



Contents lists available at SciVerse ScienceDirect

Journal of Process Control

journal homepage: [www.elsevier.com/locate/jprocont](http://www.elsevier.com/locate/jprocont)

# Convex formulations for optimal selection of controlled variables and measurements using Mixed Integer Quadratic Programming

Ramprasad Yelchuru, Sigurd Skogestad\*

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

## ARTICLE INFO

### Article history:

Received 21 October 2011  
 Received in revised form 29 April 2012  
 Accepted 29 April 2012  
 Available online 31 May 2012

### Keywords:

Control structure selection  
 Self-optimizing control  
 Mixed Integer Quadratic Programming  
 Quadratic optimization  
 Plantwide control  
 Optimal measurement selection

## ABSTRACT

The appropriate selection of controlled variables is important for operating a process optimally in the presence of disturbances. Self-optimizing control provides a mathematical framework for selecting the controlled variables as combinations of measurements,  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , with the aim to minimize the steady state loss from optimal operation. In this paper, we present (i) a convex formulation to find the optimal combination matrix  $\mathbf{H}$  for a given measurement set and (ii) a Mixed-Integer Quadratic Programming (MIQP) methodology to select optimal measurement subsets that result in minimal loss. The methods presented in this paper are exact for quadratic problems with linear measurement relations. The MIQP methods can handle additional structural constraints compared to the branch and bound (BAB) methods reported in literature. The MIQP methods are evaluated on a toy test problem, an evaporator example, a binary distillation column example with 41 stages and a Kaibel column with 71 stages.

© 2012 Elsevier Ltd. All rights reserved.

## 1. Introduction

Control structure selection deals with the selection of controlled variables (CVs/outputs) and manipulated variables (MVs/inputs), and the pairings or interconnections of these variables [1,2]. A comprehensive review of input/output selection methods was provided by [3]. These input/output selection methods use desirable control system properties, (state, structural, input–output) controllability, achievable performance as criteria to arrive at CVs that are easy to control. However, these CV selection criteria fail to take into account more overall objectives, like economic profitability or cost ( $J$ ). The selection of control structure based on economics is stressed by Narraway and co-workers [4,5] for the effect of disturbances, but they do not formulate rules or procedures to select controlled variables.

In this paper, we consider the link between (economic) optimization and control as illustrated in Fig. 1. Self-optimizing control (SOC) [2] aims at achieving acceptable operation by maintaining selected CVs ( $\mathbf{c}$  in Fig. 1) at constant or slowly varying setpoints. The idea dates back to [6], who stated 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, and with it, the optimal operating condition”. Self-optimizing control makes use of the degrees of freedom in  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , which link the optimization and

control layers. There are three elements in the self-optimizing control approach. They are off-line static optimization to compute  $\mathbf{H}$  to find controlled variables  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , on-line slow time-scale RTO to compute  $\mathbf{c}_s$  and fast time-scale feedback control that adjusts  $\mathbf{u}$ . In this paper, we present the off-line static optimization approach to select  $\mathbf{H}$ , based on steady-state economics, but because the variables  $\mathbf{c}$  are controlled in the feedback layer, one gets much faster updates in the inputs  $\mathbf{u}$  than with the online slow time-scale RTO that computes  $\mathbf{c}_s$ . The dynamic performance of control structures obtained from self-optimizing control for various processes are reported [7–9]. The idea of self-optimizing control is to put as much optimization as possible into the control layer. That is, when there is a disturbance, we want the system “go in the right direction” on the fast time scale, and not have to wait for optimization layer (RTO) to take the optimal action, which may take a long time, since RTO needs to estimate the disturbances (e.g., using data reconciliation) before taking action.

For example, consider the process of cake baking. The (original) physical degree of freedom is the oven heat input ( $\mathbf{u} = Q$ ). However, baking the “optimal” cake is difficult when using the heat input directly for optimization (with the human as the RTO), and would require frequent changes in  $Q$ . However, we have available other measurements, including the oven temperature  $T$ . Consider the two candidate “measurements”

$$\mathbf{y} = [Q \quad T]^T$$

Clearly, the best variable to keep constant is  $T$ , so we choose  $\mathbf{c} = \mathbf{H}\mathbf{y} = h_{11}Q + h_{12}T = T$  as the controlled variable, that is, we choose  $\mathbf{H} = [0 \quad 1]$ . With a temperature controller (thermostat), we (the

\* Corresponding author. Tel.: +47 73 59 41 54; fax: +47 73 59 40 80.  
 E-mail address: [skoge@chemeng.ntnu.no](mailto:skoge@chemeng.ntnu.no) (S. Skogestad).

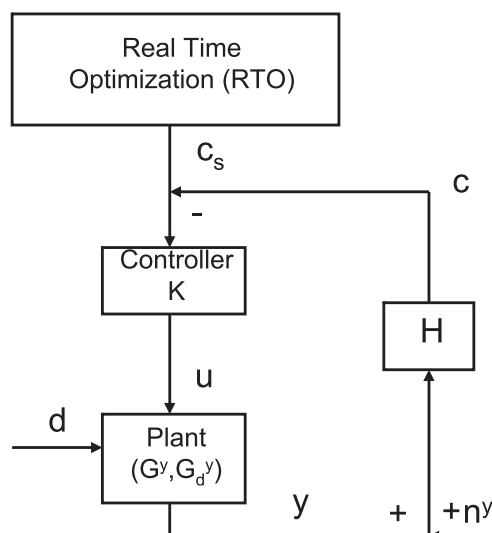


Fig. 1. Feedback implementation of optimal operation with separate layers for optimization and control [15,16]. The controller  $K$  could be any controller including MPC. Self-optimizing control deals with selection of the controlled variables  $\mathbf{c} = \mathbf{H}\mathbf{y}$ .

human RTO) may use the temperature set point ( $c_s$ ) as the optimization variable. Clearly, the introduction of the self-optimizing variable  $\mathbf{c} = T$ , simplifies the real-time optimization effort and requires less frequent changes than when using  $Q$ .

