (15). Therefore, we consider the problem from a slightly different perspective.

We assume for the moment that $n_{\mathbf{y}} = n_{\mathbf{z}} + n_{\mathbf{d}}$, and that $\tilde{\mathbf{G}^{\mathbf{y}}} = [\mathbf{G^y}\ \mathbf{G_d^y}]$ is invertible. Consider the elements of the reduced gradient vector (14), one at a time, together with all the measurement equations (13). Let the superscript (i) denote the i-th row of a matrix or a vector. We write the reduced gradient (14) together with the measurement equations (13) as a sequence of square linear systems

$$\underbrace{\begin{bmatrix} [\mathbf{N^T J_{zz}}]^{(i)} & [\mathbf{N^T J_{zd}}]^{(i)} & 0 \\ \mathbf{G^y} & \mathbf{G_d^y} & \mathbf{y} \end{bmatrix}}_{\mathbf{M}^{(i)}} \begin{bmatrix} \mathbf{z} \\ \mathbf{d} \\ -1 \end{bmatrix} = 0 \quad i = 1 \ldots n_{DOF}. \tag{18}$$

Here, the $\mathbf{M}^{(i)}$ are square matrices of size $(n_{\mathbf{y}} + 1)$. We want to find a particular output combination which satisfies (18). A unique solution for $[\mathbf{z}, \mathbf{d}]^{\mathbf{T}}$ exists only if $\mathrm{rank}(\mathbf{M^{(i)}}) = n_{\mathbf{y}} = n_{\mathbf{z}} + n_{\mathbf{d}}$. The submatrix $[\mathbf{G^y}\ \mathbf{G_d^y}\ \mathbf{y}]$ already has rank $n_{\mathbf{y}}$, irrespective of the value of $\mathbf{y}$ (or the control policy that generates the input $\mathbf{u}$ which in turn generates $\mathbf{y}$). This follows because $[\mathbf{G^y}\ \mathbf{G_d^y}\ \mathbf{y}]$ has more columns than rows, and because $\mathrm{rank}([\mathbf{G^y}\ \mathbf{G_d^y}]) = n_{\mathbf{y}}$. Therefore, the condition for a nontrivial common solution is:

$$\det(\mathbf{M}^{(i)}) = 0 \quad \text{for all } i = 1..n_{DOF}. \tag{19}$$

This condition guarantees that a common solution to (18) exists, so the elements of the controlled variable $\mathbf{c}$ are selected as $c_i = \det(\mathbf{M}^{(i)})$.

It remains to show that controlling the determinants $c_i = \det(\mathbf{M}^{(i)})$ gives the inputs which lead to the optimum. Since the system is linear and the rank of the measurement equations is $n_{\mathbf{y}}$, there is a unique linear invertible

Table 2: Gain values for Example 1

Variable	Value
G_1^y	0.9
$G_{d,1}^y$	0.1
G_2^y	0.5
$G_{d,2}^y$	-1.0

mapping between the measurements $\mathbf{y}$ and the vector $[\mathbf{z}, \mathbf{d}]^{\mathbf{T}}$. Therefore every value of $\mathbf{y}$ corresponds uniquely to some value in $\mathbf{z}$.

In the case with more measurements, $n_{\mathbf{y}} > n_{\mathbf{z}} + n_{\mathbf{d}}$, any subset of $n_{\mathbf{z}} + n_{\mathbf{d}}$ measurements may be chosen such that rank($[\mathbf{G^y}\ \mathbf{G_d^y}]$)=$n_{\mathbf{z}} + n_{\mathbf{d}}$.

Remark 3. *For simplicity, we chose to use the measurements to eliminate the internal states. In practice we would use the constraint equations* $\mathbf{Az} - \mathbf{b} = 0$ *in addition to the measurements for elimination in the matrices* $\mathbf{M}^{(i)}$*. Then we only need* $n_{\mathbf{y}} \geq n_{\mathbf{u}} + n_{\mathbf{d}}$ *independent measurements, where we assume that the degrees of freedom* $\mathbf{u}$ *are included in the measurement vector* $\mathbf{y}$*. Thus we need as many equations (measurements+constraints) as variables to eliminate.*

Example 1 (Linear model and quadratic objective)**.** *This example demonstrates that the "determinant method" gives the same result as the previously published null-space method [6]. Consider a system from [20]. The quadratic cost to minimize is*

$$J = (u - d)^2, \tag{20}$$

and the measurement relations are

$$\begin{aligned} y_1 &= G_1^y u + G_{d,1}^y d \\ y_2 &= G_2^y u + G_{d,2}^y d. \end{aligned} \tag{21}$$

The values of the gains are given in Table 2. We are searching for a condition on the measurements y_1 *and* y_2 *such that the optimality condition is satisfied. The gradient is* $\nabla_u J = 2(u - d)$ *and* $J_{uu} = 2$*,* $J_{ud} = -2$*. It is easily verified that measurements are linearly independent. Using* (18)*, this gives*

an equation system of 3 equations in 2 unknowns:

$$\mathbf{M}\begin{bmatrix} u \\ d \\ -1 \end{bmatrix} = 0, \tag{22}$$

where

$$\mathbf{M} = \begin{bmatrix} J_{uu} & J_{ud} & 0 \\ G_1^y & G_{d,1}^y & y_1 \\ G_2^y & G_{d,2}^y & y_2 \end{bmatrix}. \tag{23}$$

Equation (22) *has a nontrivial solution if and only if* $\det(\mathbf{M}) = 0$. *Therefore the necessary and sufficient condition for the existence of a nontrivial solution is*

$$\begin{aligned} c = \det(\mathbf{M}) &= -y_1(J_{uu}G_{d,2}^y - G_2^y J_{ud}) + y_2(J_{uu}G_{d,1}^y - G_1^y J_{ud}) \\ &= 0. \end{aligned} \tag{24}$$

Inserting the parameter values from Table 2 gives

$$c = \det(\mathbf{M}) = y_1 + 2y_2. \tag{25}$$

Thus, controlling $c = y_1 + 2y_2$ *to zero yields optimal operation. This is the same result as found by applying the null-space method in [20].*

Even though obtaining the invariants via the determinant (Approach 2) may seem cumbersome, it eliminates the necessity of inverting the measurements and solving for the unknowns (Approach 1). While this is of little advantage for systems of linear equations, Approach 2 can be generalized for systems of polynomial equations which cannot easily be solved for the right set of unknowns.

4. Elimination for systems of polynomial equations

4.1. The problem

We consider the optimization problem (8), where all functions are polynomials in $\mathbf{u}, \mathbf{x}$ and $\mathbf{d}$. Let $\hat{\mathbf{d}}$ now denote the vector of all unmeasured (unknown) variables,

$$\hat{\mathbf{d}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{d} \end{bmatrix}, \tag{26}$$

not only including disturbances $\mathbf{d}$, but also unknown states $\mathbf{x}$, and let $\mathbf{y}$ include all measurements, including all inputs. Thus, every variable belongs either to $\hat{\mathbf{d}}$ or $\mathbf{y}$, and we write the optimality conditions as

$$\begin{aligned} \mathbf{J}_{\mathbf{z},red}(\mathbf{y},\hat{\mathbf{d}}) &= 0 \\ \mathbf{g}(\mathbf{y},\hat{\mathbf{d}}) &= 0, \end{aligned} \tag{27}$$

and the measurement relations as

$$\mathbf{m}(\mathbf{y},\hat{\mathbf{d}}) = 0. \tag{28}$$

Remark 4. *Note that in the elimination step, we do not distinguish between internal states variables* $\mathbf{x}$ *and external disturbances* $\mathbf{d}$*. All variables which are not available as measurements (that is,* $\hat{\mathbf{d}} = [\mathbf{x},\mathbf{d}]^{\mathbf{T}}$*) have to be eliminated from the optimality conditions using* $\mathbf{g}$ *and* $\mathbf{m}$*.*

For polynomial equations, eliminating the unknown variables from $\mathbf{J}_{\mathbf{z},red}$ is not as straightforward as in the linear case, as we cannot just solve the measurement equations for the unknowns and insert them in to the expression of $\mathbf{J}_{\mathbf{z},red}$ (Approach 1). Even for the case of a univariable polynomial of degree 5 and higher, for example $d^5 - d + 1 = 0$, there exist no general analytic solution formulas, as was proven by [21]. Therefore we need to find another way to eliminate the unknown variables $\hat{\mathbf{d}}$ from $\mathbf{J}_{\mathbf{z},red}(\mathbf{y},\hat{\mathbf{d}}) = 0$ without solving $\mathbf{g}$ and $\mathbf{m}$ for them first. For linear systems, we used the determinant in (18) (Approach 2). The generalization of the determinant to systems of polynomial equations is called the resultant. According to [22],

> "the resultant of an overconstrained polynomial system characterizes the existence of common roots as a condition on the input coefficients".

