Chapter 1
Measurement Polynomials as Controlled Variables

Johannes Jäschke, Sigurd Skogestad

Abstract In this chapter we present a method for finding controlled variables, which are nonlinear combinations of measurements. The procedure extends the concept of the null-space method (Alstad and Skogestad, 2007) to processes described by polynomial equations. The method consists of three main steps. First, the active constraints are determined. If the disturbance causes the set of active constraints to change, regions of constant active constraints are defined in the disturbance space. Second, optimally invariant variable combinations are determined for the remaining unconstrained degrees of freedom in each region. Third, unknown internal variables (states) and disturbances are eliminated to obtain new invariant variable combinations containing only known variables (measurements). Furthermore we show that if the 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 a case study, a four component isothermal CSTR.

1.1 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) (Marlin and Hrymak, 1997). The real-time optimizer generally uses a nonlinear steady state model in

Johannes Jäschke
Department of Chemical Engineering, NTNU Trondheim, Norway, e-mail: jaschke@chemeng.ntnu.no

Sigurd Skogestad
Department of Chemical Engineering, NTNU Trondheim, Norway, e-mail: skoge@chemeng.ntnu.no

order to recompute new optimal setpoints for the controlled variables in the control layer below. This 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.

A second approach for 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 Morari et al (1980), who write that we "want to find a function c of the process variables which when held constant leads automatically to the optimal adjustments of the manipulated variables", but 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, or may even make it unnecessary (Jäschke and Skogestad, 2010).

The term "self-optimizing control" was defined in the context of controlled variable selection with the purpose of describing the practical goal of finding "smart" controlled variables $\boldsymbol{c}$. Skogestad (2000) writes:

> 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).

Many industrial processes are operated using self-optimizing control, although it is not always called that. For example, optimally active constraints may be viewed as self-optimizing variables, e.g. maximum cooling of an air stream before entering a compressor. However, the more difficult problem is to identify self-optimizing control variables associated with unconstrained degrees of freedom. In most cases, engineering insight and experience leads to the choice of self-optimizing controlled variables, and the optimization problem is not formulated explicitly. An example for the unconstrained case is is controlling the air/fuel ratio entering a combustion engine at a constant value.

It has been noted previously (Halvorsen and Skogestad, 1997; Bonvin et al, 2001; Cao, 2003; Halvorsen et al, 2003), that the gradient of the cost function with respect to the degrees of freedom $\boldsymbol{u}$ would be the ideal controlled variable, $\boldsymbol{c} = J_{\boldsymbol{u}}$. However, the gradient $J_{\boldsymbol{u}}$ is usually not directly measurable, and analytical expressions for the gradient generally contain unknown disturbances. Therefore, the methods in self-optimizing control theory can be thought of as an approximation (in some "best" way) of the gradient using a measurement model.

In the last decade, several contributions have been made on the systematic search of controlled variables which have self-optimizing properties (Halvorsen et al, 2003; Alstad and Skogestad, 2007; Kariwala et al, 2008; Alstad et al, 2009; Heldt, 2009), but to the authors knowledge, self-optimizing control has only been considered locally, that is, using linear process models and a quadratic approximation of the cost function. This results in linear

measurement combinations $\boldsymbol{c} = \boldsymbol{H}\boldsymbol{y}$ as controlled variables. In cases where a strong curvature is present at the optimum, the loss imposed by using linear measurement combinations may not be acceptable any more, and the controlled variables are not self-optimizing.

The main contribution of this work is to extend the ideas of self-optimizing control, in particular the concept of the null-space method (Alstad and Skogestad, 2007) to constrained systems described by multivariable polynomials. This results in controlled variables which are polynomials in the measurements, $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{y})$.

Second, we show that under some assumptions, the controlled variables can be used to determine when the set of active constraints changes and which set to change to.

1.2 Overview

The procedure for achieving optimal operation is summarized in Fig. 1.1. 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, the optimality conditions are formulated, and the Lagrangian multipliers are eliminated. Then the unknown variables, i.e. the disturbances and the internal state variables are eliminated from the optimality conditions to obtain an invariant variable combination 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.

Finally, we monitor the active constraints and the invariants of the neighbouring regions to determine when to switch to a new region.

1.3 Achieving Optimal Operation using Measurement Invariants

1.3.1 Problem Formulation

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

$$\min_{\boldsymbol{u},\boldsymbol{x}} J(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{d}) \quad \text{s.t} \quad \begin{cases} g(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{d}) = 0 \\ h(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{d}) \leq 0 \end{cases} \tag{1.1}$$

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 $\mathcal{A}_i$
   - i Formulate optimality conditions
   - ii Eliminate:
     - a Lagrangian multipliers λ from optimality conditions to obtain invariants $J_{\boldsymbol{z},red}$ (reduced gradient)
     - b Unknown variables to obtain measurement invariants $\boldsymbol{c}(\boldsymbol{y})$, such that

$$\boldsymbol{c}(\boldsymbol{y}) = 0 \Longleftrightarrow J_{\boldsymbol{z},red} = 0$$

   - iii In each region $\mathcal{A}_i$, control the
     - a Active constraints
     - b Invariants $\boldsymbol{c}(\boldsymbol{y})$
4. Use controlled variables and measured constraints for changing regions

Fig. 1.1 Procedure for finding nonlinear invariants as controlled variables

Here $\boldsymbol{u}$, $\boldsymbol{x}$, $\boldsymbol{d}$ denote the manipulated input variables, the internal state variables, and the unmeasured disturbance variables, respectively. We assume that, in addition, we have measurements $\boldsymbol{y}$ satisfying the relation, $m(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{d}, \boldsymbol{y}) = 0$, which provide information about internal states, inputs, and disturbances. To simplify notation, we combine state and input variables in a vector $\boldsymbol{z} = [\boldsymbol{u}\,, \boldsymbol{x}]^{\boldsymbol{T}}$.

This is the same problem which is solved online at given sample times when using RTO. In this work, however, we do not wish to solve the optimization problem online, instead, we analyse the problem using offline calculations, in order to find good controlled variables which yield optimal operation when controlled at their setpoints.

1.3.1.1 Optimality Conditions

Let $\boldsymbol{z}^*$ be a feasible point of the optimization problem (1.1), and assume that all gradient vectors $\nabla_{\boldsymbol{z}} g_i(\boldsymbol{z}^*, \boldsymbol{d})$ and $\nabla_{\boldsymbol{z}} h_i(\boldsymbol{z}^*, \boldsymbol{d})$ associated with $g_i(\boldsymbol{z}^*, \boldsymbol{d}) = 0$ and the active constraints, $h_i(\boldsymbol{z}^*, \boldsymbol{d}) = 0$, are linearly independent. Then $\boldsymbol{z}^*$ is locally optimal if there exist Lagrangian multiplier vectors λ and ν, such that the following conditions, known as the KKT conditions are satisfied (Nocedal and Wright, 2006):

$$\begin{aligned} \nabla_{\boldsymbol{z}} J(\boldsymbol{z}^*, \boldsymbol{d}) + [\nabla_{\boldsymbol{z}} g(\boldsymbol{z}^*, \boldsymbol{d})]^{\boldsymbol{T}} \lambda + [\nabla_{\boldsymbol{z}} h(\boldsymbol{z}^*, \boldsymbol{d})]^{\boldsymbol{T}} \nu &= 0 \\ g(\boldsymbol{z}^*, \boldsymbol{d}) &= 0 \\ h(\boldsymbol{z}^*, \boldsymbol{d}) &\leq 0 \\ [h(\boldsymbol{z}^*, \boldsymbol{d})] \nu^{\boldsymbol{T}} &= 0 \\ \lambda, \nu &\leq 0 \end{aligned} \tag{1.2}$$

The condition that the Jacobian of the active constraints has independent rows (has full rank) is called the linear independence constraint qualification (LICQ) and guarantees that the Lagrangian multipliers λ and ν are uniquely defined at the optimum $\boldsymbol{z}^*$.