Instead of the two layer structure in Fig. 1, one could combine the layers and use real time optimization more directly by using a dynamic or steady state process model online to obtain an optimal input  $\mathbf{u}_{opt}(\mathbf{d})$  for a disturbance  $\mathbf{d}$ . However, such a centralized solution would be costly in terms of modeling, implementation and maintenance [10] and would normally operate at a slower time scale than the feedback layer in Fig. 1. A related alternative is optimizing controllers where the MVs ( $\mathbf{u}$ ) are updated directly to maintain the gradient of the Lagrangian function associated with the optimal process operation at zero [11]. Based on how the gradient is obtained, these methods are categorized as necessary conditions of optimality (NCO) tracking [11,12] or extremum seeking approaches [13,14]. The former approaches use analytical gradients, whereas the latter use operational data to estimate gradients. Although these optimizing controllers may be useful, slow speed of convergence caused by inaccurate gradient information usually makes these difficult to use in practice.

Importantly, self-optimizing control which deals with the selection of  $\mathbf{H}$  should not be viewed as an alternative to these other methods, including real time optimization or model predictive control (MPC), but rather as a complement, as illustrated in Fig. 1. By appropriate selection of the variables  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , we may reduce or eliminate the need for reoptimizing  $c_s$  independently of the approach we use for online optimization.

To quantify “acceptable operation” we introduce a scalar cost function  $J$  which should be minimized for optimal operation. In this paper, we assume that the (economic) cost mainly depends on the (quasi) steady-state behavior, which is a good assumption for most continuous plants in the process industry. When selecting  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , the cost function  $J$  is further assumed to be quadratic and the steady-state process model is assumed linear. Almost all steady-state unconstrained optimal operation problems can be approximated this way, usually by linearizing at the nominally optimal point. The scope of this paper is to provide systematic and good methods to select controlled variables (CVs,  $\mathbf{c} \in \mathbb{R}^{n_c}$ ) associated with the unconstrained steady state degrees of freedom ( $\mathbf{u} \in \mathbb{R}^{n_u}$ ) that minimize the loss,  $L(\mathbf{u}, \mathbf{d}) = J(\mathbf{u}, \mathbf{d}) - J_{opt}(\mathbf{d})$ , from

economically optimal operation. The number of selected CVs is equal to the number of steady state degrees of freedom ( $n_c = n_u$ ).

More specifically, the objective is to find a linear measurement combination,

$$\mathbf{c} = \mathbf{H}\mathbf{y} \tag{1}$$

such that control of these indirectly leads to acceptable operation with a small loss  $L(\mathbf{u}, \mathbf{d}) = J(\mathbf{u}, \mathbf{d}) - J_{opt}(\mathbf{d})$ , in spite of unknown disturbances,  $\mathbf{d}$ , and measurement noise (error),  $\mathbf{n}^y$ . If the original optimization problem is constrained, then we assume that all optimally active constraints are kept constant (controlled) and we consider the lower-dimensional unconstrained subspace. Depending on the disturbance range considered, there may be several constrained regions, and the procedure of finding  $\mathbf{H}$  needs to be repeated in each constrained region.

In this paper, we consider three problems related to finding optimal controlled variables,  $\mathbf{c} = \mathbf{H}\mathbf{y}$ ,

**Problem 1** Full  $\mathbf{H}$ , where the CVs are combinations of all measurements  $\mathbf{y}$ .

**Problem 2** Measurement selection problems, where some columns in  $\mathbf{H}$  are zero.

Case 2.1 Given subset of measurements.

Case 2.2 Optimal subset of measurements.

Case 2.3 Best individual measurements for decentralized control.

Compared to previous work [17], some additional restrictions are allowed for:

Case 2.4 Restriction on number of measurements from specified sections of the process.

Case 2.5 Addition of extra measurements to a given set.

**Problem 3** Structured  $\mathbf{H}$ , where specified elements in  $\mathbf{H}$  are zero; for example a block diagonal  $\mathbf{H}$ .

The problem of finding CVs as optimal measurement combinations (Problem 1) in the presence of disturbances and measurement noise was originally believed to be non convex and thus difficult to solve numerically [18], but later it has been shown that this problem may be reformulated as a quadratic optimization problem with linear constraints [19]. The same problem was solved using generalized singular value decomposition method [20,21]. However, the problems of selecting individual measurements or linear combinations of a subset of measurements as controlled variables (Problems 2 and 3) are more difficult because of their combinatorial nature.

To solve Problem 2, effective partial bidirectional branch and bound (PB<sup>3</sup>) methods have been developed [22] that exploit the monotonicity properties. However, these methods cannot be used directly in the presence of the restrictions in Cases 2.4 and 2.5 as the monotonicity is not guaranteed. In this paper, we propose a different method to solve Problem 2 by reformulating the minimum loss method problem as a Mixed-Integer Quadratic Programming (MIQP) problem. The MIQP formulations are simple and intuitive. The proposed MIQP formulations solve a convex quadratic optimization problem at each node in the search tree. These form a subclass of MIQP that are convex and hence these methods give globally optimal  $\mathbf{H}$  that results in measurement combinations as CVs. The additional restrictions Cases 2.4 and 2.5 can easily be handled with the MIQP based methods, whereas the branch and bound methods [22] would require further customization. Problem 3 is non-convex and cannot be solved by the methods presented in this paper and will be the topic of future work.