4.2. Results from polynomial elimination theory

For the elimination procedure, we consider multivariate polynomials $f \in \mathbb{R}(\mathbf{y})[\hat{\mathbf{d}}]$, that is, polynomials in the variables $\hat{\mathbf{d}}$, whose coefficients are functions of $\mathbf{y}$ (that is, polynomials with variables $\mathbf{y}$ and coefficients in $\mathbb{R}$). Given an $n_{\hat{\mathbf{d}}}$-tuple,

$$\boldsymbol{\alpha}_{i,j} = \left(\alpha_{i,j}(1), \alpha_{i,j}(2), \ldots, \alpha_{i,j}(n_{\hat{\mathbf{d}}})\right), \tag{29}$$

we use the shorthand notation

$$\hat{\mathbf{d}}^{\boldsymbol{\alpha}_{i,j}} = \hat{\mathbf{d}}_1^{\alpha_{i,j}(1)} \hat{\mathbf{d}}_2^{\alpha_{i,j}(2)} \ldots \hat{\mathbf{d}}_{n_{\hat{\mathbf{d}}}}^{\alpha_{i,j}(n_{\hat{\mathbf{d}}})}. \tag{30}$$

Then we can write a system of n polynomials in compact form

$$f_i(\mathbf{y}, \hat{\mathbf{d}}) = \sum_{j=0}^{k_i} a_{i,j}(\mathbf{y}) \hat{\mathbf{d}}^{\boldsymbol{\alpha}_{i,j}}, \quad i = 1..n, \tag{31}$$

where the coefficients $a_{ij}(\mathbf{y}) \neq 0$ are polynomials in $\mathbb{R}[\mathbf{y}]$, that is, polynomials in $\mathbf{y}$ with coefficients in $\mathbb{R}$.

We consider the functions $a_{i,j}(\mathbf{y})$ as polynomial coefficients, and $\hat{\mathbf{d}}$ as variables. For every polynomial f_i, we collect the exponent vectors in the set $\mathcal{E}_i = \{\boldsymbol{\alpha}_{i,1}, \ldots, \boldsymbol{\alpha}_{i,k_i}\}$. This set is called support of the polynomial f_i.

The support of the polynomial $f = d_1^2 + d_1 d_2 - 1$, for example, is $\mathcal{E} = \{(2,0), (1,1), (0,0)\}$. We denote as $Q_i = \text{conv}(\mathcal{E}_i)$ the convex hull of the support of a polynomial (that is the smallest convex set in $\mathbb{R}^{n_{\hat{\mathbf{d}}}}$ containing $\mathcal{E}$).

Further, we denote the set of complex numbers without zero as $\mathbb{C}^*$ ($\mathbb{C}^* = \mathbb{C} \setminus 0$).

Next we present some basic concepts from algebraic geometry taken from [23].

Definition 1 (Affine variety)**.** *Consider $f_1, \ldots, f_n$ polynomials in $\mathbb{C}[\hat{\mathbf{d}}_1, \ldots \hat{\mathbf{d}}_{n_{\hat{\mathbf{d}}}}]$. The affine variety defined by $f_1, \ldots, f_n$ is the set*

$$V(f_1, \cdots, f_n) = \left\{ (\hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_{n_{\hat{\mathbf{d}}}}) \in \mathbb{C}^{n_{\hat{\mathbf{d}}}} : f_i(\hat{\mathbf{d}}_1, \ldots \hat{\mathbf{d}}_{n_{\hat{\mathbf{d}}}}) = 0 \quad i = 1 \ldots n \right\} \tag{32}$$

Casually speaking, the variety is the set of all solutions in $\mathbb{C}^{n_{\hat{\mathbf{d}}}}$.

Definition 2 (Zariski closure)**.** *Given a subset $S \subset \mathbb{C}^m$, there is a smallest affine variety $\bar{S} \subset \mathbb{C}^m$ containing S. We call $\bar{S}$ the Zariski closure of S.*

Let $L(\mathcal{E}_i)$ be the set of all polynomials whose terms all have exponents in the support $\mathcal{E}_i$:

$$L(\mathcal{E}_i) = \left\{ a_{i,1} \hat{\mathbf{d}}^{\boldsymbol{\alpha}_{i1}} + \cdots + a_{i,k_i} \hat{\mathbf{d}}^{\boldsymbol{\alpha}_{ik_i}} : a_{i,j} \in \mathbb{C} \right\} \tag{33}$$

The coefficients $a_{i,j}$ of a given polynomial then define a point in $\mathbb{C}^{k_i}$. Now let

$$Z(\mathcal{E}_1, \ldots \mathcal{E}_n) \subset L(\mathcal{E}_1) \times \cdots \times L(\mathcal{E}_n) \tag{34}$$

be the Zariski closure of the set of all $(f_1, \dots f_n)$ for which (31) has a solution in $(\mathbb{C}^*)^{n_{\hat{\mathbf{a}}}}$ (that is the Zariski closure of the points defined by all coefficients $a_{i,j} \in \mathbb{C}$ for which (31) has a root). For an overdetermined system of $n_{\hat{\mathbf{d}}} + 1$ polynomials in $n_{\hat{\mathbf{d}}}$ variables we have following result:

Lemma 1 (Sparse resultant)**.** *Assume that $Q_i = \mathrm{conv}\,(\mathcal{E}_\mathrm{i})$ is an $n_{\hat{\mathbf{d}}}$-dimensional polytope for $i = 1, \dots, n_{\hat{\mathbf{d}}} + 1$. Then there is an irreducible polynomial $\mathcal{R}$ in the coefficients of the f_i such that*

$$(f_1, \dots, f_{n_{\hat{\mathbf{d}}}+1}) \in Z(\mathcal{E}_1, \dots, \mathcal{E}_{n_{\hat{\mathbf{d}}}+1}) \Longleftrightarrow \mathcal{R}(f_1, \dots, f_{n_{\hat{\mathbf{d}}}+1}) = 0. \tag{35}$$

In particular, if

$$f_1(d_1 \dots d_{n_{\hat{\mathbf{d}}}}) = \dots = f_{n_{\hat{\mathbf{d}}}+1}(d_1 \dots d_{n_{\hat{\mathbf{d}}}}) = 0 \tag{36}$$

has a solution $(\hat{\mathbf{d}}_1, \dots, \hat{\mathbf{d}}_{n_{\hat{\mathbf{d}}}})$ in $(\mathbb{C}^*)^{n_{\hat{\mathbf{a}}}}$, then

$$\mathcal{R}(f_1, \dots, f_{n_{\hat{\mathbf{d}}}+1}) = 0. \tag{37}$$

For a proof and a detailed treatment of the sparse resultant, we refer to e.g. [24, 23].

Remark 5. *The requirement that Q_i has to be $n_{\hat{\mathbf{d}}}$-dimensional is no restriction and can be relaxed, [25]. However, for simplicity, we chose to present this result here.*

Depending on the allowed space for the roots, there are other resultant types (e.g. Bezout resultants and Dixon resultants for system of homogeneous polynomials), with different algorithms to generate them. Generally, they will be conditions for roots in the projective space with homogeneous (or homogenized) polynomials. For more details on different resultants, we refer to [24, 25, 23].

We choose to use the sparse resultant, since most polynomial systems encountered in practice are sparse in the supports. That means, for example, a polynomial of degree 5 in two variables x, y will not contain all 21 possible combinations of monomials $x^5, y^5, x^4y, xy^4, \dots, x^4, y^4, x^3y, \dots, y, x, 1$. Just as in linear algebra, this sparseness can be exploited for calculating the resultant. Another reason for using the sparse resultant is that it gives the necessary and sufficient conditions for toric roots, that is, roots in $(\mathbb{C}^*)^{n_{\hat{\mathbf{a}}}}$, such that the input polynomials need not be homogeneous (or homogenized),

as for other resultants. Finally, the sparse resultant enables us to work with Laurent polynomials, that is, polynomials with positive and negative integer exponents.

Usually, resultant algorithms set up a matrix in the coefficients of the system. The determinant of this matrix is then the resultant or a multiple of it. Generating the coefficient matrices and their determinants efficiently is a subject to ongoing research, but there are some useful algorithms freely available. An overview of different matrix constructions in elimination theory is given in [22]. In this work, we use the maple software package `multires` [26], which can be downloaded from the internet[1]. For more details on the theory of sparse resultants, we refer to [24, 22, 27, 28].