When optimizing nonlinear systems, such as polynomial systems, there are several complications which may arise. The optimality conditions, (1.2), 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 (1.2) 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). So before controlling $\boldsymbol{c}(\boldsymbol{y})$ to zero, it has to be assured that the process actually is at the desired optimum.

This 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 $\boldsymbol{c}(\boldsymbol{y})$ which is zero for all points that satisfy the KKT conditions, and which is nonzero whenever the KKT conditions are not satisfied.

1.3.2 Partitioning into Sets of Active Constraints

Generally, the set of inequality constraints $h_i(\boldsymbol{z}, \boldsymbol{d}) \leq 0$ that are active varies with the value of the elements in $\boldsymbol{d}$. 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 (1.1) into a sequence of equality constrained optimization problems, which are valid in the corresponding critical region. This idea is also applied in multi-parametric programming (Pistikopoulos et al, 2007). However, we do not search for an explicit expression for the inputs $\boldsymbol{u}^*$, as in multi-parametric programming. We rather use each subproblem to find good controlled variables $\boldsymbol{c}$ 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 has to be controlled for each unconstrained degree of freedom, $n_c = n_{DOF}$.

The number unconstrained degrees of freedom, n_{DOF} is calculated according to

$$n_{DOF} = n_z - n_g - n_{h,active} \tag{1.3}$$

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

In the rest of the paper, by abuse of notation, all active constraints $h_i(\boldsymbol{z}, \boldsymbol{d}) = 0$ are included in the equality constraint vector $g(\boldsymbol{z}, \boldsymbol{d}) = 0$. Then in every critical region, the optimization problem (1.1) can be written as:

$$\begin{aligned} \min\, & J(\boldsymbol{z}, \boldsymbol{d}) \\ \text{s.t.} \quad & \\ & g(\boldsymbol{z}, \boldsymbol{d}) = 0 \end{aligned} \tag{1.4}$$

The KKT first order optimality conditions, (1.2), simplify for problem (1.4) in each critical region, to

$$\begin{aligned} \nabla_{\boldsymbol{z}} J(\boldsymbol{z}, \boldsymbol{d}) + [\nabla_{\boldsymbol{z}} g(\boldsymbol{z}, \boldsymbol{d})]^{\boldsymbol{T}} \lambda = 0, \\ g(\boldsymbol{z}, \boldsymbol{d}) = 0. \end{aligned} \tag{1.5}$$

These expressions cannot be used for control yet, because they still contain unknown variables, in particular $\boldsymbol{x}$ (in $\boldsymbol{z} = [\boldsymbol{u}, \boldsymbol{x}]$), $\boldsymbol{d}$, and the Lagrangian multipliers λ, which have to be eliminated.

1.3.3 Eliminating the Lagrangian Multipliers $\boldsymbol{\lambda}$

We consider one equality constrained sub-problem (1.4) at a time. Every control structure that gives optimal operation has to satisfy (1.5). Recall that the LICQ holds, i.e. $\nabla_{\boldsymbol{z}} g(\boldsymbol{z}, \boldsymbol{d})$ has full row rank for every value of $\boldsymbol{d}$ within the critical region. In addition, we assume that we have strict complementarity (either the constraint is active, or the corresponding value in λ is zero).

Proposition 1.1. *Let $\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d}) \in \mathbb{R}^{(n_z - n_g) \times n_g}$ be a basis for the null space of $\nabla_{\boldsymbol{z}} g(\boldsymbol{z}, \boldsymbol{d})$. Controlling the active constraints $g(\boldsymbol{z}, \boldsymbol{d}) = 0$, and the variable combination $[\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d})]^{\boldsymbol{T}} \nabla_{\boldsymbol{z}} J(\boldsymbol{z}, \boldsymbol{d}) = 0$ results in optimal steady state operation.*

Proof. Select $\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d})$ such that $\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d}) \nabla_{\boldsymbol{z}} g(\boldsymbol{z}, \boldsymbol{d}) = 0$. Since the LICQ are satisfied, the constraint Jacobian $\nabla_{\boldsymbol{z}} g(\boldsymbol{z}, \boldsymbol{d})$ has full row rank and $\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d})$ is well defined and does not change dimension within the region. The first equation in (1.5) is premultiplied by $[\boldsymbol{N}(\boldsymbol{z}, \boldsymbol{d})]^{\boldsymbol{T}}$ to get:

$$[\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})]^{\boldsymbol{T}}\left(\nabla_{\boldsymbol{z}}J(\boldsymbol{z},\boldsymbol{d})+[\nabla_{\boldsymbol{z}}g(\boldsymbol{z},\boldsymbol{d})]^{\boldsymbol{T}}\lambda\right)=[\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})]^{\boldsymbol{T}}\nabla_{\boldsymbol{z}}J(\boldsymbol{z}.\boldsymbol{d})+\underline{0}\lambda$$
$$=[\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})]^{\boldsymbol{T}}\nabla_{\boldsymbol{z}}J(\boldsymbol{z},\boldsymbol{d}) \tag{1.6}$$

Since $\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})$ and has full rank, we have that (1.5) are satisfied whenever $g(\boldsymbol{z},\boldsymbol{d})=0$ and $\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})\nabla_{\boldsymbol{z}}J(\boldsymbol{z},\boldsymbol{d})=0$.

We introduce $J_{\boldsymbol{z},red}=\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})\nabla_{\boldsymbol{z}}J(\boldsymbol{z},\boldsymbol{d})$, and call $J_{\boldsymbol{z},red}$ the reduced gradient. By construction, the reduced gradient has $n_{DOF}=n_z-n_g$ elements. Controlling

$$J_{\boldsymbol{z},red}=[\boldsymbol{N}(\boldsymbol{z},\boldsymbol{d})]^{\boldsymbol{T}}\nabla_{\boldsymbol{z}}J(\boldsymbol{z},\boldsymbol{d})=0 \tag{1.7}$$

together with the active constraints, $g(\boldsymbol{z},\boldsymbol{d})=0$, fully specifies the system at the optimum and is equivalent to controlling the first order optimality conditions (1.5). However, $J_{\boldsymbol{z},red}$ cannot be used for control directly because it still depends on unknown variables, $\boldsymbol{d}$ and $\boldsymbol{x}$ ($\boldsymbol{x}$ enters through $\boldsymbol{z}=[\boldsymbol{u}\,,\boldsymbol{x}]^{\boldsymbol{T}}$). For this purpose, the disturbance and the internal states have to be eliminated from the expression (1.7).

A first naive approach would be to solve the measurement model equations $m(\boldsymbol{x},\boldsymbol{u},\boldsymbol{d},\boldsymbol{y})=0$ and the active constraints $g(\boldsymbol{z},\boldsymbol{d})=0$ for the unknowns $\boldsymbol{d}$ and $\boldsymbol{x}$, and substitute the solution into $J_{\boldsymbol{z},red}$. As we show, this is straightforward in case of linear equations, but it becomes significantly more complicated when working with polynomials of higher degree.

1.4 Elimination for Systems of Linear Equations (Zero Loss Method)

In this section we describe the basic concept of how the unknowns are eliminated form $J_{\boldsymbol{z},red}$. Our procedure is demonstrated step by step for minimizing a quadratic cost function subject to linear constraints and a linear measurement model. Solving the model and measurement equations for the unknowns and substituting into $J_{\boldsymbol{z},red}$ is avoided, because this is difficult to extend to the polynomial case. Instead, we search for necessary and sufficient conditions which guarantee that the measurement model $m(\boldsymbol{x},\boldsymbol{u},\boldsymbol{d},\boldsymbol{y})=0$, the active constraints and the model $g(\boldsymbol{z},\boldsymbol{d})=0$, and the reduced gradient $J_{\boldsymbol{z},red}=0$ are are satisfied at the same time. We require that the necessary and sufficient condition is a function of measurements $\boldsymbol{y}$ and known parameters, only.