This paper is organized as follows: A self-contained summary of the minimum loss method formulation for SOC is presented in Section 2. The transformation of non-convex SOC problem to convex QP problem is discussed in Section 3 (Problem 1). The MIQP formulation for CV selection in SOC is presented in Section 4 (Problem 2). The evaluation of developed methods is performed on a toy

problem, on an evaporator example, on a binary distillation column example with 41 stages and on a 4-product Kaibel column with 71 stages and is discussed in Section 5. A discussion on Problem 3 is presented in Section 6. The conclusions from this work are discussed in Section 7.

## 2. Minimum loss method

The key idea in the self-optimizing framework of Skogestad and co-workers [23] is to minimize the loss ( $L = J - J_{opt}(\mathbf{d})$ ) from optimal operation when there are disturbances.

To find the minimum cost for a given disturbance  $J_{opt}(\mathbf{d})$ , we first find an expression for  $\mathbf{u}_{opt}(\mathbf{d})$ . We then evaluate the steady-state loss from this policy when  $\mathbf{u}$  is adjusted in a feedback fashion such that  $\mathbf{c} = \mathbf{H}\mathbf{y}$  is kept constant.

### 2.1. Problem formulation

#### 2.1.1. Classification of variables

- $\mathbf{u} \in \mathbb{R}^{n_u}$  – unconstrained steady state degrees of freedom (inputs) for optimization (it does not actually matter what they are as long as they form an independent set).
- $\mathbf{d} \in \mathbb{R}^{n_d}$  – disturbances, including parameter changes.
- $\mathbf{y} \in \mathbb{R}^{n_y}$  – all available measurements. The manipulated variables (MVs, often the same as the inputs  $\mathbf{u}$ ) are generally included in the measurement set  $\mathbf{y}$ . This will allow, for example, for simple control policies where the inputs are kept constant. Of course, the set  $\mathbf{y}$  can also include measured disturbances ( $\mathbf{d}_m$ , a subset of  $\mathbf{d}$ ).
- $\mathbf{n}^y$  – measurement noise (error) for  $\mathbf{y}$ ,  $\mathbf{y}_m = \mathbf{y} + \mathbf{n}^y$ .
- $\mathbf{c} \in \mathbb{R}^{n_c}$  where  $n_c = n_u$  – selected controlled variables  $\mathbf{c} = \mathbf{H}\mathbf{y}$ .

#### 2.1.2. Cost function

We consider an unconstrained optimization problem, where the objective is to adjust the input  $\mathbf{u}$  to minimize a quadratic steady-state process cost function

$$J(\mathbf{u}, \mathbf{d}) = J(\mathbf{u}^*, \mathbf{d}^*) + [\mathbf{J}_u^* \quad \mathbf{J}_d^*] \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{d} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{d} \end{bmatrix}^T \begin{bmatrix} \mathbf{J}_{uu}^* & \mathbf{J}_{ud}^* \\ \mathbf{J}_{ud}^{*T} & \mathbf{J}_{dd}^* \end{bmatrix} \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{d} \end{bmatrix} \quad (2)$$

Here  $\Delta \mathbf{u} = \mathbf{u} - \mathbf{u}^*$  and  $\Delta \mathbf{d} = \mathbf{d} - \mathbf{d}^*$  represent deviations from the nominal optimal point  $(\mathbf{u}^*, \mathbf{d}^*)$ .  $\mathbf{J}_u^*$  and  $\mathbf{J}_d^*$  are first derivatives of  $J$  with respect to  $\mathbf{u}$  and  $\mathbf{d}$ ,  $\mathbf{J}_{uu}^*$ ,  $\mathbf{J}_{ud}^*$  and  $\mathbf{J}_{dd}^*$  are second derivatives of  $J$  with respect to  $\mathbf{u}$ ,  $\mathbf{u}$  and  $\mathbf{d}$ , and  $\mathbf{d}$ , respectively at  $(\mathbf{u}^*, \mathbf{d}^*)$ . The nominal point is assumed to be optimal, which implies that  $\mathbf{J}_u^* = \mathbf{0}$ . To further simplify notation, we assume that the variables have been shifted so that the nominal optimal point is zero  $(\mathbf{u}^*, \mathbf{d}^*) = (0, 0)$  and also  $\mathbf{y}^* = \mathbf{0}$ , then we have  $\mathbf{u} = \Delta \mathbf{u}$ ,  $\mathbf{d} = \Delta \mathbf{d}$  and  $\mathbf{y} = \Delta \mathbf{y}$ . From the derivation below, we find that the values of  $\mathbf{J}_d^*$  and  $\mathbf{J}_{dd}^*$  are not needed for finding the optimal  $\mathbf{H}$ , because they do not affect the optimal input  $\mathbf{u}$ .