4.3. Finding invariant controlled variables for polynomial systems

We are now ready to apply these concepts to the problem of selecting controlled variables and self-optimizing control. As in the linear case above, we assume that the active constraints and the model equations, $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = 0$, and the measurement relations, $\mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0$, are satisfied. To obtain the $n_{\mathbf{c}} = n_{DOF}$ controlled variables needed for the unconstrained degrees of freedom we have:

Theorem 2 (Nonlinear measurement combinations as controlled variables)**.** *Given* $\hat{\mathbf{d}} \in (\mathbb{R}^*)^{n_{\hat{\mathbf{d}}}}$*, and* $n_{\mathbf{y}} + n_{\mathbf{g}} = n_{\hat{\mathbf{d}}}$*, independent relations* $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = \mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0$ *such that the system*

$$\begin{aligned} \mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) &= 0 \\ \mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) &= 0 \end{aligned} \tag{38}$$

has finitely many solutions for $\hat{\mathbf{d}} \in (\mathbb{C}^*)^{n_{\hat{\mathbf{d}}}}$*, and let* $J^{(i)}_{\mathbf{z},red}$ *denote the* i*-th element in the reduced gradient expression. Let* $\mathcal{R}(J^{(i)}_{\mathbf{z},red}, \mathbf{g}, \mathbf{m}), \quad i = 1 \dots n_{\mathbf{c}}$ *be the sparse resultants of the* $n_{\mathbf{c}}$ *polynomial systems composed of*

$$J^{(i)}_{\mathbf{z},red}(\mathbf{y}, \hat{\mathbf{d}}) = 0, \quad \mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = 0, \quad \mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0 \quad i = 1 \cdots n_{\mathbf{c}}. \tag{39}$$

Then controlling the active constraints, $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = 0$*, and the polynomial invariants* $c_i = \mathcal{R}(J^{(i)}_{\mathbf{z},red}, \mathbf{g}, \mathbf{m}) = 0, \quad i = 1, \dots, n_{\mathbf{c}}$*, yields optimal operation to first order throughout the region.*

[1] `http://www-sop.inria.fr/galaad/logiciels/multires`

Proof. The active constraints are controlled, thus $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = 0$ and $\mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0$ are satisfied. The system $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = 0, \mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0$ has only finitely many solutions for $\hat{\mathbf{d}}$, so the set of possible $\hat{\mathbf{d}}$ is fixed. Moreover, we know that a real solution $\hat{\mathbf{d}}$ to the subsystem $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) = \mathbf{m}(\mathbf{y}, \hat{\mathbf{d}}) = 0$ exists, since it is the given disturbance $\mathbf{d}$ and the actual state $\mathbf{x}$.

From Lemma 1, the sparse resultant gives the necessary and sufficient conditions for the existence of a solution $\hat{\mathbf{d}} \in (\mathbb{C}^*)^{n_{\hat{\mathbf{d}}}}$ for (39). Therefore, whenever $J^{(i)}_{z,red} = 0$, the resultant is zero (necessary condition). On the other hand, if $\mathcal{R}(J^{(i)}_{z,red}, \mathbf{g}, \mathbf{m}) = 0$ then (39) is satisfied (sufficient condition).

This holds for any solution $\hat{\mathbf{d}} \in (\mathbb{C}^*)^{n_{\hat{\mathbf{d}}}}$, and in particular the "actual" values of $\hat{\mathbf{d}}$. Because there are as many resultants as unconstrained degrees of freedom, controlling $\mathcal{R}(J^{(i)}_{z,red}, \mathbf{g}, \mathbf{m})$ for $i = 1, \ldots, n_{\mathbf{u}}$ satisfies the necessary conditions of optimality in the region. □

Remark 6. *In cases where the $\hat{\mathbf{d}} \notin (\mathbb{C}^*)^{n_{\hat{\mathbf{d}}}}$, we may apply a variable transformation to formulate the problem such we get $\hat{\mathbf{d}} \in (\mathbb{C}^*)^{n_{\hat{\mathbf{d}}}}$. For example a translation $d = \tilde{d} - 1$.*

Remark 7. *We assume that we have "well behaved" systems for each region. In particular it is assumed that there are no base points (values of $a_{i,j}(\mathbf{y})$, for which a polynomial in $\mathbf{g}$ or $\mathbf{m}$ vanishes for all values of $\hat{\mathbf{d}}$).*

Example 2 (Elimination)**.** *This simple example illustrates the computation of the resultant for the case with one disturbance d and no unmeasured states. Consider a system where we want to minimize a cost J subject to one constraint. Assume the reduced gradient is $J_{\mathbf{z},red} = \mathbf{N}^{\mathbf{T}} \nabla_{\mathbf{z}} J(y, d) = a_{1,1}(y) + a_{1,2}(y)d$, and the constraint is*

$$g(y, d) = a_{2,1}(y) + a_{2,2}(y)d + a_{2,3}(y)d^2 = 0, \tag{40}$$

where all coefficients $a_{i,j}(y)$ are known functions of the measurements. At the optimum we must have

$$J_{\mathbf{z},red} = a_{1,1}(y) + a_{1,2}(y)d = 0. \tag{41}$$

For arbitrary coefficients $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}, a_{2,3}$, this system of univariate polynomials in d does not have a common solution. However if the sparse resultant is zero, then there exist a common solution $d \neq 0$ for (40)-(41).

In the case of univariate polynomials, the sparse resultant coincides with the classical resultant, which is the determinant of the Sylvester matrix [18],

$$Syl = \begin{bmatrix} a_{1,2}(y) & a_{1,1}(y) & 0 \\ 0 & a_{1,2}(y) & a_{1,1}(y) \\ a_{2,3}(y) & a_{2,2}(y) & a_{2,1}(y) \end{bmatrix}. \tag{42}$$

The resultant is (where we omit writing explicitly the dependence on y)

$$\mathcal{R}\left(J_{\mathbf{z},red}, g(y,d)\right) = \det(Syl) = a_{1,2}^2 a_{2,1} - a_{1,2}a_{1,1}a_{2,2} + a_{2,3}a_{1,1}^2. \tag{43}$$

For a common root d^* to exist, the polynomial in the coefficients $\mathcal{R}(J_{\mathbf{z},red}, g(y,d))$ must vanish. Since the constraints are satisfied, $g(y,d) = 0$ for any disturbance $d \in \mathbb{R}$, controlling the resultant to zero is the condition for the reduced gradient $J_{\mathbf{z},red}$ to become zero. So for any real $d \neq 0$, the optimality conditions will be satisfied whenever $\mathcal{R}(J_{\mathbf{z},red}, g(y,d)) = 0$.

5. CSTR Case Study I

The purpose of this case study is to show on a small CSTR example, that the proposed polynomial method can give polynomial variable combinations that are suitable for practical implementation. Consider a CSTR (Figure 1), with a feed stream F [m^3/min] containing mainly component A, and two first-order chemical reactions,

$$\begin{aligned} A &\longrightarrow B \quad r_1 = k_1 c_A \\ B &\longrightarrow C \quad r_2 = k_2 c_B. \end{aligned} \tag{44}$$

Component B is the desired product, while C is an undesired side product. At steady state we have one degree of freedom, the feed stream $u = F$, which can be adjusted to achieve optimal operation. The operational objective is to maximize the production of component B, which for a given feed rate F corresponds to maximizing the concentration of B,

$$J = -c_B. \tag{45}$$

It is assumed that the unmeasured disturbances $\mathbf{d}$ are the rate constants k_1 and k_2, which could vary due to catalyst decay, but also imperfect temperature control in the reactor or unknown reaction mechanisms, which have

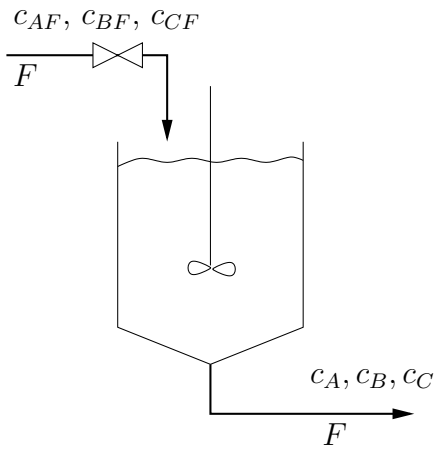

Figure 1: Isothermal CSTR (Case Study 1)

Table 3: Variables for Case Study 2

Symbol	Description	Type	Value	Unit
F	Feed flow rate	Known input u	Varying	m^3/min
c_A	Outlet concentration A	Measurement y	Varying	kmol/m^3
c_C	Outlet concentration C	”	”	kmol/m^3
V	Tank volume	Known parameter	Fixed	m^3
c_{AF}	Feed concentration A	”	”	kmol/m^3
c_{BF}	Feed concentration B	”	”	kmol/m^3
c_{CF}	Feed concentration C	”	”	kmol/m^3
c_B	Concentration of product	Unmeasured state x	Varying	kmol/m^3
k_1	Reaction constant 1	Disturbance d	Varying	1/min
k_2	Reaction constant 2	”	”	1/min

been approximated by first-order kinetics. In addition, we assume that the concentration c_B is too difficult (or expensive) to measure online.

All variables and known parameters are shown in Table 3. The task is to find a controlled variable $\mathbf{c}(\mathbf{y})$ which can be controlled using the total flow rate $u = F$, and which maximizes the desired concentration. We use the procedure from Table 1.