The optimization problem we consider is

$$\begin{gathered}J(\boldsymbol{z},\boldsymbol{d})=\min\left[\,\boldsymbol{z}^{\boldsymbol{T}}\;\boldsymbol{d}^{\boldsymbol{T}}\,\right]\begin{bmatrix}\boldsymbol{J_{zz}} & \boldsymbol{J_{zd}}\\ \boldsymbol{J_{zd}}^{\boldsymbol{T}} & \boldsymbol{J_{dd}}\end{bmatrix}\begin{bmatrix}\boldsymbol{z}\\ \boldsymbol{d}\end{bmatrix}\\ \text{s.t.}\\ \boldsymbol{A}\boldsymbol{z}-\boldsymbol{b}=0,\end{gathered} \tag{1.8}$$

and the linear measurement model is

$$\begin{aligned} m(\boldsymbol{z},\boldsymbol{d},\boldsymbol{y}) &= \boldsymbol{y} - [\boldsymbol{G}^y\ \boldsymbol{G}_d^y] \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{d} \end{bmatrix} = 0 \\ &= \boldsymbol{y} - \tilde{\boldsymbol{G}}^y \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{d} \end{bmatrix} = 0. \end{aligned} \tag{1.9}$$

We consider $[\boldsymbol{z}\,,\boldsymbol{d}]^{\boldsymbol{T}}$ as unknown and we assume that (1.8) has a solution, $\boldsymbol{J_{zz}}$ is positive definite, and $\boldsymbol{A}$ has full row rank. In addition, we assume that the measurements are linearly independent, and $\tilde{\boldsymbol{G}}^y = [\boldsymbol{G}^y\ \ \boldsymbol{G}_d^y]$ invertible.

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

$$J_{\boldsymbol{z},red} = \boldsymbol{N^T}\nabla_{\boldsymbol{z}} J(\boldsymbol{z},\boldsymbol{d}) = \boldsymbol{N^T} \begin{bmatrix} \boldsymbol{J_{zz}}\ \boldsymbol{J_{zd}} \end{bmatrix} \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{d} \end{bmatrix} = 0. \tag{1.10}$$

If the number of independent measurements (n_y) is equal to the number of unknown variables variables ($n_z + n_d$), the measurement relations (1.9) can be solved for the unknowns and substituted into the gradient expression (1.10) to obtain

$$\boldsymbol{c}(\boldsymbol{y}) = \boldsymbol{N^T} \begin{bmatrix} \boldsymbol{J_{zz}}\ \boldsymbol{J_{zd}} \end{bmatrix} \left[\tilde{\boldsymbol{G}}^y\right]^{-1} \boldsymbol{y}. \tag{1.11}$$

Controlling $\boldsymbol{c}(\boldsymbol{y}) =$ and $\boldsymbol{Az} - \boldsymbol{b}$ to zero, then results in optimal operation.

In the case of polynomial equations of higher degrees, it is generally not possible to solve for the unknown variables. Therefore, we consider the problem from a slightly different perspective. Suppose $n_y = n_z + n_d$, then for any output and disturbance pair $(\boldsymbol{y},\boldsymbol{d})$, there exist a unique $\boldsymbol{z}$, which satisfies the measurement equations (1.9). However, an arbitrary pair ($\boldsymbol{y}$,$\boldsymbol{d}$) will fail to satisfy the first order optimality condition (1.10). More accurately, since $\boldsymbol{J_{zz}} > 0$, only one pair $(\boldsymbol{y},\boldsymbol{d})$ satisfies the first order optimality conditions.

Consider the elements of the reduced gradient vector (1.10), one at a time, together with all the measurement equations (1.9). Let the superscript (i) denote the i-th row of a matrix or a vector. We write the reduced gradient together with the measurements as a sequence of square linear systems

$$\underbrace{\begin{bmatrix} [\boldsymbol{NJ_{zz}}]^{(i)} & [\boldsymbol{NJ_{zd}}]^{(i)} & 0 \\ \boldsymbol{G}^y & \boldsymbol{G}_d^y & y \end{bmatrix}}_{\boldsymbol{M}^{(i)}} \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{d} \\ -1 \end{bmatrix} = 0 \tag{1.12}$$

Here $\boldsymbol{M}^{(i)}$, $i = 1..n_{DOF}$ are square matrices of size $(n_y + 1) \times (n_y + 1)$.

We want to find a solution $[\boldsymbol{z},\boldsymbol{d}]$ which satisfies (1.12). For this system to have a solution for $[\boldsymbol{z},\boldsymbol{d}]^{\boldsymbol{T}}$, we must have $\text{rank}(\boldsymbol{M^{(i)}}) = n_y = n_z + n_d$.

The submatrix $[\boldsymbol{G}^y\ \boldsymbol{G}_d^y\ \boldsymbol{y}]$ already has rank n_y, irrespective of the value of $\boldsymbol{y}$ (or the control policy that generates the input $\boldsymbol{u}$ which in turn generates

$\boldsymbol{y}$). This follows because $[\boldsymbol{G}^y\ \boldsymbol{G}_d^y\ \boldsymbol{y}]$ has more columns than rows, and because $\text{rank}([\boldsymbol{G}^y\ \boldsymbol{G}_d^y]) = n_y$. Therefore, the condition for a common solution is:

$$\det(\boldsymbol{M}^{(i)}) = 0 \quad \text{for all } i = 1..n_{DOF}. \tag{1.13}$$

This condition guarantees that a common solution to (1.12) exists, so the elements of the controlled variable $\boldsymbol{c}$ are selected as $c_i = \det(\boldsymbol{M}^{(i)})$.

It remains to show that controlling the determinants $c_i = \det(\boldsymbol{M}^{(i)})$ gives the inputs which lead to the optimum. Since the system is linear and the rank of the measurement equations is n_y, there is a unique linear invertible mapping between the measurements $\boldsymbol{y}$ and the vector $[\boldsymbol{z}\,\boldsymbol{d}]^{\boldsymbol{T}}$. Therefore every value of $\boldsymbol{y}$ corresponds uniquely to some value in $\boldsymbol{z}$.

In the case with more measurements, $n_y > n_z + n_d$, any subset of $n_z + n_d$ measurements may be chosen such that $\text{rank}([\boldsymbol{G}^y\ \boldsymbol{G}_d^y]) = n_z + n_d$.

Remark 1.1. Actually, in the case of (1.8), we can use the constraint equations to eliminate the unmeasured internal states $\boldsymbol{x}$. Then we only $n_y = n_u + n_d$ measurements, and the matrices (1.12) become:

$$\boldsymbol{M}^{(i)} = \begin{bmatrix} (\boldsymbol{N}\boldsymbol{J}_{\boldsymbol{zz}})^{(i)} & (\boldsymbol{N}^{\boldsymbol{T}}\boldsymbol{J}_{\boldsymbol{zd}})^{(i)} & 0 \\ \boldsymbol{A} & 0 & \boldsymbol{b} \\ \boldsymbol{G}^y & \boldsymbol{G}_d^y & \boldsymbol{y} \end{bmatrix}, \tag{1.14}$$

and we must require, that

$$\text{rank}\left(\begin{bmatrix} \boldsymbol{A} & 0 \\ \boldsymbol{G}^y & \boldsymbol{G}_d^y \end{bmatrix}\right) = n_z + n_d \tag{1.15}$$

Remark 1.2. When there are no constraints, we have that $\boldsymbol{z} = \boldsymbol{u}$, and this method results in the null space method (Alstad and Skogestad, 2007). In this case, $\boldsymbol{N}$ may be set to any nonsingular matrix, for example the identity matrix $\boldsymbol{N} = \text{i}$. Then we have that

$$\boldsymbol{c}_{\text{Nullspace}} = [\boldsymbol{J}_{\boldsymbol{uu}}\ \ \boldsymbol{J}_{\boldsymbol{ud}}][\tilde{\boldsymbol{G}}^y]^{-1}\boldsymbol{y}, \tag{1.16}$$

as has been derived in Alstad et al (2009).