A special case of (2) is indirect control, which is further studied for a distillation column in Example 4, where  $\mathbf{y}_1$  are the primary variables. Here, the cost function is

$$J = (\mathbf{y}_1 - \mathbf{y}_{1s})^T \mathbf{W}_1^T \mathbf{W}_1 (\mathbf{y}_1 - \mathbf{y}_{1s}) \quad (3a)$$

where  $\mathbf{W}_1$  is a weighting matrix,  $\mathbf{y}_{1s}$  are set points for  $\mathbf{y}_1$ , and with a linear model for  $\mathbf{y}_1$

$$\mathbf{y}_1 = \mathbf{G}_1^y \mathbf{u} + \mathbf{G}_{d_1}^y \mathbf{d} \quad (3b)$$

where  $\mathbf{G}_1^y$  and  $\mathbf{G}_{d_1}^y$  are steady state gains, further we get

$$\mathbf{J}_{uu} = \mathbf{G}_1^{yT} \mathbf{W}_1^T \mathbf{W}_1 \mathbf{G}_1^y, \quad \mathbf{J}_{ud} = \mathbf{G}_1^{yT} \mathbf{W}_1^T \mathbf{W}_1 \mathbf{G}_{d_1}^y \quad (3c)$$

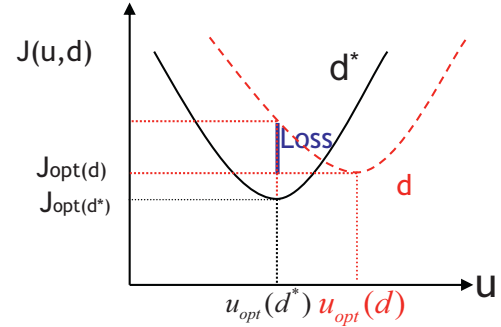


Fig. 2. Illustration of loss by keeping input  $\mathbf{u}$  constant at  $\mathbf{u} = \mathbf{u}_{opt}(\mathbf{d}^*)$  when there is a disturbance  $\mathbf{d}$ .

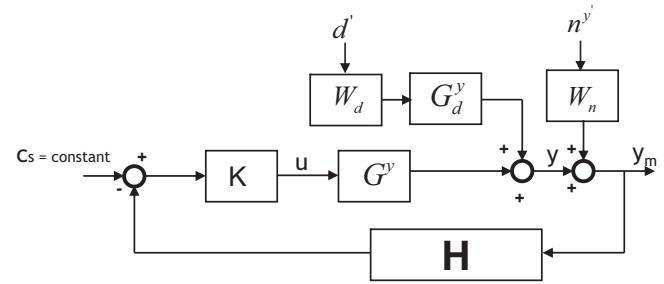


Fig. 3. Feedback diagram.

#### 2.1.3. Measurement model

A linear steady-state model is assumed for the effect of  $\mathbf{u}$  and  $\mathbf{d}$  on the measurements  $\mathbf{y}$

$$\mathbf{y} = \mathbf{G}^y \mathbf{u} + \mathbf{G}_d^y \mathbf{d} = \tilde{\mathbf{G}}^y \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} \quad (4)$$

In Fig. 1,  $\mathbf{G}^y$  and  $\mathbf{G}_d^y$  are transfer functions, but in this paper only steady-state gains in (4) are used for selecting  $\mathbf{H}$ .

#### 2.1.4. Further assumptions

- Any active constraints are controlled and  $\mathbf{u}$  spans the remaining unconstrained subspace.
- We want to find as many controlled variables  $\mathbf{c}$  as there are degrees of freedom, that is,  $n_c = \dim(\mathbf{c}) = \dim(\mathbf{u}) = n_u$ . Then  $\mathbf{H}\mathbf{G}^y$  is a square  $n_u \times n_u$  matrix.
- We need at least as many independent measurements  $\mathbf{y}$  as there are degrees of freedom ( $\text{rank}(\mathbf{G}^y) = n_u$ ) to get offset free control of all CVs ( $\mathbf{c}$ ). This requires  $n_y \geq n_u = n_c$ .
- We write  $\mathbf{d} = \mathbf{W}_d \mathbf{d}'$  where  $\mathbf{W}_d$  is a diagonal matrix giving the expected magnitude of each disturbance and  $\mathbf{d}'$  is of unit magnitude (see below for further definition of “unit magnitude”).
- Similarly,  $\mathbf{n}^y = \mathbf{W}_{n^y} \mathbf{n}^y$  where  $\mathbf{W}_{n^y}$  is a diagonal matrix with the magnitude of the noise for each measurement, and the vector  $\mathbf{n}^y$  is of unit magnitude (see below).

#### 2.1.5. Problem

For any disturbance  $\mathbf{d}$ , having inputs  $\mathbf{u}$  other than  $\mathbf{u}_{opt}(\mathbf{d})$  will result in a loss. For example, keeping the inputs  $\mathbf{u}$  constant at  $\mathbf{u}_{opt}(\mathbf{d}^*)$  when there is a disturbance  $\mathbf{d}$  will result in a loss as illustrated in Fig. 2. In this paper, we use a sub-optimal policy, which is to adjust inputs  $\mathbf{u}$  in a feedback fashion (see Figs. 1 and 3) to keep the measured controlled variables  $\mathbf{c}_m$  at a constant set point  $\mathbf{c}_s = 0$ . Mathematically, we have

$$\mathbf{c}_m = \underbrace{\mathbf{H}(\mathbf{y} + \mathbf{n}^y)}_{\mathbf{y}_m} = \mathbf{c}_s = 0 \quad (5)$$

With this policy, there are two problems of interest. First, to find the “magnitude” of the loss,  $L = J(\mathbf{u}, \mathbf{d}) - J_{\text{opt}}(\mathbf{d})$ , for a given  $\mathbf{H}$  (see solution in Section 2.2.5) and second to find the optimal  $\mathbf{H}$  with a minimum loss (see Theorem 1 in Section 2.2.7) for the expected  $\mathbf{d}'$  and  $\mathbf{n}^y$ , when  $\mathbf{u}$  is adjusted such that  $\mathbf{c}_m = 0$  in (5) is satisfied.