Step 1: Formulate the optimization problem. We collect the input $u = F$ and the states $\mathbf{x} = [c_A, c_B, c_C]^\mathbf{T}$ into a vector

$$\mathbf{z} = [F, c_A, c_B, c_C]^\mathbf{T}. \tag{46}$$

Then, the optimization problem is

$$\begin{aligned} \min_{\mathbf{z}} J &= -c_B \\ s.t. & \\ \mathbf{g}(\mathbf{z}) &= 0, \end{aligned} \tag{47}$$

where the model equations $\mathbf{g}(\mathbf{z}) = 0$ are derived from the mass balances,

$$\begin{aligned} g_1 &= Fc_{AF} - Fc_A - k_1 c_A V = 0 \\ g_2 &= Fc_{BF} - Fc_B + k_1 c_A V - k_2 c_B V = 0 \\ g_3 &= Fc_{CF} - Fc_C + k_2 c_B V = 0. \end{aligned} \tag{48}$$

Step 2: Find regions of active constraints. In our example, there are no other constraints than the model equations. Therefore we have only one region of active constraints, which is defined by (48). Since we have four variables and three constraints, the number of unconstrained degrees of freedom is

$$n_{DOF} = n_\mathbf{z} - n_\mathbf{g} = 4 - 3 = 1, \tag{49}$$

and thus the number of controlled variables which we want to find, is $n_c = n_{DOF} = 1$.

Step 3a: Formulate optimality conditions. Using $\mathbf{z} = [F, c_A, c_B, c_C]^\mathrm{T}$, the first-order optimality conditions are

$$\begin{aligned} \nabla_\mathbf{z} J(\mathbf{z}) + [\nabla_\mathbf{z} \mathbf{g}(\mathbf{z})]^\mathbf{T} \boldsymbol{\lambda} &= 0, \\ \mathbf{g}(\mathbf{z}) &= 0. \end{aligned} \tag{50}$$

Step 3b: Eliminate Lagrangian multipliers. We calculate the null-space of the constraint Jacobian $\mathbf{N} = [n_1, n_2, n_3, n_4]^{\mathbf{T}}$ with

$$\begin{aligned} n_1 &= -F(F + k_2V)(F + k_1V) && (51)\\ n_2 &= -(F + k_2V)F(c_{AF} - c_A) && (52)\\ n_3 &= F(-k_1Vc_{AF} + k_1c_AV - Fc_{BF} - c_{BF}k_1V + Fc_B + c_Bk_1V) && (53)\\ n_4 &= k_1[(-c_{BF} + c_B - c_{AF} + c_A - c_{CF} + c_C)V^2k_2 + V(-Fc_{CF} + Fc_C)] \\ &\quad +(Fc_B - Fc_{CF} + Fc_C - Fc_{BF})Vk_2 + F^2(c_C - c_{CF}). && (54) \end{aligned}$$

The reduced gradient for our system is defined as $J_{\mathbf{z},red} = [\mathbf{N}(\mathbf{z})]^{\mathbf{T}}\nabla_{\mathbf{z}}J(\mathbf{z}) = 0$. Using $\nabla_{\mathbf{z}}J(\mathbf{z}) = [0,\, 0,\, -1,\, 0]^{\mathbf{T}}$ we have that

$$\begin{aligned} J_{\mathbf{z},red} &= -n_3 \\ &= -F(-k_1Vc_{AF} + k_1c_AV - Fc_{BF} - c_{BF}k_1V + Fc_B + c_Bk_1V). \end{aligned} \tag{55}$$

Step 3c: Eliminating unknowns k_1, k_2 and c_B. We have three model equations $\mathbf{g}(\mathbf{z}) = 0$, (48), and three unknowns

$$\hat{\mathbf{d}} = \begin{bmatrix} c_B \\ k_1 \\ k_2 \end{bmatrix}, \tag{56}$$

which need to be eliminated from $J_{\mathbf{z},red}$. Before we apply Theorem 2, we check the assumptions:

1. Under normal operation (nonzero feed, etc.), when all other variables are given, $\mathbf{g}(\mathbf{z}) = 0$ has one solution for k_1, k_2, c_B (finite number of solutions).
2. Under normal operation we have that $k_1 \neq 0$, $k_2 \neq 0$ and $c_B \neq 0$. Therefore we have that $\mathbf{d} \in (\mathbb{C}^*)^3$.

Since all requirements are fulfilled, we can use the sparse resultant $\mathcal{R}(J_{\mathbf{z},red}, g_1, g_2, g_3)$ as controlled variable. We use the software `multires` [26] to calculate the sparse resultant and obtain for the controlled variable

$$c = \mathcal{R}(J_{\mathbf{z},red}, g_1, g_2, g_3) = c_{AF}c_A + c_{AF}c_{CF} - c_{AF}c_C - c_A^2. \tag{57}$$

This variable combination is simple and should be well suited for practical implementation.

Step 4: Control the invariant. Controlling the invariant such that

$$c = 0 \tag{58}$$

yields optimal operation.

Remark 8. *We note with interest that the self-optimizing invariant* (57) *is simpler than the expression for the reduced gradient* (55)*. This is good for implementation, in other cases, however, it may become more complicated. Generally it is difficult to make statements about the form of the invariant a-priori.*

6. Discussion: Changes in active constraints.

In this section we present a brief discussion on a method for detecting when a disturbance $\mathbf{d}$ causes changes in the active set. Since we have derived a set of controlled variables, which is equivalent to controlling the optimality conditions, the idea is to use these controlled variables for detecting changes in the active set.

This important topic has received some attention in literature, for example Baotić et al. [29] worked on linear systems with quadratic objectives, while [30] present an extremum seeking method, which can handle changing active constraints. Other references are [31, 32, 33].

From an optimization perspective, there is no difference between a constraint and a controlled variable $\mathbf{c}(\mathbf{y})$, as the controlled variable may be simply seen as an active constraint, and, similarly, an active constraint may be considered a variable which is controlled at its constant setpoint. From this perspective, there is no difference between an active constraint and the model equations, either.

However, from an implementation point of view, there are differences between the model, the active constraints and the controlled variables $\mathbf{c}(\mathbf{y})$. First of all, the active constraints and the controlled variables $\mathbf{c}(\mathbf{y}) = 0$ are not satisfied automatically, that is one must control them to their setpoints. Secondly, since their values are known (or calculated using known measurements) they may be used for detecting when to switch control structures. To do this, we make following main assumptions:

Assumption 1. *The regions are adjacent and only two regions share a boundary.*

Assumption 2. *The disturbance moves the system continuously from one region to another, and the system cannot jump over regions.*

Assumption 3. *Controlling the invariant* $\mathbf{c}(\mathbf{y}) = 0$ *and the constraint* $\mathbf{g} = 0$ *is equivalent to controlling the optimality conditions, and the system is operated optimally in the current region (*$\mathbf{c} = \mathbf{g} = 0$*). Moreover, we assume that controlling* $\mathbf{c} = \mathbf{g} = 0$ *minimizes the cost* J*.*

Assumption 4. *The optimality conditions of two neighbouring regions are simultaneously satisfied only on the interface between the regions.*

In most practical cases, only one constraint will become active or inactive at a time. However, it is also possible that several constraints become active or inactive simultaneously. Starting in the correct region, we use following rules to track the set of active constraints:

1. (One or more new constraints become active) When a new constraint is hit, change the control structure to the corresponding region.
2. (One or more constraints become inactive) As soon as the controlled variable $\mathbf{c}$ in one of the neighboring regions becomes zero (reaches its optimal setpoint), change the control structure to the corresponding region.

The reasoning behind the rules is that we start with an optimally operated system in region 1, and that controlling the $\mathbf{c}_1 = \mathbf{g}_1 = 0$ is equivalent to controlling the optimality conditions in region 1 (Assumption 3). A slowly varying (quasi-steady-state) disturbance will move the system gradually towards the boundary (Assumption 2). On the interface between two regions, the optimality conditions of both regions are satisfied. This is when the control structure is switched because either $\mathbf{c}_2$ of the new region will become zero, or a constraint of the new region $\mathbf{g}_2$ will become active (Assumption 4). As the disturbance moves the system further into the new region, the system stays optimal, because the optimality conditions of the new region, $\mathbf{c}_2 = 0$ and $\mathbf{g}_2 = 0$, are controlled. Thus, if the boundary is the only place where the optimality conditions of both regions are satisfied simultaneously (Assumption 4), and if only two region share a common boundary (Assumption 1), then we may use the controlled variables for determining when to switch regions.

Although the assumptions will not generally hold for all polynomial systems, in many practical cases the rules can be used to detect when the control structure should be switched. The probability of the system operating

at steady state on the boundary is zero (zero measure set), so this does not affect the controllability of the whole system [33].

Similar to our approach, [30] present a method which detects active set changes based on the optimality conditions. Their approach will be applicable in the same cases as our approach. However, there are significant differences between [30] and our approach. We separate the steady-state optimization problem and the dynamic control problem, by using self-optimizing controlled variables. Once the steady-state optimal regions of active constraints are known, and control structures are set up for each region, we start with designing the dynamic controllers and an appropriate switching law, which can handle the dynamic system and avoids e.g. switching back and forth for high frequency disturbances. In contrast, [30] aim at directly designing a dynamic (extremum seeking) controller, which can detect changing active constraints.