The null space method was originally derived by Alstad and Skogestad (2007) using the optimal sensitivity matrix $\boldsymbol{F} = \frac{\partial \boldsymbol{y}^{opt}}{\partial \boldsymbol{d}}$. However, it escaped the authors notice then, that controlling $\boldsymbol{c} = \boldsymbol{H}\boldsymbol{y}$ with $\boldsymbol{H}$ selected such that $\boldsymbol{H}\boldsymbol{F} = 0$, is indeed the same as controlling the gradient to zero.

Example 1.1. Consider a system from Alstad (2005). The cost to minimize is

$$J = (u - d)^2, \tag{1.17}$$

and the measurement relations (model equations) are

$$\begin{aligned} y_1 &= G_1^y u + G_{d,1} d \\ y_2 &= G_2^y u + G_{d,2}^y d \end{aligned} \tag{1.18}$$

where the variables u, d, y denote the input, the disturbance and the measurements, respectively. The values of the gains are given in table 1.1. We are

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

Table 1.1 Gain values for small example

searching for a condition on 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 independent. This gives an equation system of 3 equations in 2 unknowns:

$$\boldsymbol{M} \begin{bmatrix} u \\ d \\ -1 \end{bmatrix} = 0 \tag{1.19}$$

where

$$\boldsymbol{M} = \begin{bmatrix} J_{uu} & J_{ud} & 0 \\ G_1^y & G_{d,1}^y & y_1 \\ G_2^y & G_{d,2}^y & y_2 \end{bmatrix} \tag{1.20}$$

Equation (1.19) has a solution $\begin{bmatrix} u \\ d \\ -1 \end{bmatrix}$ if and only if

$$\det(\boldsymbol{M}) = 0 \tag{1.21}$$

Therefore the necessary and sufficient condition for the existence of a non-trivial solution is

$$\begin{aligned} \det\left(\begin{bmatrix} J_{uu} & J_{ud} & 0 \\ G_1^y & G_{d,1}^y & y_1 \\ G_2^y & G_{d,2}^y & y_2 \end{bmatrix}\right) &= -y_1(J_{uu}G_{d,2}^y - G_2^y J_{ud}) + y_2(J_{uu}G_{d,1}^y - G_1^y J_{ud}) \\ &= 0 \end{aligned} \tag{1.22}$$

On inserting the parameter values from table 1.1, we obtain

$$c = \det(\boldsymbol{M}) = y_1 + 2y_2. \tag{1.23}$$

Controlling $c = y_1 + 2y_2$ to zero therefore yields optimal operation. This the same variable combination as found by applying the null-space method in Alstad (2005).

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

1.5 Elimination for Systems of Polynomial Equations

Let $\hat{\boldsymbol{d}}$ now denote the vector of all unmeasured variables, $\hat{\boldsymbol{d}} = [\boldsymbol{x}, \boldsymbol{d}]$, not only including disturbances, but also unknown states and unknown inputs, and let $\boldsymbol{y}$ include all measurements and known inputs. Thus, every variable belongs either to $\hat{\boldsymbol{d}}$ or $\boldsymbol{y}$, and we write the optimality conditions as

$$\begin{aligned} J_{\boldsymbol{z},red}(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0 \\ g(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0, \end{aligned} \tag{1.24}$$

and the measurement relations as

$$m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0. \tag{1.25}$$

To be able to use the reduced gradient $J_{\boldsymbol{z},red}$ for control, all unknown variables $\hat{\boldsymbol{d}}$ have to be eliminated from it. For polynomial equations, this is not as straightforward as in the linear case. 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 Abel (1826). Therefore we are interested in finding a way to eliminate the unknown variables $\hat{\boldsymbol{d}}$ from $J_{\boldsymbol{z},red}(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$ without solving g and m for them first. This is exactly what was done in the previous section, where we used the determinant of a carefully constructed coefficient matrix, which characterizes the existence of a common solution in $\boldsymbol{d}$, to replace $J_{\boldsymbol{z},red}$. The determinant is a function of the known variables only, that is, the measurements $\boldsymbol{y}$ and the parameters $\tilde{\boldsymbol{G}}^y$, $\boldsymbol{J}_{\boldsymbol{zz}}$ and $\boldsymbol{J}_{\boldsymbol{zd}}$.

The generalization of the determinant to systems of polynomial equations is called resultant. According to Emiris and Mourrain (1999), "the resultant of an overconstrained polynomial system characterizes the existence of common roots as a condition on the input coefficients".

To be more specific, we consider multivariate polynomials $f \in \mathbb{R}[\boldsymbol{y}, \hat{\boldsymbol{d}}]$, that is real polynomial functions with coefficients in $\mathbb{R}$, and variables $\boldsymbol{y} = [y_1, y_2, \ldots, y_{n_y}]$ and $\hat{\boldsymbol{d}} = [\boldsymbol{x}, \boldsymbol{d}] = [\hat{d}_1, \hat{d}_2, \ldots, \hat{d}_{n_{\hat{d}}}]$. Given a $n_{\hat{d}}$-tuple, $\alpha_{i,j} = \big(\alpha_{i,j}(1), \alpha_{i,j}(2), \ldots, \alpha_{i,j}(n_{\hat{d}})\big)$, we use the shorthand notation

$$\hat{\boldsymbol{d}}^{\alpha_{i,j}} = \hat{d}_1^{\alpha_{i,j}(1)} \hat{d}_2^{\alpha_{i,j}(2)} \dots \hat{d}_{n_{\hat{d}}}^{\alpha_{i,j}(n_{\hat{d}})}.$$

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

$$f_i(\boldsymbol{y}, \hat{\boldsymbol{d}}) = \sum_{j=0}^{k_i} a_{i,j}(\boldsymbol{y}) \hat{\boldsymbol{d}}^{\alpha_{i,j}}, \quad i = 1..n \tag{1.26}$$

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

We consider the functions $a_{i,j}(\boldsymbol{y})$ as polynomial coefficients, and $\hat{\boldsymbol{d}}$ as variables. For every polynomial f_i, we collect the exponent vectors in the set $\mathcal{E}_i = \{\alpha_{i,1}, \dots, \alpha_{i,k_i}\}$. 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 the convex hull of the support of a polynomial, $Q_i = \text{conv}(\mathcal{E}_i)$.

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

Next present some basic concepts from algebraic geometry taken from Cox et al (2005).

Definition 1.1 (Affine variety). Consider $f_1, \dots, f_n$ polynomials in $\mathbb{C}[\hat{d}_1, \dots \hat{d}_{n_{\hat{d}}}]$. The affine variety defined by $f_1, \dots, f_n$ is the set

$$V(f_1, \cdots, f_n) = \left\{ (\hat{d}_1, \dots, \hat{d}_{n_{\hat{d}}}) \in \mathbb{C}^{n_{\hat{d}}} : f_i(\hat{d}_1, \dots \hat{d}_{n_{\hat{d}}}) = 0 \quad i = 1 \dots n \right\} \tag{1.27}$$

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

Definition 1.2 (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 $\mathcal{E}_i$:

$$L(\mathcal{E}_i) = \left\{ a_{i,1} \hat{\boldsymbol{d}}^{\alpha_{i1}} + \dots + a_{i,k_i} \hat{\boldsymbol{d}}^{\alpha_{ik_i}} : a_{i,j} \in \mathbb{C} \right\} \tag{1.28}$$

Then the coefficients $a_{i,j}$ of n polynomials define a point in $\mathbb{C}^{n \times k_i}$.

Now let

$$Z(\mathcal{E}_1, \dots \mathcal{E}_n) \subset L(\mathcal{E}_1) \times \dots \times L(\mathcal{E}_n) \tag{1.29}$$

be the Zariski closure of the set of all $(f_1, \dots f_n)$ for which (1.26) has a solution in $(\mathbb{C}^*)^{n_{\hat{d}}}$ (that is the Zariski closure of all coefficients $a_{i,j} \in \mathbb{C}$ for which (1.26) has a solution).