The “magnitude” of the loss and the “unit magnitude” of the expected  $\mathbf{d}'$  and  $\mathbf{n}^y$  still need to be defined. Two possibilities are considered.

- Worst-case loss,  $L_{wc}$ , when the combined normalization vectors for disturbances and measurement noise have 2-norm less than 1,

$$\left\| \begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^y \end{bmatrix} \right\|_2 \leq 1 \quad (6)$$

- Average or expected loss,  $L_{avg} = \mathcal{E}(L)$ , for a normal distributed set

$$\begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^y \end{bmatrix} \in \mathcal{N}(0, 1) \quad (7)$$

$\mathcal{E}(\cdot)$  is expectation operator.

It is sometimes argued that the worst-case loss is not likely to occur, but this is not really true in this case since we use the combined 2-norm for disturbances and noise in (6). This means that the “unlikely” combination with all  $\mathbf{d}'$ 's and  $\mathbf{n}^y$ 's being 1 at the same time will not occur. This is discussed in more detail in Appendix of [18].

## 2.2. Solution to minimum loss problem

The objective is to derive the solution to the above problem. This solution has previously been called the “exact local method” [18].

### 2.2.1. Expression for $\mathbf{u}_{\text{opt}}(\mathbf{d})$

We first want to find the optimal input  $\mathbf{u}$  for a given disturbance  $\mathbf{d}$ . Expanding the gradient  $\mathbf{J}_u$  around the nominal optimal point  $(\mathbf{u}^*, \mathbf{d}^*) = (0, 0)$  gives

$$\mathbf{J}_u(\mathbf{u}, \mathbf{d}) = \underbrace{\mathbf{J}_u^*(\mathbf{u}^*, \mathbf{d}^*)}_{=0} + \mathbf{J}_{uu}^* \mathbf{u} + \mathbf{J}_{ud}^* \mathbf{d} \quad (8)$$

where  $\mathbf{J}_u^*(\mathbf{u}^*, \mathbf{d}^*) = 0$  because the nominal point is assumed to be optimal. We assume that we change the input to remain optimal, i.e. we have  $\mathbf{u} = \mathbf{u}_{\text{opt}}(\mathbf{d})$  and  $\mathbf{J}_u(\mathbf{u}, \mathbf{d}) = 0$ , and we get

$$\mathbf{u}_{\text{opt}} = -\mathbf{J}_{uu}^{-1} \mathbf{J}_{ud} \mathbf{d} \quad (9)$$

Note that we are considering a quadratic problem (2), where the Hessian matrices are assumed constant, i.e.  $\mathbf{J}_{uu} = \mathbf{J}_{uu}^*$  and  $\mathbf{J}_{ud} = \mathbf{J}_{ud}^*$ .

### 2.2.2. Expression for the loss $L$ in terms of $\mathbf{u} - \mathbf{u}_{\text{opt}}(\mathbf{d})$

Consider a given disturbance  $\mathbf{d}$  and a non-optimal input  $\mathbf{u}$ . A second order Taylor's expansion of the cost  $J$  around the “moving” optimum point,  $\mathbf{u}_{\text{opt}}(\mathbf{d})$ , gives

$$\begin{aligned} J(\mathbf{u}, \mathbf{d}) &= \underbrace{J(\mathbf{u}_{\text{opt}}(\mathbf{d}), \mathbf{d})}_{J_{\text{opt}}(\mathbf{d})} + \underbrace{\mathbf{J}_{u,\text{opt}}}_{=0}(\mathbf{u} - \mathbf{u}_{\text{opt}}(\mathbf{d})) \\ &\quad + \frac{1}{2}(\mathbf{u} - \mathbf{u}_{\text{opt}}(\mathbf{d}))^T \mathbf{J}_{uu,\text{opt}}(\mathbf{u} - \mathbf{u}_{\text{opt}}(\mathbf{d})) \end{aligned} \quad (10)$$

Note that for a truly quadratic problem, this is an exact expression and  $\mathbf{J}_{uu,\text{opt}} = \mathbf{J}_{uu}^* = \mathbf{J}_{uu}$ . Because we are expanding around an optimal point  $\mathbf{J}_{u,\text{opt}} = 0$  and we get the following expression for the loss

$$L(\mathbf{u}, \mathbf{d}) = J(\mathbf{u}, \mathbf{d}) - J_{\text{opt}}(\mathbf{d}) = \frac{1}{2} \mathbf{z}^T \mathbf{z} = \frac{1}{2} \|\mathbf{z}\|_2^2 \quad (11)$$

where we have introduced

$$\mathbf{z} \triangleq \mathbf{J}_{uu}^{1/2}(\mathbf{u} - \mathbf{u}_{\text{opt}}(\mathbf{d})) \quad (12)$$

This simple expression for the loss is a key result that allows us to end up with a convex optimization problem.