The main focus of this work is to find steady-state optimal controlled variables for different regions in the disturbance space. The actual dynamic implementation with switching control structures is beyond the scope of this paper and has to be studied separately.

7. CSTR Case Study II

The purpose of this case study, taken from [34], is to show how to find variable combinations for use as controlled variables in different regions of active constraints, and how to switch between the regions. In this case the resulting polynomials are probably too complicated for practical implementation. Nevertheless, we illustrate by dynamic simulation that a simple feedback control structure based on these variables gives optimal steady state operation.

We consider an isothermal CSTR with two parallel reactions, as depicted in Figure 2. The reactor is fed with two feed streams F_A and F_B which contain the reactants A and B in the concentrations c_A and c_B. In the main vessel, the two components react to the desired product C, and the undesired side product D. The reactants A and B are not consumed completely during the reaction, so the outflow contains all four products. The CSTR is operated isothermally, and we assume that perfect temperature control has been implemented.

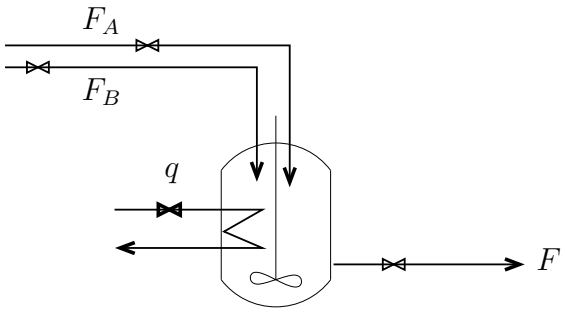

Figure 2: CSTR with two reactions (Case Study II)

The products C and D are formed by the reactions:

$$\begin{aligned} \text{A} + \text{B} &\xrightarrow{k_1} \text{C} \\ 2\,\text{B} &\xrightarrow{k_2} \text{D}. \end{aligned} \tag{59}$$

We wish to maximize the amount of desired product $(F_A+F_B)c_C$, weighted by a yield factor $(F_A + F_B)c_C/(F_A c_{A,in})$ [34]. The amount of removed heat and the maximum flow rate are limited by the equipment, and we formulate the mathematical optimization problem as follows [34]:

$$\max_{F_A,F_B} \frac{(F_A + F_B)c_C}{F_A c_{A_{in}}}(F_A + F_B)c_C \tag{60}$$

subject to

$$\begin{aligned} F_A c_{A_{in}} - (F_A + F_B)c_A - k_1 c_A c_B V &= 0 \\ F_B c_{B_{in}} - (F_A + F_B)c_B - k_1 c_A c_B V - 2k_2 c_B^2 V &= 0 \\ -(F_A + F_B)c_C + k_1 c_A c_B V &= 0 \\ F_A + F_B &\leq F_{max} \\ k_1 c_A c_B V(-\Delta H_1) + 2k_2 c_B^2 V(-\Delta H_2) &\leq q_{max}. \end{aligned} \tag{61}$$

Here, k_1 and k_2 are the rate constants for the two reactions, $(-\Delta H_1)$ and $(-\Delta H_2)$ are the reaction enthalpies, q_{max} the maximum allowed heat production, V the reactor volume, and F_{max} the maximum total flow rate. The measured variables ($\mathbf{y}$), the manipulated variables ($\mathbf{u}$), the disturbance variables ($\mathbf{d}$), and the internal states ($\mathbf{x}$) are given in Table 4, and the parameter values of the system are listed in Table 5.

Table 4: Overview of variables (Case Study II)

Symbol	Description	Comment
F_A	Inflow stream A	Measured input $\mathbf{u}$
F_B	Inflow stream B	”
F	total flow	Measured variable $\mathbf{y}$
q	Heat produced	”
c_B	Concentration of B	”
c_A	Concentration of A	Unmeasured state $\mathbf{x}$
c_C	Concentration of C	”
k_1	Rate constant reaction 1	Unmeasured disturbance d

Table 5: Parameters (Case Study II)

Symbol	Unit	Value
k_1	l/(mol h)	0.3 - 1.5
k_2	l/(mol h)	0.0014
$(-\Delta H_1)$	J/mol	7×10^4
$(-\Delta H_2)$	J/mol	5×10^4
$c_{A,in}$	mol/l	2
$c_{B,in}$	mol/l	1.5
V	l	500
$F_{\max}$	l/h	22
q_{max}	kJ/h	1000

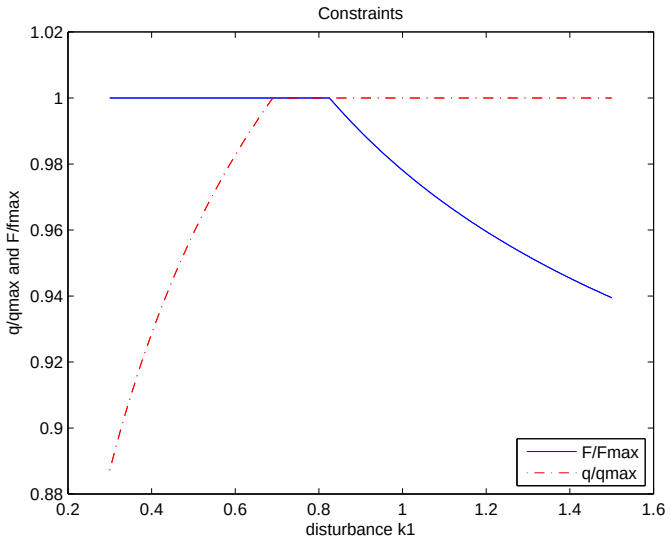

Figure 3: Optimal values of the constrained variables (Case Study II)

We write the combined vector of states $\mathbf{x} = [c_A,\, c_B,\, c_C]$ and manipulated variables $\mathbf{u} = [F_A,\, F_B]$ as

$$\mathbf{z} = \left[\; c_A, \quad c_B, \quad c_C, \quad F_A, \quad F_B \;\right]^{\mathbf{T}}. \tag{62}$$

7.1. Identifying operational regions

Next, the system is optimized off-line for the range of possible disturbances, which is assumed to be the single disturbance $d = k_1$. Based on which constraints are active, the system can be partitioned into three adjacent critical regions. The critical regions are visualized in Figure 3, where the normalized constraints are plotted over the disturbance range. In the first region, for disturbances below about $k_1 = 0.65\frac{\mathrm{l}}{\mathrm{mol\,h}}$, the flow constraint is the only active constraint. The second critical region, for values between about $k_1 = 0.65\frac{\mathrm{l}}{\mathrm{mol\,h}}$ and $k_1 = 0.8\frac{\mathrm{l}}{\mathrm{mol\,h}}$, is characterized by two active constraints, i. e. both the flow constraint and the heat constraint are active. Finally, in the third region, above about $k_1 = 0.8\frac{\mathrm{l}}{\mathrm{mol\,h}}$, only the heat constraint remains.

7.2. Eliminating $\boldsymbol{\lambda}$

In each critical region, the set of controlled variables contains the active constraints (we know that they should be controlled at the optimum). This leaves the unconstrained degrees of freedom, which is the difference between the number of manipulated variables and the active constraints, $n_{DOF} = n_{\mathbf{z}} - n_{\mathbf{g}}$. For each of the unconstrained degrees of freedom one controlled variable is needed.

In the first critical region this gives $n_{DOF,1} = 5 - 4 = 1$ unconstrained degrees of freedom, so apart from the active constraint, which is the first controlled variable, we need to control one more variable (invariant).

To obtain the reduced gradient, we calculate the null space of Jacobian of the active set $\mathbf{N}_{\mathbf{z}}^{\mathbf{T}}$ and multiply it with the gradient of the objective function $\nabla_{\mathbf{z}} J(\mathbf{z}, \mathbf{d})$ to obtain $J_{\mathbf{z},red,1} = \mathbf{N}_{\mathbf{z}}^{\mathbf{T}} \nabla_{\mathbf{z}} J$. Depending on the algorithm to compute the null space, this may become a fractional expression, but since we want to control the process at the optimum, i. e. we control $J_{\mathbf{z},red,1}$ to zero, it is sufficient to consider only the numerator of $J_{\mathbf{z},red,1}$. This is possible because a fraction vanishes if the numerator is zero (provided the denominator is nonzero which is the case here because $\nabla_{\mathbf{z}} \mathbf{g}$ has full rank). For the critical region 1, we obtain from (11) the reduced gradient

$$
\begin{aligned}
J_{\mathbf{z},red,1} = -(F_A + F_B)^2 c_C \big[& -3c_C F_B^2 F_A - 3c_C F_A^2 F_B \\
& - 4c_C c_B F_A^2 k_2 V - 4c_C k_2 V^2 k_1 c_B^2 F_A - c_C F_A^3 \\
& - c_C F_B^3 - 4c_C k_2 V^2 k_1 c_B^2 F_B - c_C c_B F_A^2 k_1 V \\
& - 4c_C c_B F_B^2 k_2 V - c_C c_B F_B^2 k_1 V - c_C F_A^2 c_A k_1 V \\
& - c_C F_B^2 c_A k_1 V - 8c_C F_A c_B F_B k_2 V \\
& - 2c_C F_A c_B F_B k_1 V - 2c_C F_A F_B c_A k_1 V \\
& + 8F_A k_1 V^2 c_{A,in} k_2 c_B^2 + 2F_A^2 k_1 V c_B c_{A,in} \\
& + 2F_A k_1 V F_B c_B c_{A,in} - 2F_A^2 k_1 V c_{B,in} c_A \\
& - 2F_A k_1 V F_B c_{B,in} c_A \big] ,
\end{aligned}
\tag{63}
$$

which should be controlled to zero. This expression may be simplified slightly, since it is known that $(F_A + F_B)^2 c_C \neq 0$. It is therefore sufficient to find an invariant which is equivalent to controlling the factor in square brackets in (63).