For an overdetermined system of $n_{\hat{d}} + 1$ polynomials in $n_{\hat{d}}$ variables we have following result:

Theorem 1.1 (Sparse resultant (Gelfand et al, 1994; Cox et al, 2005)). *Assume that* $Q_i = \mathrm{conv}\,(\mathcal{E}_\mathrm{i})$ *is an* $n_{\hat{d}}$*-dimensional polytope for* $i = 1, \ldots, n_{\hat{d}} + 1$*. Then there is an irreducible polynomial* $\mathcal{R}$ *in the coefficients of the* f_i *such that*

$$(f_1, \ldots, f_{n_{\hat{d}}+1}) \in Z(\mathcal{E}_1, \ldots, \mathcal{E}_{n_{\hat{d}+1}}) \Longleftrightarrow \mathcal{R}(f_1, \ldots, f_{n_{\hat{d}}+1}) = 0. \tag{1.30}$$

In particular, if

$$f_1(d_1 \ldots d_{n_{\hat{d}}}) = \cdots = f_{n_{\hat{d}}+1}(d_1 \ldots d_{n_{\hat{d}}}) = 0 \tag{1.31}$$

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

$$\mathcal{R}(f_1, \ldots, f_{n_{\hat{d}}+1}) = 0. \tag{1.32}$$

Remark 1.3. The requirement that Q_i has to be $n_{\hat{d}}$-dimensional is no restriction and can be relaxed, (Sturmfels, 1994). 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 the coefficient matrix. 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 Gelfand et al (1994); Sturmfels (1994); Cox et al (2005). An overview of different matrix constructions in elimination theory is given in Emiris and Mourrain (1999).

We choose to use the sparse resultant, because 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 $x^5, y^5, x^4y, xy^4, \ldots, x^4, y^4, x^3y, \ldots, y, x, 1$. 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 $(\mathbb{C}^*)^{n_{\hat{d}}}$, such that the input polynomials need not be homogeneous (or homogenized), as in other resultants.

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

Generally, 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. In this work, we use the maple software package `multires` Busé and Mourrain (2003), which can be downloaded from the internet[1].

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

For more details on the theory of sparse resultants, we refer to Gelfand et al (1994); Emiris and Mourrain (1999); Sturmfels (2002); Dickenstein and Emiris (2005).

1.5.1 Finding Invariant Controlled Variables for Polynomial Systems

After introducing the concepts above, we are ready to apply them in the context of controlled variable selection and self-optimizing control. As in the linear case above, we assume that the active constraints and the model equations, $g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$, and the measurement relations, $m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$, are satisfied.

Let $J^{(i)}_{\boldsymbol{z},red}$ denote the i-th element in the reduced gradient expression. To obtain the n_c controlled variables needed for the unconstrained degrees of freedom we have:

Theorem 1.2 (Nonlinear measurement combinations as controlled variables). *Given $\hat{\boldsymbol{d}} \in (\mathbb{R}^*)^{n_{\hat{d}}}$, and $n_y + n_g = n_{\hat{d}}$, independent relations $g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$ such that the system*

$$\begin{aligned} g(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0 \\ m(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0 \end{aligned} \tag{1.33}$$

has finitely many solutions for $\hat{\boldsymbol{d}} \in (\mathbb{C}^*)^{n_{\hat{d}}}$. Let $\mathcal{R}(J^{(i)}_{z,red}, g, m), \quad i = 1 \dots n_{\boldsymbol{c}}$ be the sparse resultants of the n_c polynomial systems composed of

$$J^{(i)}_{z,red}(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0, \quad g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0, \quad m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0 \quad i = 1 \cdots n_c, \tag{1.34}$$

then controlling the active constraints, $g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$, and $c_i = \mathcal{R}(J^{(i)}, g, m) \quad i = 1, \dots, n_c$, yields optimal operation throughout the region.

Proof. The active constraints are controlled, so $g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$ and $m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$ are satisfied always, and there is no condition on the parameters for this part of the system.

The system $g(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0, m(\boldsymbol{y}, \hat{\boldsymbol{d}}) = 0$ has only finitely many solutions for $\hat{\boldsymbol{d}}$, so the set of possible $\hat{\boldsymbol{d}}$ is fixed. Moreover, we know that a real solution to the subsystem $g(\boldsymbol{y}, \boldsymbol{d}) = m(\boldsymbol{y}, \boldsymbol{d}) = 0$ exists, since it is the given disturbance.

From Theorem 1.1, the sparse resultant gives the necessary and sufficient conditions for the existence of a solution for (1.34) in $\boldsymbol{d} \in (\mathbb{C}^*)^{n_d}$. Therefore, whenever $J^{(i)}_{z,red} = 0$, the resultant is zero (necessary condition). On the other hand if $\mathcal{R}(J_{z,red}, g, m) = 0$ then the system (1.34) is satisfied, too (sufficient condition).

This holds for any solution $\hat{\boldsymbol{d}} \in (\mathbb{C}^*)^{n_{\hat{d}}}$, and in particular the "actual" values of $\hat{\boldsymbol{d}}$. Because there are as many resultants as unconstrained degrees of

freedom, controlling $\mathcal{R}(J^{(i)}_{z,red}, g, m)$ for $i = 1, \ldots, n_u$ satisfies the necessary conditions of optimality in the region.

Remark 1.4. In cases where the $\hat{\boldsymbol{d}} \notin (\mathbb{C}^*)^{n_{\hat{d}}}$, we may apply a variable transformation to formulate the problem such we get $\hat{\boldsymbol{d}} \in (\mathbb{C}^*)^{n_{\hat{d}}}$. For example a translation $d = \tilde{d} - 1$.

Remark 1.5. By partitioning the overall optimization problem into several regions of active constraints, we assume that we have obtained well behaving systems for each region. In particular it is assumed that there are no base points (values of $a_{i,j}(\boldsymbol{y})$, where a polynomial in g or m vanishes for all values of $\hat{\boldsymbol{d}}$).

Remark 1.6. In some cases, the matrix of coefficients may be singular, yielding an identically zero determinant. These cases can be handled by a perturbation of the system at that point. This is a standard method of handling degeneracies in resultants Canny (1990); Rojas (1999).

Example 1.2 (Case with one disturbance). Consider the system of two polynomials in one unknown variable d, with one measurement relation $m(y, d) = 0$. At the optimum we must have:

$$\begin{aligned} J_{z,red} = \boldsymbol{N}\nabla_{\boldsymbol{z}} J(y, d) = a_{1,1}(y) + a_{1,2}(y)d = 0 \\ m(y, d) = a_{2,1}(y) + a_{2,2}(y)d + a_{2,3}(y)d^2 = 0 \end{aligned} \tag{1.35}$$

This system of univariate polynomials in d is overdetermined, and does not have a common solution d^* for arbitrary coefficients $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}, a_{2,3}$. The sparse resultant coincides in the case of univariate polynomials with the classical resultant, which is the determinant of the Sylvester matrix (Cox et al, 1992), and the vanishing of the resultant is the necessary and sufficient condition for the existence of a common root. We construct the Sylvester matrix

$$Syl = \begin{bmatrix} a_{1,2}(y) & a_{1,1}(y) & 0 \\ 0 & a_{1,2}(y) & a_{1,1}(y) \\ a_{2,3}(y) & a_{2,2}(y) & a_{2,1}(y) \end{bmatrix}, \tag{1.36}$$

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

$$\mathcal{R}\left(J_{z,red}, m(y, d)\right) = \det(Syl) = a_{1,2}^2 a_{2,1} - a_{1,2}a_{1,1}a_{2,2} + a_{2,3}a_{1,1}^2 \tag{1.37}$$

For a common root d^* to exist, the polynomial in the coefficients, $\mathrm{Res}(f_1, f_2)$ must vanish. Since the model $m(y, d)$ is chosen such that it is always satisfied, $m(y, d) = 0$ for any disturbance $d \in \mathbb{R}$, controlling the resultant to zero is the condition for the reduced gradient $J_{\boldsymbol{z},red}$ to become zero. So for a particular given exogenous $d \in \mathbb{R}$, the optimality conditions will be satisfied, and operation will be optimal.