### 2.2.3. Optimal sensitivities

Note from (9) that we can write  $\mathbf{u}_{\text{opt}} = \mathbf{F}^u \mathbf{d}$  where  $\mathbf{F}^u = -\mathbf{J}_{uu}^{-1} \mathbf{J}_{ud}$ . More generally, we can write

$$\mathbf{y}_{\text{opt}} = \mathbf{F} \mathbf{d} \quad (13)$$

where  $\mathbf{F}$  is the optimal sensitivity of the outputs (measurements) with respect to the disturbances. Here,  $\mathbf{F}$  can be obtained using (4) and (9),

$$\mathbf{y}_{\text{opt}} = \mathbf{G}^y \mathbf{u}_{\text{opt}} + \mathbf{G}_d^y \mathbf{d} = (-\mathbf{G}^y \mathbf{J}_{uu}^{-1} \mathbf{J}_{ud} + \mathbf{G}_d^y) \mathbf{d}$$

that is,

$$\mathbf{F} = (-\mathbf{G}^y \mathbf{J}_{uu}^{-1} \mathbf{J}_{ud} + \mathbf{G}_d^y) \quad (14)$$

However, (14) is not generally a robust way to obtain  $\mathbf{F}$ , for example  $\mathbf{J}_{uu}$ ,  $\mathbf{J}_{ud}$  can be difficult to obtain numerically, and taking the difference in (14) can also be unreliable numerically. Thus, for practical use it is usually better to obtain  $\mathbf{F}$  directly from its definition,  $\mathbf{F} = (d\mathbf{y}_{\text{opt}}/d\mathbf{d})$ . This typically involves numerical reoptimization for each disturbance.

### 2.2.4. The loss $L$ as a function of disturbances and noise

We present the derivation of the main result [18]. We start from the loss expression in (11),  $L = (1/2) \|\mathbf{z}\|_2^2$  where  $\mathbf{z} = \mathbf{J}_{uu}^{1/2}(\mathbf{u} - \mathbf{u}_{\text{opt}})$ . We want to write  $\mathbf{z}$  as a function of  $\mathbf{d}$  and  $\mathbf{n}^y$ , given that the input  $\mathbf{u}$  should be adjusted to satisfy (5). We start by writing  $\mathbf{u} - \mathbf{u}_{\text{opt}}$  as a function of  $\mathbf{c} - \mathbf{c}_{\text{opt}}$ . We have  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , so

$$\mathbf{c} = \mathbf{H}\mathbf{y} = \mathbf{H}\mathbf{G}^y \mathbf{u} + \mathbf{H}\mathbf{G}_d^y \mathbf{d}$$

$$\mathbf{c}_{\text{opt}} = \mathbf{H}\mathbf{y}_{\text{opt}} = \mathbf{H}\mathbf{G}^y \mathbf{u}_{\text{opt}} + \mathbf{H}\mathbf{G}_d^y \mathbf{d}$$

Thus,  $\mathbf{c} - \mathbf{c}_{\text{opt}} = \mathbf{H}\mathbf{G}^y(\mathbf{u} - \mathbf{u}_{\text{opt}})$ , or

$$(\mathbf{u} - \mathbf{u}_{\text{opt}}) = (\mathbf{H}\mathbf{G}^y)^{-1}(\mathbf{c} - \mathbf{c}_{\text{opt}}) \quad (15a)$$

where  $\mathbf{H}\mathbf{G}^y$  is the square gain matrix from the inputs  $\mathbf{u}$  to the selected controlled variables  $\mathbf{c}$ .

The next step is to express  $(\mathbf{c} - \mathbf{c}_{\text{opt}})$  as a function of  $\mathbf{d}$  and  $\mathbf{n}^y$ . From (13) we have that

$$\mathbf{c}_{\text{opt}} = \mathbf{H}\mathbf{F}\mathbf{d} \quad (15b)$$

From (5) we have that  $\mathbf{H}(\mathbf{y} + \mathbf{n}^y) = \mathbf{c}_s$  (constant), or

$$\mathbf{c} = \mathbf{H}\mathbf{y} = -\mathbf{H}\mathbf{n}^y + \mathbf{c}_s \quad (15c)$$

Here,  $\mathbf{c}_s = 0$ , since we assume the nominal point is optimal. Since the signs for  $\mathbf{n}^y$  and  $\mathbf{d}$  do not matter for the expressions we derive below (from (6) we can have both positive and negative changes), we can write

$$\begin{aligned} \mathbf{u} - \mathbf{u}_{\text{opt}} &= (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}(\mathbf{F}\mathbf{d} + \mathbf{n}^y) = (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}(\mathbf{F}\mathbf{W}_d \mathbf{d}' + \mathbf{W}_{ny} \mathbf{n}^y) \\ &= (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}\mathbf{Y} \begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^y \end{bmatrix} \end{aligned} \quad (15d)$$

where we have introduced

$$\mathbf{Y} = [\mathbf{F}\mathbf{W}_d \quad \mathbf{W}_{ny}] \quad (16)$$

Note that  $\mathbf{W}_d$  and  $\mathbf{W}_{ny}$  are usually diagonal matrices, representing the magnitude of the disturbances and measurement noises, respectively.