Similarly, in the second critical region $n_{DOF,2} = 5 - 5 = 0$, and here we simply control the active constraints, keeping q at $q_{\max}$ and F at $F_{\max}$.

In the third critical region $n_{DOF,3} = 5 - 4 = 1$, and we use one of the manipulated variables to control the active constraint ($q = q_{max}$) while the other one is used to control the invariant derived from $J_{\mathbf{z},red,3}$, which is an expression similar to (63).

7.3. Eliminating unknown variables

The reduced gradients for the first and the third critical region $J_{\mathbf{z},red,1}$ and $J_{\mathbf{z},red,3}$ still contain unknown variables, namely k_1, c_A and c_C, and cannot be used for feedback control directly. To arrive at variable combinations which can be used for control, we include all known variables into $\mathbf{y}$, and all unknown variables into $\hat{\mathbf{d}}$, such that $\hat{\mathbf{d}} = [k_1, c_A, c_C]^{\mathbf{T}}$. Then we write the necessary conditions for optimality for each region as

$$\begin{aligned} J_{\mathbf{z},red}(\mathbf{y}, \hat{\mathbf{d}}) &= 0 \\ \mathbf{g}(\mathbf{y}, \hat{\mathbf{d}}) &= 0. \end{aligned} \tag{64}$$

Considering the known variables $\mathbf{y}$ as parameters of the system, we want to find conditions on these parameters such that (64) is satisfied. The system has $n_{\hat{\mathbf{d}}} = 3$ unknown variables, k_1, c_A and c_C, of which we know that they are not zero. This corresponds to solutions $[k_1, c_A, c_C] \in (\mathbb{C}^*)^3$. According to Section 4 we have that (64) is satisfied if and only if the sparse resultant is zero.

For the first region, we use the sparse resultant of the system consisting of the invariant (63), the model equations (the first three equality constraints in (61)) and the first (active) inequality constraint in (61) to eliminate k_1, c_A, c_C and F_B and to calculate the controlled variable combination. The computations were performed using the `multires` software [26], and the controlled variable for region 1 is

$$\begin{aligned} c_1 = &-c_{b,in}^2 F_A^2 - F_A^2 c_{A,in} c_{b,in} + 6 F_A c_{A,in} k_2 c_b^2 V + 2 F_A c_{A,in} F_{max} c_b \\ &- F_A c_{A,in} F_{max} c_{b,in} + F_{max}^2 c_b^2 + c_{b,in}^2 F_{max}^2 + 4 V^2 k_2^2 c_b^4 \\ &- 2 c_{b,in} F_{max}^2 c_b - 4 V k_2 c_b^2 c_{b,in} F_{max} + 4 V k_2 c_b^3 F_{max}. \end{aligned} \tag{65}$$

Note that this invariant has become simpler than the reduced gradient (63).

In the second critical region control is simple; the two degrees of freedom are used to control the two active constraints $F = F_{max}$ and $q = q_{max}$.

The third critical region is controlled similar to the first one. One degree of freedom is used to control the active constraint, and the second degree

of freedom is used to control the resultant. The model equations (the first three equations together with the energy constraint) in (61) and the reduced gradient are used to compute the resultant. Thus, the unknown variables k_1, c_A, c_C, and F_B are eliminated from the reduced gradient. The controlled variable for region 3 is

$$\begin{aligned}
c_3 = &-4Vc_B^2k_2\Delta H_2F_Ac_{A,in}c_{B,in}q_{max}\Delta H_1 + F_Ac_{B,in}^2q_{max}^2\Delta H_1 \\
&+ 4V^2c_B^4k_2^2\Delta H_2F_Ac_{A,in}c_{B,in}\Delta H_1^2 - 4V^2c_B^4k_2^2\Delta H_2^2F_Ac_{A,in}c_{B,in}\Delta H_1 \\
&- 2Vc_B^2k_2F_Ac_{A,in}c_{B,in}\Delta H_1^2q_{max} - 4Vc_B^2k_2\Delta H_2F_Ac_{B,in}^2\Delta H_1q_{max} \\
&- 2Vc_B^2k_2\Delta H_2F_A^2c_{A,in}c_{B,in}^2\Delta H_1^2 + 8Vc_B^3k_2\Delta H_2\Delta H_1F_Ac_{A,in}q_{max} \\
&- 8V^2c_B^4k_2^2\Delta H_2c_{B,in}\Delta H_1q_{max} - 12V^2c_B^4k_2^2F_A\Delta H_2^2c_{B,in}^2\Delta H_1 \\
&- 8V^2c_B^5k_2^2\Delta H_2F_Ac_{A,in}\Delta H_1^2 + 8V^2c_B^5k_2^2\Delta H_2^2\Delta H_1F_Ac_{A,in} \\
&+ 8V^2c_B^5k_2^2F_A\Delta H_2^2c_{B,in}\Delta H_1 - q_{max}^3c_{B,in} + 2c_Bq_{max}^3 \\
&- \Delta H_1c_{B,in}F_Ac_{A,in}q_{max}^2 + 2c_BF_Ac_{A,in}q_{max}^2\Delta H_1 + F_A^2c_{A,in}c_{B,in}^2\Delta H_1^2q_{max} \\
&- 2c_BF_Ac_{B,in}q_{max}^2\Delta H_1 + 8Vc_B^3k_2\Delta H_2q_{max}^2 + 8V^2c_B^5k_2^2\Delta H_2^2q_{max} \\
&+ 8V^3c_B^6k_2^3\Delta H_2^3c_{B,in} - 2c_BF_A^2c_{A,in}c_{B,in}\Delta H_1^2q_{max} \\
&- 2Vc_B^2k_2\Delta H_1q_{max}^2c_{B,in} - 2Vc_B^2k_2\Delta H_2q_{max}^2c_{B,in} \\
&+ 4V^2c_B^4k_2^2\Delta H_2^2c_{B,in}q_{max} - 8V^3c_B^6k_2^3\Delta H_2^2c_{B,in}\Delta H_1.
\end{aligned} \tag{66}$$

Due to the structure of the polynomials in region 3, here the invariant has become more complicated after eliminating the unknown variables.

Although the expressions are quite complicated, they contain only known quantities, and can be simply evaluated and used for control. Before actually using the measurement combinations for control, they are scaled so that the order of magnitude is similar. That is, c_1 is scaled (divided) by 10, and c_3 is scaled by $\Delta H_1^2\Delta H_2F_AF_B$.

7.4. Using measurement invariants for control and region identification

Having established the controlled variables for the three critical regions, it remains to determine when to switch between the regions. Starting in the first critical region, the flow rate is controlled such that $F_A + F_B = F_{max}$, and the first measurement combination c_1 is controlled to zero. As the value of the disturbance k_1 rises, the reaction rate increases as well as the required cooling to keep the system isothermal, until maximum cooling is reached, Figure 4. When the constraint is reached, the control structure is switched

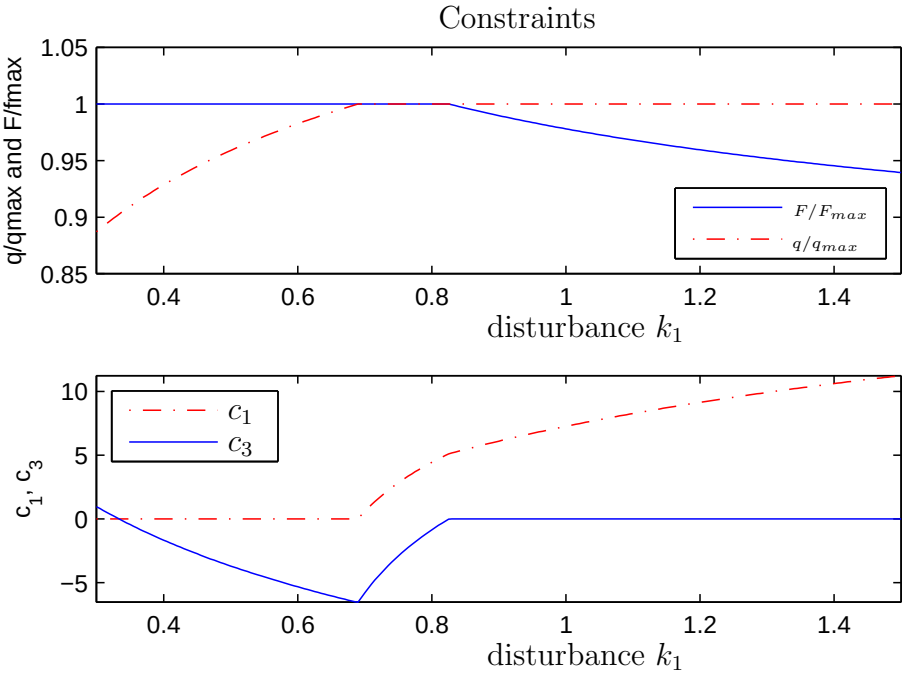

Figure 4: Optimal values of controlled variables (Case Study II)

to the next critical region, where the inputs are used to control $q = q_{max}$ and $F_A + F_B = F_{max}$. While operating in the second region, the controlled variables of the neighboring regions are monitored. As soon as one of the variables c_1 or c_3 reaches its optimal setpoint (i. e. 0) for its region the control structure is changed accordingly. Specifically, when k_1 is further increased, such that $c_3 = 0$ is reached, we must keep $F_A + F_B < F_{max}$ such to maintain the value $c_3 = 0$.