1.6 Switching Operating Regions

In this section, we present a pragmatic approach for detecting when to change the control structure, because of changes in the active set. This task is a research field in itself (Baotić et al (2008) has worked on linear systems with quadratic objectives), and an exhaustive study is outside the scope of this paper. However, we would like to present a procedure, which may be used as starting point for a more thorough investigation of this problem in future work.

From a pure optimization perspective, there is no difference between a constraint and a controlled variable $\boldsymbol{c}(\boldsymbol{y})$, 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 $\boldsymbol{c}(\boldsymbol{y})$. First of all, the active constraints and the controlled variables $\boldsymbol{c}(\boldsymbol{y}) = 0$ are not satisfied automatically, that is one has to 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. The basic idea is to monitor the controlled variables and the active constraints of all neighbouring regions.

The main assumptions are that the regions are adjacent, that the disturbance moves the system continuously from one region to another, and that the system cannot jump over regions. In addition, we assume that controlling $\boldsymbol{c}(\boldsymbol{y}) = 0$ is equivalent to controlling the gradient to zero, as shown in the previous section. To determine when the control structure should be switched, we propose two rules:

1. (A new constraint becomes active) When a new constraint becomes active, change the control structure to the corresponding region
2. (A constraint becomes inactive) As soon as the controlled variable $\boldsymbol{c}$ in one of the neighbouring regions becomes zero (reaches its optimal setpoint), change the control structure to the corresponding region.

However, since we are working with systems of polynomial equations, there are some potential pitfalls here. The first one is that we are assuming that the regions of active constraints are adjacent, and that a changing disturbance moves the system continuously to from one region into another. Although this is the case for many systems in practice, it has to be confirmed that this holds for the particular case in consideration.

The second pitfall is that since our controlled variables are derived from the optimality conditions, this method will give optimal operation (and switching), as long as the same optimality conditions cannot be satisfied at two

distinct $\boldsymbol{d}$. This will hold if the optimization problem is convex in the disturbance space of interest.

1.7 Case Study

We consider an isothermal CSTR with two parallel reactions, as depicted in Fig. 1.2, taken from Srinivasan et al (2008). 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.

The products C and D are formed by the reactions:

$$\begin{aligned} \mathrm{A} + \mathrm{B} &\xrightarrow{k_1} \mathrm{C} \\ 2\,\mathrm{B} &\xrightarrow{k_2} \mathrm{D} \end{aligned} \tag{1.38}$$

We wish to maximize the amount of desired product $(F_A+F_B)c_C$, weighted by a yield factor $(F_A+F_B)c_C/(F_A c_{A,in})$ (Srinivasan et al, 2008). The amount of heat to remove and the maximum flow rate are limited by the equipment, and we formulate the mathematical optimization problem as follows (Srinivasan et al, 2008):

$$\max_{F_A,F_B} \frac{(F_A+F_B)c_C}{F_A c_{A_{in}}}(F_A+F_B)c_C \tag{1.39}$$

subject to

$$\begin{aligned} F_A c_{A_{in}} - (F_A+F_B)c_A - k_1 c_A c_B V &= 0 \\ F_B c_{B_{in}} - (F_A+F_B)c_B - k_1 c_A c_B V - 2k_2 c_B^2 V &= 0 \\ -(F_A+F_B)c_C + k_1 c_A c_B V &= 0 \\ F_A + F_B &\le F_{max} \\ k_1 c_A c_B V(-\Delta H_1) + 2k_2 c_B V(-\Delta H_2) &\le q_{max} \end{aligned} \tag{1.40}$$

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, V the reactor volume, and F_{max} the maximum total flow rate. The measured variables ($\boldsymbol{y}$), the manipulated variables ($\boldsymbol{u}$), the disturbance variables ($\boldsymbol{d}$), and the internal states are given in table 1.2, and the parameter values of the system are listed in table 1.3.

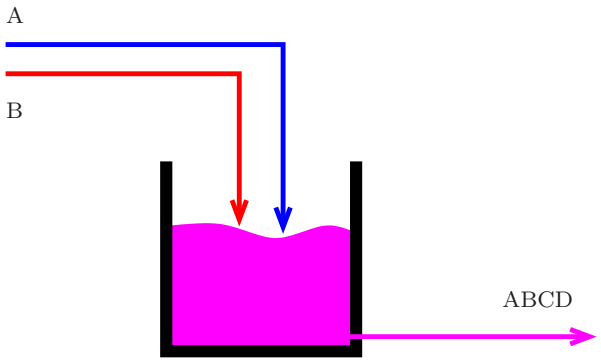

Fig. 1.2 CSTR with two reactions

Table 1.2 Overview of variables

Symbol	Description	Comment
F_A	Inflow stream A	measured input
F_B	Inflow stream B	measured input
F	total flow	measured variable
q	heat produced	measured variable
c_B	concentration of B	measured variable
c_A	concentration of A	unmeasured variable
c_C	concentration of C	unmeasured variable
k_1	rate constant reaction 1	unmeasured disturbance

Table 1.3 Parameters

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	22
q_{max}	kJ/h	1000

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

$$\boldsymbol{z} = \left[c_A, c_B, c_C, F_A, F_B \right]^{\boldsymbol{T}} . \tag{1.41}$$

1.7.1 Identifying Operational Regions

Following the procedure from Section 1.3, the system is optimized off-line for the range of possible disturbances $d = k_1$. This shows that the system can be partitioned into three adjacent critical regions, defined by their active constraints.

The critical regions are visualized in Fig. 1.3, where the normalized constraints are plotted over the disturbance range. In the first region, for disturbances below about $k_1 = 0.65\frac{\text{l}}{\text{mol h}}$, the flow constraint is the only active constraint. The second critical region for values between about $k_1 = 0.65\frac{\text{l}}{\text{mol h}}$ and $k_1 = 0.8\frac{\text{l}}{\text{mol h}}$ 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 $k_1 = 0.8\frac{\text{l}}{\text{mol h}}$ only the heat constraint remains.

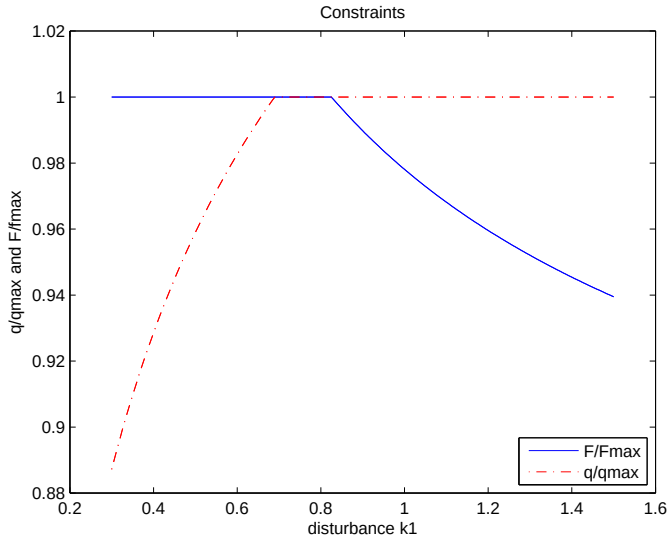

Fig. 1.3 Optimal values of the constrained variables

1.7.2 Eliminating λ

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, $n_{DOF} =$

$n_{\boldsymbol{z}} - n_g$. 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 $\boldsymbol{N}_z^{\boldsymbol{T}}$ and multiply it with the gradient of the objective function $\nabla_{\boldsymbol{z}} J(\boldsymbol{z}, \boldsymbol{d})$ to obtain $J_{z,red,1} = \boldsymbol{N}_z^{\boldsymbol{T}} \nabla_{\boldsymbol{z}} J$. 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 $J_{z,red,1}$ to zero, it is sufficient to consider only the numerator of $J_{z,red,1}$. This is possible because a fraction vanishes if the numerator is zero (provided the denominator is nonzero which is the case here because $\nabla_{\boldsymbol{z}} g$ has full rank). For the critical region 1, we obtain from (1.7) the invariant