In summary, we have derived that for the given normalized disturbances  $\mathbf{d}'$  and for the given normalized measurement noises  $\mathbf{n}^y$  the loss is given by [18]

$$L = \frac{1}{2} \mathbf{z}^T \mathbf{z} \quad (17)$$

where

$$\mathbf{z} = \mathbf{J}_{uu}^{1/2} (\mathbf{u} - \mathbf{u}_{\text{opt}}) = \underbrace{\mathbf{J}_{uu}^{1/2} (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}\mathbf{Y}}_{\mathbf{M}(\mathbf{H})} \begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^y \end{bmatrix} \quad (18)$$

### 2.2.5. Worst-case and average loss for a given $\mathbf{H}$ (analysis using loss method)

The above expressions give the loss for a given  $\mathbf{d}'$  and  $\mathbf{n}^y$ , but the goal is to find the “magnitude” of the loss  $L$  for the expected set for example as given in (6). Here “magnitude” can be defined in different ways, see (6) and (7), and for a given  $\mathbf{H}$  the worst-case loss [18] and average expected loss [24] are given by

$$L_{wc}(\mathbf{H}) = \frac{1}{2} \bar{\sigma}(\mathbf{M})^2 \quad (19)$$

$$L_{avg}(\mathbf{H}) = \varepsilon(L) = \frac{1}{2} \|\mathbf{M}\|_F^2 \quad (20)$$

where

$$\mathbf{M}(\mathbf{H}) = \mathbf{J}_{uu}^{1/2} (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}\mathbf{Y} \quad (21)$$

Here  $\bar{\sigma}(\mathbf{M})$  denotes the maximum singular value (induced 2-norm) of the matrix  $\mathbf{M}(\mathbf{H})$ , and  $\|\mathbf{M}\|_F = \sqrt{\sum_{i,j} \mathbf{M}_{ij}^2}$  denotes the Frobenius norm of the matrix  $\mathbf{M}$ . Use of the norm of  $\mathbf{M}$  to analyze the loss is known as the “exact local method” [18]. Note that these loss expressions are for a given matrix  $\mathbf{H}$ .

*Comment:* A uniform distribution for  $\mathbf{d}'$  and  $\mathbf{n}^y$  is sometimes assumed, resulting in an average loss  $(1/(6(n_y + n_d))) \|\mathbf{M}\|_F^2$  [24]. However, as discussed in Section 6.2, this is not meaningful from an engineering point of view.

### 2.2.6. Null space method and maximum gain rule

Two special methods for analyzing or finding  $\mathbf{H}$  can be derived from the expression for  $\mathbf{H}$  in (21). First, the null space method of selecting  $\mathbf{H}$  such that  $\mathbf{H}\mathbf{F} = \mathbf{0}$  [25] follows if we neglect measurement noise such that  $\mathbf{Y} = [\mathbf{F}\mathbf{W}_d \quad \mathbf{0}]$ , where  $\mathbf{0}$  is zero matrix of  $n_y \times n_y$  size, and assume that we have enough measurements to make  $\mathbf{H}\mathbf{F} = \mathbf{0}$ . Second, the approximate maximum gain rule [23] of maximizing the norm of  $\mathbf{S}_1 \mathbf{H}\mathbf{G}^y \mathbf{S}_2$  follows from (21) if we select the scaling factors as  $\mathbf{S}_2 = \mathbf{J}_{uu}^{1/2}$  and the appropriate  $\mathbf{S}_1$  as a diagonal matrix with the elements of  $\mathbf{S}_1^{-1}$  equal to the expected optimal variation in each  $\mathbf{c}$  variable (the norm of the corresponding rows in  $\mathbf{H}\mathbf{Y}$ ).

### 2.2.7. Loss method for finding optimal $\mathbf{H}$

The objective of this paper is to find methods for obtaining the optimal  $\mathbf{H}$  by minimizing either the worst-case loss (19) or the average loss (20). Fortunately, [24] proves that the  $\mathbf{H}$  that minimizes the average loss in Eq. (20) is super optimal, in the sense that the same  $\mathbf{H}$  minimizes the worst case loss in (19). Hence, only minimization of the Frobenius norm in (20) is considered in the rest of the paper. Note that square does not effect the optimal solution and can be omitted. In summary, the problem is to find the combination matrix  $\mathbf{H}$  that minimizes  $\|\mathbf{M}\|_F$ :

**Theorem 1** (Minimum loss method [19]). *To minimize the average and worst case loss,  $L_{avg}(\mathbf{H})$  and  $L_{wc}(\mathbf{H})$ , for expected combined disturbances and noise, find the  $\mathbf{H}$  that solves the problem*

$$\min_{\mathbf{H}} \left\| \mathbf{J}_{uu}^{1/2} (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}\mathbf{Y} \right\|_F \quad (22)$$

where  $\mathbf{Y} = [\mathbf{F}\mathbf{W}_d \quad \mathbf{W}_{ny}]$ .

The objective in (22) is to find the non-square  $n_c \times n_y$  matrix  $\mathbf{H}$ .

Here,  $\mathbf{H}$  may have a specified structure and we consider the three problems mentioned in Section 1. For the full  $\mathbf{H}$  case (Problem 1), it may be recast as a convex optimization problem as discussed in Section 3. For the measurement selection problem (Problem 2), where some columns in  $\mathbf{H}$  are zero, convex formulations in each MIQP node are derived in Section 4.

### 3. Convex formulations of minimum loss method (Problem 1)

We here consider the standard “full”  $\mathbf{H}$  case with no restriction on the structure of the matrix  $\mathbf{H}$  (Problem 1), that is we want to find optimal combination of all the measurements.