7.5. Implementation and dynamic simulations

In the steady-state case, we have assumed that we have ideal temperature and level control. In practice this has to be achieved by control, so we modified the model (61) such that the reactor holdup (level) and the temperature can vary dynamically. A detailed description of the dynamic model with its parameters is given in [35].

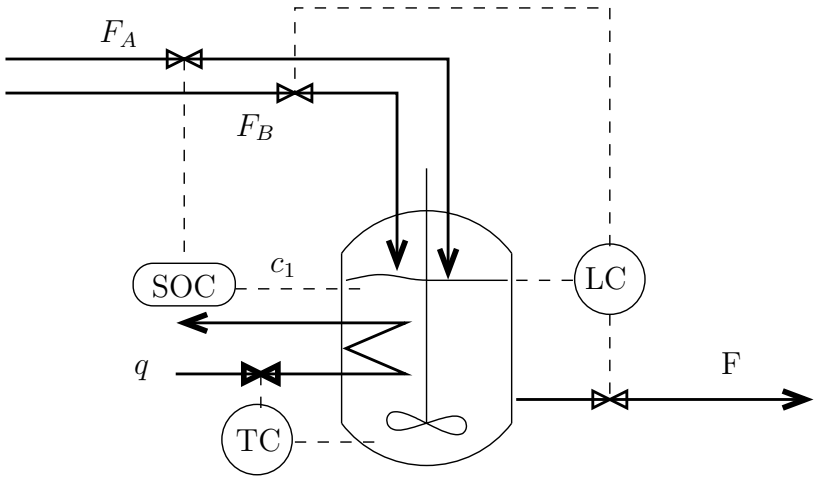

Figure 5: Variable pairings for Region 1 (Case Study II)

Control structure in region 1. All variables are controlled using PI controllers. The control structure in region 1 is presented in Figure 5. The cooling duty q is used to control the temperature and the feed flow F_A is used to control the invariant c_1. Further, we use the outflow F to control the level, and the feed F_B to control the throughput to $F = F_{max}$. Since we are controlling a constraint with a PI controller, we need some back-off in order not to become infeasible. We assume that this has already been taken into consideration when formulating the constraints, such that 0.1 l/h deviations from $F_{max} = 22$ l/h can be tolerated.

Control structure in region 2. In this region we simply keep $F = F_{max}$ and $q = q_{max}$. Using PI controllers, the temperature is controlled by manipulating the input F_A, and the level is controlled by F_B.

Control structure in region 3. All variables are controlled using PI controllers. The selected pairing is shown in Figure 6. The optimal value for cooling, $q = q_{max}$ is set in open loop. The feed flow F_A is used to control the invariant c_3, and the feed flow F_B is used to control the temperature. As in region 1, the outflow F is used to stabilize the level of the reactor.

In Figure 7 we show the dynamic behavior of the system. Starting in region 3, we control the heat constraint $q = q_{max}$ and the invariant c_3. The disturbance decreases stepwise until the flow constraint becomes active. When

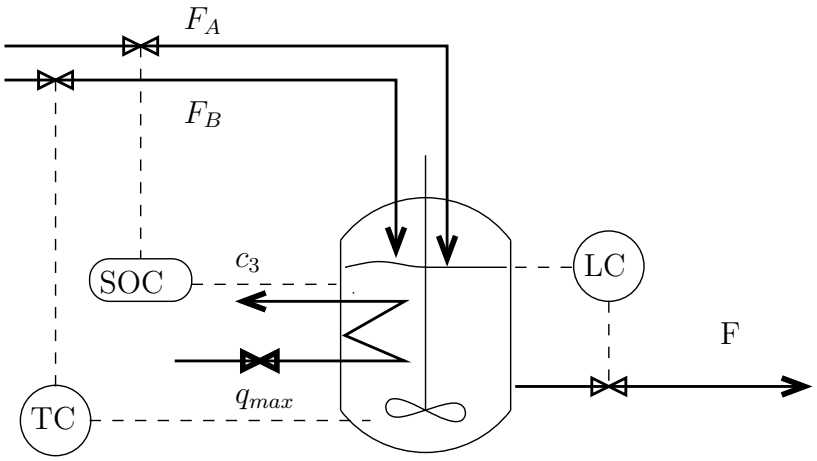

Figure 6: Variable Pairings for Region 3 (Case Study II)

the constraint is hit, we change the corresponding control structure to region 2, where we control F at F_{max} and q at q_{max}. We monitor the controlled variables of the neighboring regions, and as the disturbance decreases further, we switch the control structure to region 1 when the controlled variable c_1 becomes zero.

The simulations demonstrate nicely that it is possible to obtain optimal operation by controlling the invariants using simple PI controllers. Moreover we see that the control structure including active constraints can be changed based on monitoring the controlled variables and the constraints.

8. Discussion

Applicability. Since the sparse resultants can give "large" expressions, our method is best suited for small systems with not too many constraints and measurement equations. This is further emphasized by the fact that calculating the analytical determinant for large matrices is computationally demanding and that the construction of the resultant matrices is based on the computation of the mixed volume, which is a hard enumerative problem [23]. However, large systems can often be decomposed into smaller subsystems which can be considered (optimized) independently. In those cases our method can be applied to a subsystem.

Depending on the problem structure, the invariants may become more

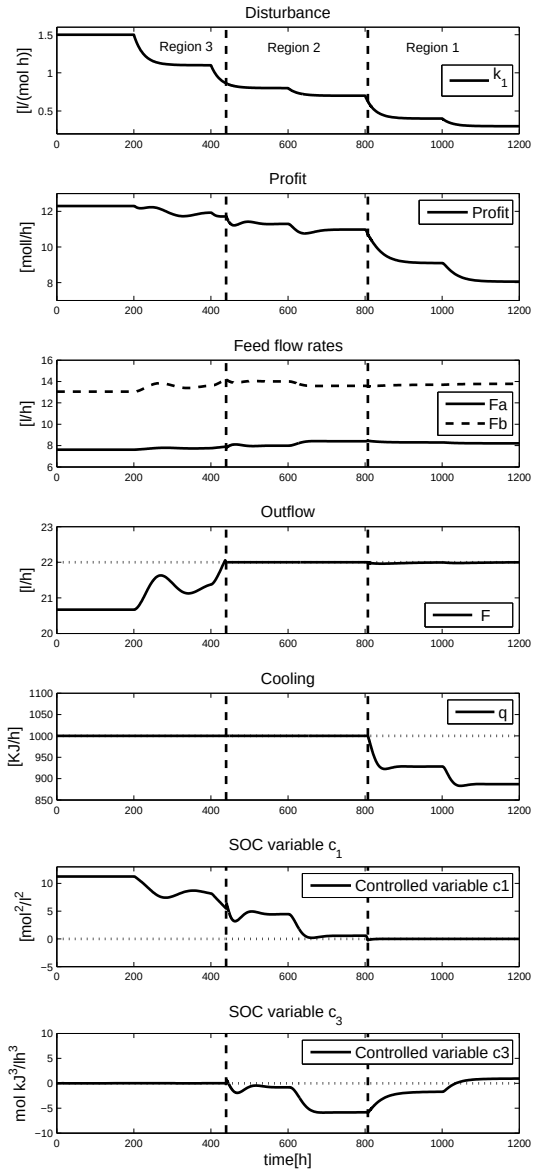

Figure 7: Case Study II: Starting in region 3 (left), where we control $c_3 = 0$ and $q = q_{max}$, the disturbance moves the system to region 2 (middle), where we control $F = F_{max}$ and $q = q_{max}$, and subsequently to region 1 (right) where we control $c_1 = 0$ and $F = F_{max}$.

complicated than the reduced gradient, as in region 3 of case study II, or much simpler, as in case study I and region 1 of case study II.