$$\begin{aligned} J_{z,red,1} = -(F_A + F_B)^2 c_C \big[& -3c_C F_B^2 F_A - 3c_C F_A^2 F_B \\ & - 4c_C c_B F_A^2 k_2 V - 4c_C k_2 V^2 k_1 c_B^2 F_A - c_C F_A^3 \\ & - c_C F_B^3 - 4c_C k_2 V^2 k_1 c_B^2 F_B - c_C c_B F_A^2 k_1 V \\ & - 4c_C c_B F_B^2 k_2 V - c_C c_B F_B^2 k_1 V - c_C F_A^2 c_A k_1 V \\ & - c_C F_B^2 c_A k_1 V - 8c_C F_A c_B F_B k_2 V \\ & - 2c_C F_A c_B F_B k_1 V - 2c_C F_A F_B c_A k_1 V \\ & + 8F_A k_1 V^2 c_{A,in} k_2 c_B^2 + 2F_A^2 k_1 V c_B c_{A,in} \\ & + 2F_A k_1 V F_B c_B c_{A,in} - 2F_A^2 k_1 V c_{B,in} c_A \\ & -2F_A k_1 V F_B c_{B,in} c_A \big] \end{aligned} \tag{1.42}$$

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 control the factor in square brackets in (1.42) to zero.

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 control the active constraint ($q = q_{max}$) while the other one is used to control the invariant measurement combination $J_{z,red,3}$, which is an expression similar to (1.42).

1.7.3 Eliminating Unknown Variables

The invariant variable combinations for the first and the third critical region $J_{z,red,1}$ and $J_{z,red,3}$ 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 $\boldsymbol{y}$, and all unknown variables into $\hat{\boldsymbol{d}}$, such that $\hat{\boldsymbol{d}} = [k_1, c_a, c_C]$. Then we write the necessary conditions for optimality as for each region as

$$\begin{aligned} J_{z,red}(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0 \\ g(\boldsymbol{y}, \hat{\boldsymbol{d}}) &= 0. \end{aligned} \tag{1.43}$$

Considering the known variables $\boldsymbol{y}$ as parameters of the system, we want to find conditions on these parameters such that (1.43) is satisfied. The system has $n_{\hat{d}} = 3$ unknown variables, k_1, c_a and c_c, 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 1.5 we have that (1.43) is satisfied if and only if the sparse resultant is zero.

In critical region 1 and 3, the number of equations $n_{eq} = 5$ (model equations+active constrains+invariant), and the number of unknowns $n_{\hat{d}} = 3$. So we have more equations than necessary. Since we assume no measurement noise, all measurements are equally good, and we may select a subset of $n_{\hat{d}} + 1$ equations from (1.43) to compute the sparse resultant for the subset of equations. Obviously, the reduced gradient must be contained in this set of equations. Alternatively, as we do in the following, we can eliminate one more variable from the invariant.

For the first region, we use the sparse resultant of the system consisting of the invariant (1.42), the model equations (the first three equality constraints in (1.40)) and the first (active) inequality constraint in (1.40) 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 (Busé and Mourrain, 2003). After division by nonzero factors, the controlled variable for region 1 becomes:

$$\begin{aligned} c_1 = &-c_{b,in}^2 F_A^2 - F_A^2 c_{A,in} c_{b,in} + 6 F_A c_{A,in} k_2 c_b^2 V + 2 F_A c_{A,in} F_{max} c_b \\ &- F_A c_{A,in} F_{max} c_{b,in} + F_{max}^2 c_b^2 + c_{b,in}^2 F_{max}^2 + 4 V^2 k_2^2 c_b^4 \\ &- 2 c_{b,in} F_{max}^2 c_b - 4 V k_2 c_b^2 c_{b,in} F_{max} + 4 V k_2 c_b^3 F_{max} \end{aligned} \tag{1.44}$$

In the second critical region, control is simple; the two manipulated variables 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 input variable is used to control the active constraint, and the second input is used to control the resultant. The model equations (the first three equations) together with the energy constraint) in (1.40) 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:

$$
\begin{aligned}
c_3 = &-4Vc_B^2k_2\Delta H_2F_Ac_{A,in}c_{B,in}q_{max}\Delta H_1 + F_Ac_{B,in}^2q_{max}^2\Delta H_1 \\
&+ 4V^2c_B^4k_2^2\Delta H_2F_Ac_{A,in}c_{B,in}\Delta H_1^2 - 4V^2c_B^4k_2^2\Delta H_2^2F_Ac_{A,in}c_{B,in}\Delta H_1 \\
&- 2Vc_B^2k_2F_Ac_{A,in}c_{B,in}\Delta H_1^2q_{max} - 4Vc_B^2k_2\Delta H_2F_Ac_{B,in}^2\Delta H_1q_{max} \\
&- 2Vc_B^2k_2\Delta H_2F_A^2c_{A,in}c_{B,in}^2\Delta H_1^2 + 8Vc_B^3k_2\Delta H_2\Delta H_1F_Ac_{A,in}q_{max} \\
&- 8V^2c_B^4k_2^2\Delta H_2c_{B,in}\Delta H_1q_{max} - 12V^2c_B^4k_2^2F_A\Delta H_2^2c_{B,in}^2\Delta H_1 \\
&- 8V^2c_B^5k_2^2\Delta H_2F_Ac_{A,in}\Delta H_1^2 + 8V^2c_B^5k_2^2\Delta H_2^2\Delta H_1F_Ac_{A,in} \\
&+ 8V^2c_B^5k_2^2F_A\Delta H_2^2c_{B,in}\Delta H_1 - q_{max}^3c_{B,in} + 2c_Bq_{max}^3 \\
&- \Delta H_1c_{B,in}F_Ac_{A,in}q_{max}^2 + 2c_BF_Ac_{A,in}q_{max}^2\Delta H_1 + F_A^2c_{A,in}c_{B,in}^2\Delta H_1^2q_{max} \\
&- 2c_BF_Ac_{B,in}q_{max}^2\Delta H_1 + 8Vc_B^3k_2\Delta H_2q_{max}^2 + 8V^2c_B^5k_2^2\Delta H_2^2q_{max} \\
&+ 8V^3c_B^6k_2^3\Delta H_2^3c_{B,in} - 2c_BF_A^2c_{A,in}c_{B,in}\Delta H_1^2q_{max} \\
&- 2Vc_B^2k_2\Delta H_1q_{max}^2c_{B,in} - 2Vc_B^2k_2\Delta H_2q_{max}^2c_{B,in} \\
&+ 4V^2c_B^4k_2^2\Delta H_2^2c_{B,in}q_{max} - 8V^3c_B^6k_2^3\Delta H_2^2c_{B,in}\Delta H_1
\end{aligned}
\tag{1.45}
$$

Although these expressions are quite involved, they contain only known quantities, and can be easily 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 F_{max}, and c_2 is scaled by $\Delta H_1^2\Delta H_2F_AF_B$.

1.7.4 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 and the required cooling to keep the system isothermal, until maximum cooling is reached, Fig. 1.4. When the constraint is reached, the control structure is switched 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 neighbouring 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 $c_3 = 0$.