Alternative elimination of $\boldsymbol{\lambda}$. The elimination of the Lagrangian multipliers could also been done simultaneously with the other unknown variables using the resultant. Under the strict complementarity condition (λ_i is nonzero whenever the corresponding constraint is active), the solutions for $\boldsymbol{\lambda}$ lie in the toric variety, and therefore the sparse resultant gives necessary and sufficient conditions on the known variables so that the KKT system has a solution. We chose to apply the two-step procedure, where we first form the reduced gradient, and then eliminate the unknown variables using the resultant, because this results in lower computational load when computing the resultants.

Gröbner bases. As an alternative to using resultants, our initial approach was to compute the controlled variable combinations by Gröbner bases [18]. We calculated a Gröbner basis for the ideal generated by the optimality conditions using a suitable elimination ordering. Then we used a polynomial from the elimination ideal as controlled variable. However, in this approach it is not straightforward to find an ordering that eliminates the unknown variables while not yielding the "trivial solution" (i. e. the invariant is always zero when the constraints are satisfied). Another disadvantage with this Gröbner basis approach is that the selected invariant may give rise to additional "artificial solutions" which are not solutions of the original optimality conditions.

A similar approach is to calculate a Göbner basis for the ideal generated by the active constraints $\mathbf{g}(\mathbf{y}, \hat{\mathbf{d}})$ and $\mathbf{m}(\mathbf{y}, \hat{\mathbf{d}})$ using some elimination ordering, and to reduce $\mathbf{N^T}\nabla_{\mathbf{z}} J$ modulo the ideal. This avoids the trivial solution, however, the problem of choosing a monomial ordering which eliminates all unknown variables, remains. Generally the Gröbner basis approach tends to give even more complicated expressions than the sparse resultant approach presented here.

Number of equations/measurements. If there are more polynomial equations than unknowns, the engineer must choose a set of $n_{\hat{\mathbf{d}}}$ polynomials to use for eliminating the unknowns. For different sets of polynomials, however, the controlled variables will look quite different. The best (in terms of simplicity) choice depends on the structure of the equations, and is thus specific to the

problem. However, as a general guideline, we would advise to keep the degrees of the polynomials low.

Although we can specify which $n_{\hat{\mathbf{d}}}$ variables must be eliminated from the reduced gradient, the number of variables which remain depends strongly on the structure of the model equations and the eliminated variables. In some cases all information about the optimum is contained in very few variables, in other cases many variables are needed to specify the optimum.

In the case where $n_{\mathbf{y}} + n_{\mathbf{g}} \leq n_{\hat{\mathbf{d}}}$, that is, fewer equations than unknown variables, it is generally not possible to find invariants, which are equivalent to controlling the optimality conditions. Instead, we need to find the best possible approximation in terms of the cost. For linear systems with quadratic cost, the method in [8] can be used. How this can be generalized to polynomial and nonlinear systems is still an open problem.

Noise, plant-model mismatch. The resultant method, as presented in this paper, does not take into account measurement noise or model error. One possible approach to compensate for plant-model mismatch could be to use an experimental method such as NCO tracking [12] to adjust the setpoints or other parameters in the invariants. However, this is beyond the scope of this work; our goal was to generalize the null-space method [6] and to demonstrate that the concept of finding variables which remain constant at optimal operation is possible also for polynomial systems.

Controllability. Since our approach separates the controlled variables selection procedure and the controller design, it must be verified that the measurement invariants are controllable using simple controllers. That is, the controlled variables $\mathbf{c}$ must cross zero. Proving this is not trivial, and beyond the scope of this paper. However, in our experience so far, the measurement invariants could be controlled by simple controllers.

Relationship to NCO tracking. The presented method is based on the same idea as NCO tracking [12, 34]. However, in contrast to [12] and [34], where the optimality conditions are solved for the optimizing *inputs*, this work focuses on finding the right *outputs* which express the optimality conditions. The problem of generating the inputs which control the outputs to zero is dealt with separately. In most cases, inputs can be generated by feedback control, e.g. PI controllers.

9. Conclusions

Previously, the concept of self-optimizing control was only treated in the framework of linear models with quadratic cost functions. This paper contains the first contribution to extend the ideas to the more general class of polynomial systems. Although further work should be dedicated to handling the effect of noisy measurements and finding ways to approximate the invariant using simple and robust measurement polynomials, the polynomials obtained by the sparse resultant can become quite simple for some systems, and then they may also be implemented in real processes.

[1] T. E. Marlin, A. Hrymak, Real-time operations optimization of continuous processes, in: Proceedings of CPC V, AIChE Symposium Series vol. 93., 1997, pp. 156–164.

[2] M. Morari, G. Stephanopoulos, Y. Arkun, Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the classification of the control task. Analysis of the optimizing control structures, AIChE Journal 26 (2) (1980) 220–232.

[3] J. Jäschke, S. Skogestad, Self-optimizing control and NCO tracking in the context or real-time optimization, Accepted for publication in Journal of Process control.

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

[5] I. J. Halvorsen, S. Skogestad, J. C. Morud, V. Alstad, Optimal selection of controlled variables, Industrial & Engineering Chemistry Research 42 (14) (2003) 3273–3284.

[6] V. Alstad, S. Skogestad, Null space method for selecting optimal measurement combinations as controlled variables, Industrial & Engineering Chemistry Research 46 (2007) 846–853.

[7] V. Kariwala, Y. Cao, S. Janardhanan, Local self-optimizing control with average loss minimization, Industrial & Engineering Chemistry Research 47 (2008) 1150–1158.

[8] V. Alstad, S. Skogestad, E. S. Hori, Optimal measurement combinations as controlled variables, Journal of Process Control 19 (1) (2009) 138–148.

[9] S. Heldt, On a new approach for self-optimizing control structure design, in: ADCHEM 2009, July 12-15, Istanbul, 2009, pp. 807–811.

[10] I. J. Halvorsen, S. Skogestad, Indirect on-line optimization through set-point control, AIChE 1997 Annual Meeting, Los Angeles; paper 194h.

[11] Y. Cao, Self-optimizing control structure selection via differentiation, in: Proceedings of the European Control Conference, 2003, pp. 445–453.

[12] G. François, B. Srinivasan, D. Bonvin, Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty, Journal of Process Control 15 (6) (2005) 701 – 712.

[13] B. Chachuat, B. Srinivasan, D. Bonvin, Adaptation strategies for real-time optimization, Computers & Chemical Engineering 33 (10) (2009) 1557 – 1567.

[14] J. Nocedal, S. Wright, Numerical Optimization, Springer, 2006.

[15] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, Nonlinear Programming: Theory and Algorithms, John Wiley & Sons, 2006.

[16] E. Pistikopoulos, M. Georgiadis, V. Dua, Multiparametric programming, Wiley-VCH, 2007.

[17] M. Reid, Undergraduate Algebraic Geometry, Cambridge University Press, 1988.

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

[19] M. Coste, An Introduction to Semialgebraic Geometry, RAAG network school, 2002.

[20] V. Alstad, Studies on selection of controlled variables, Ph.D. thesis, Norwegian University of Science and Technology, Department of Chemical Engineering (2005).

[21] N. H. Abel, Beweis der Unmöglichkeit, algebraische Gleichungen von höheren Graden als dem vierten allgemein aufzulösen, Journal für die reine und angewandte Mathematik (1826) 65–84.

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

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

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

[25] B. Sturmfels, On the newton polytope of the resultant, Journal of Algebraic Combinatorics 3 (1994) 207–236.

[26] L. Busé, B. Mourrain, Using the maple multires package (2003).
URL `http://www-sop.inria.fr/galaad/software/multires`

[27] B. Sturmfels, Solving systems of polynomial equations (2002).

[28] A. Dickenstein, I. Z. Emiris, Solving Systems of Polynomial Equations, Springer Berlin Heidelberg, 2005.

[29] M. Baotić, F. Borrelli, A. Bemporad, M. Morari, Efficient on-line computation of constrained optimal control, SIAM Journal on Control and Optimization 47 (2008) 2470–2489.

[30] L. Woodward, M. Perrier, B. Srinivasan, Real-time optimization using a jamming-free switching logic for gradient projection on active constraints, Computers & Chemical Engineering 34 (11) (2010) 1863 – 1872.

[31] M. G. Jacobsen, S. Skogestad, Active constraint regions for optimal operation of chemical processes - Application to a reactor-separator-recycle system, submitted to Industrial & Engineering Chemistry Research (2011).

[32] H. Manum, Simple implementation of optimal control for process systems, Ph.D. thesis, Norwegian University of Science and Technology (2010).

[33] V. Lersbamrungsuk, T. Srinophakun, S. Narasimhan, S. Skogestad, Control structure design for optimal operation of heat exchanger networks, AIChE Journal 54 (1) (2008) 150–162.

[34] B. Srinivasan, L. T. Biegler, D. Bonvin, Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function, Computers & Chemical Engineering 32 (2008) 572–279.

[35] J. Jäschke, Invariants for optimal operation of process systems, Ph.D. thesis, Norwegian University of Science and Technology (2011).