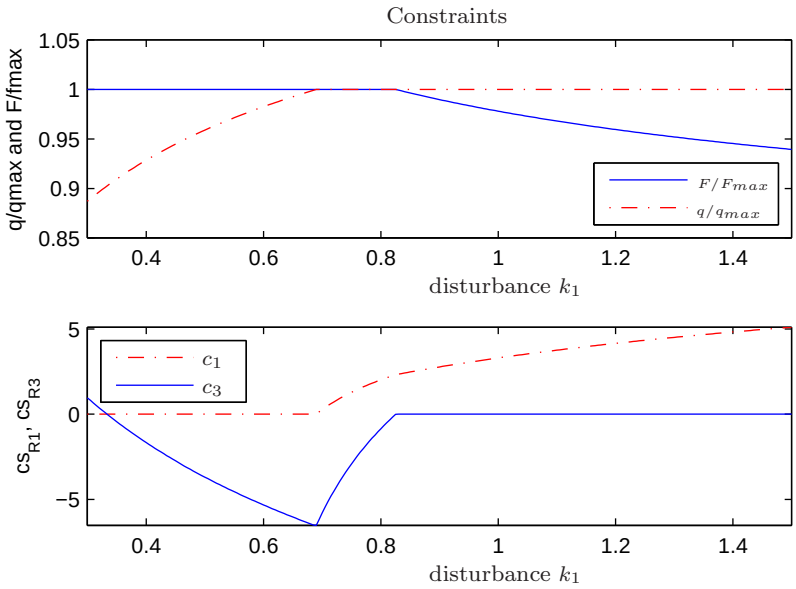

Fig. 1.4 Optimal values of controlled variables

1.8 Discussion

The presented method is based on the same idea as NCO tracking (François et al, 2005). However in contrast to NCO tracking, where the optimality conditions are solved for the optimizing *inputs*, this work focuses on finding the right *outputs* which express the optimality conditions. The corresponding inputs are generated by PI controllers and feedback control.

The method was developed as an alternative derivation and a generalization of the existing null space method (Alstad and Skogestad, 2007) for linear systems.

In the linear case, eliminating the constraints is straight forward, while this is not trivial in the polynomial case. However, by premultiplying ∇J by the null space of the constraints $\boldsymbol{N^T}$, we eliminate the Lagrangian multipliers from the equation set, and obtain the reduced gradient for the nonlinear case. The elimination of the Lagrangian multipliers can also be done by the resultant. Under the strict complementarity condition (either $\lambda = 0$ *or* the constraint is active), the solutions for λ 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 in this work, since this results in lower computational load when computing the resultants.

As an alternative to calculating resultants, the controlled variable combinations could be computed using Gröbner bases with an appropriate elimination ordering (Cox et al, 1992). One could use an appropriate monomial ordering which eliminates the unknown variables, and then use a polynomial from the elimination ideal as controlled variable. However, this Gröbner basis approach has some disadvantages, as it is not straightforward to find an elimination order which eliminates the unknown variables from the equation system while not yielding the "trivial solution" (i. e. the invariant is always zero when the constraints are satisfied). Another problem is that the selected invariant might give rise to "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 $g(\boldsymbol{y}, \hat{\boldsymbol{d}})$ and $m(\boldsymbol{y}, \hat{\boldsymbol{d}})$, and to reduce the $\boldsymbol{N}\nabla_{\boldsymbol{z}} J$ modulo the ideal. This avoids the trivial solution, however, the problem of choosing a monomial ordering which eliminates all unknown variables, remains.

Another argument against using a Gröbner basis for calculation the invariant, is, that it can yield very large and complicated expressions.

Since also the sparse resultants can give large expressions, in practice the method is best suited for small systems, with few constraints and 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 mixed subdivision, which is a hard enumeration problem (Cox et al, 2005).

In many cases (and in our case study) there are more polynomial equations than unknowns. Then the engineer has to chose which model polynomials to use in the resultant calculations in addition to the reduced gradient. From a purely mathematical view, this does not make any difference, as long as the set of model equations has finitely many solutions for $\boldsymbol{d}$. However the controlled variables will look quite different for different choices. 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, it would be advisable to keep the degrees of the polynomials low in the unknown variables. This leads to simpler resultants.

The resultant method, as presented above, does not take into account errors in the model and measurements. This is beyond the scope of this work. Our goal was to extend the idea of the null-space method (Alstad and Skogestad, 2007) and to demonstrate on a small example that the concept of finding variables which remain constant at optimal operation is possible also for polynomial systems.

Apart from handling noisy measurements and model mismatch, a further subject for future research is to find methods which reproduce not all solutions of the optimality conditions, but only a certain set of interest. This could be all the real solutions or solutions which reside in some further specified semialgebraic set.

1.9 Conclusions

In this chapter we have presented an approach to obtain optimal steady state operation which does not require online calculations. We have shown that, after identifying the critical regions, there exist optimally invariant variable combination for each region. If there are enough measurement/model relations ($n_g + n_m \geq n_{\hat{d}}$), the unknown variables can be eliminated by measurements and system equations, and the invariant combinations can be used for control using a decentralized self-optimizing control structure.

Further, we have shown that the measurement invariants can be used for detecting changes in the active set and for finding the right region to switch to.

Using methods form elimination theory, we have shown that, in principle, the idea of using polynomials in the measurements as self-optimizing control variables can be used to control the process optimally.

References

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

Alstad V (2005) Studies on selection of controlled variables. PhD thesis, Norwegian University of Science and Technology, Department of Chemical Engineering

Alstad V, Skogestad S (2007) Null space method for selecting optimal measurement combinations as controlled variables. Ind Eng Chem Res 46:846–853

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

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

Bonvin D, Srinivasan B, Ruppen D (2001) Dynamic Optimization in the Batch Chemical Industry. in: Chemical Process Control-VI

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

Canny JF (1990) Generalised characteristic polynomials. J Symbolic Computation 9:241–250

Cao Y (2003) Self-optimizing control structure selection via differentiation. In: Proceedings of the European Control Conference

Cox D, Little J, O'Shea D (1992) Ideals, Varieties, and Algorithms. Springer-Verlag

Cox D, Little J, O'Shea D (2005) Using Algebraic Geometry. Springer

Dickenstein A, Emiris IZ (2005) Solving Systems of Polynomial Equations. Springer Berlin Heidelberg

Emiris IZ, Mourrain B (1999) Matrices in elimination theory,. Journal of Symbolic Computation 28(1-2):3–43, DOI DOI:10.1006/jsco.1998.0266

François G, Srinivasan B, Bonvin D (2005) Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty. Journal of Process Control 15(6):701 – 712, DOI DOI:10.1016/j.jprocont.2004.11.006

Gelfand IM, Kapranov MM, Zelevinsky AV (1994) Discriminants, Resultants and Multidimensional Determinants. Birkhäuser, Boston, MA

Halvorsen IJ, Skogestad S (1997) Indirect on-line optimization through setpoint control. AIChE 1997 Annual Meeting, Los Angeles; paper 194h

Halvorsen IJ, Skogestad S, Morud JC, Alstad V (2003) Optimal selection of controlled variablesâĂă. Industrial & Engineering Chemistry Research 42(14):3273–3284

Heldt S (2009) On a new approach for self-optimizing control structure design. In: ADCHEM 2009, July 12-15, Istanbul, pp 807–811

Jäschke J, Skogestad S (2010) Self-optimizing control and NCO tracking in the context or real-time optimization. In: Proceedings of the 9th International Symposium on Dynamics and Control of Process Systems, DYCOPS 2010, July 5-7, Leuven, Belgium, pp 593–598

Kariwala V, Cao Y, Janardhanan S (2008) Local self-optimizing control with average loss minimization. Ind Eng Chem Res 47:1150–1158

Marlin T, Hrymak A (1997) Real-time operations optimization of continuous processes. In: Proceedings of CPC V, AIChE Symposium Series vol. 93., pp 156–164

Morari M, Stephanopoulos G, Arkun Y (1980) 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. analsysis of the optimizing control structures. AIChE Journal 26(2):220–232

Nocedal J, Wright SJ (2006) Numerical Optimization. Springer

Pistikopoulos E, Georgiadis M, Dua V (2007) Multiparametric programming. Wiley-VCH

Rojas JM (1999) Solving degenerate sparse polynomial systems faster. J Symbolic Computation 28:155–186

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

Srinivasan B, Biegler LT, Bonvin D (2008) Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function. Computers and Chemical Engineering 32:572–279

Sturmfels B (1994) On the newton polytope of the resultant. Journal of Algebraic Combinatorics 3:207–236

Sturmfels B (2002) Solving systems of polynomial equations