**Theorem 2** (Convex reformulation for full  $\mathbf{H}$  case [19]). *The problem in Eq. (22) may seem non-convex, but for the standard case where  $\mathbf{H}$  is a “full” matrix (with no structural constraints), it can be reformulated as a convex constrained quadratic programming problem*

$$\begin{aligned} \min_{\mathbf{H}} & \|\mathbf{H}\mathbf{Y}\|_F \\ \text{s.t. } & \mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2} \end{aligned} \quad (23)$$

**Proof.** From the original problem in Eq. (22), we have that the optimal solution  $\mathbf{H}$  is non-unique because if  $\mathbf{H}$  is a solution then  $\mathbf{H}_1 = \mathbf{D}\mathbf{H}$  is also a solution for any non-singular matrix  $\mathbf{D}$  of size  $n_c \times n_c$ . This follows because

$$\mathbf{J}_{uu}^{1/2} (\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{H}\mathbf{Y} = \mathbf{J}_{uu}^{1/2} (\mathbf{D}\mathbf{H}\mathbf{G}^y)^{-1} \mathbf{D}\mathbf{H}\mathbf{Y} = \mathbf{J}_{uu}^{1/2} (\mathbf{H}_1\mathbf{G}^y)^{-1} \mathbf{H}_1\mathbf{Y}$$

One implication is that we can freely choose  $\mathbf{G} = \mathbf{H}\mathbf{G}^y$ , which is a  $n_c \times n_c$  matrix representing the effect of  $\mathbf{u}$  on  $\mathbf{c}$  ( $\mathbf{c} = \mathbf{G}\mathbf{u}$ ). Thus, in (22) we may use the non-uniqueness of  $\mathbf{H}$  to set the first part of the expression equal to the identity matrix, which is equivalent to setting  $\mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2}$ . This must be added as a constraint in the optimization as shown in (23).  $\square$

**Theorem 3** (Analytical solution [19]). *For a “full”  $\mathbf{H}$  in (22) and (23), an analytical solution is*

$$\mathbf{H}^T = (\mathbf{Y}\mathbf{Y}^T)^{-1} \mathbf{G}^y (\mathbf{G}^y (\mathbf{Y}\mathbf{Y}^T)^{-1} \mathbf{G}^y)^{-1} \mathbf{J}_{uu}^{1/2} \quad (24)$$

*Comment:* We also require that  $\mathbf{Y}\mathbf{Y}^T$  is full rank, which is always satisfied if we have nonzero measurement noise.

**Theorem 4** (Simplified analytical solution (new result)). *For a full  $\mathbf{H}$ , another analytical solution to (22) is*

$$\mathbf{H}^T = (\mathbf{Y}\mathbf{Y}^T)^{-1} \mathbf{G}^y \mathbf{Q}_1 \quad (25)$$

where  $\mathbf{Q}_1$  is any non-singular matrix of  $n_c \times n_c$ , for example  $\mathbf{Q}_1 = \mathbf{I}$ .

**Proof.** This follows trivially from Theorems 2 and 3, since if  $\mathbf{H}^T$  is a solution then so is

$$\mathbf{H}_1^T = \mathbf{H}^T \mathbf{D}^T \text{ and we simply select}$$

$$\mathbf{D}^T = \mathbf{J}_{uu}^{-1/2} (\mathbf{G}^y (\mathbf{Y}\mathbf{Y}^T)^{-1} \mathbf{G}^y) \mathbf{Q}_1$$

which is a  $n_c \times n_c$  matrix.  $\square$

**Corollary 1. Important insight (new result).** *Theorem 4 gives the very important insight that  $\mathbf{J}_{uu}$  is not needed for finding the optimal full  $\mathbf{H}$  in (22) and (23).*

This means that in (22) we can replace  $\mathbf{J}_{uu}^{1/2}$  by any non-singular matrix  $\mathbf{Q}$  and still get an optimal  $\mathbf{H}$ . This can simplify practical calculations, because  $\mathbf{J}_{uu}$  may be difficult to obtain numerically because it involves the second derivative and because  $\mathbf{Q}$  may be in some cases be selected for numerical reasons. On the other hand, we have that  $\mathbf{F}$ , which enters in  $\mathbf{Y}$ , is relatively straightforward

to obtain numerically [7,9], because it only needs first derivative,  $\mathbf{F}=(dy_{\text{opt}}/d\mathbf{d})$ , as mentioned earlier. Although  $\mathbf{J}_{uu}$  is not needed for finding the optimal  $\mathbf{H}$ , it would be required for finding a numerical value for the loss, and it is needed if  $\mathbf{H}$  is structured (Problems 2 and 3) as discussed below.

**Vectorized QP formulation:** As the numerical software packages, such as Matlab, cannot deal with the matrix formulations, the problem (23) is vectorized (see Appendix A). First, the decision matrix

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n_y} \\ h_{21} & h_{22} & \dots & h_{2n_y} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_u 1} & h_{n_u 2} & \dots & h_{n_u n_y} \end{bmatrix}$$

is vectorized along the rows of  $\mathbf{H}$  to form a long vector

$$\mathbf{h}_\delta = [h_{11} \dots h_{1n_y} \ h_{21} \dots h_{2n_y} \dots h_{n_u 1} \dots h_{n_u n_y}]^T \in \mathbb{R}^{n_u n_y \times 1}$$

The equivalent QP is then formulated as

$$\begin{aligned} \min_{\mathbf{h}_\delta} \mathbf{h}_\delta^T \mathbf{F}_\delta \mathbf{h}_\delta \\ \text{s.t. } \mathbf{G}_\delta^T \mathbf{h}_\delta = \mathbf{j}_\delta \end{aligned} \quad (26)$$

where  $\mathbf{h}_\delta \in \mathbb{R}^{n_u n_y \times 1}$ ,  $\mathbf{j}_\delta \in \mathbb{R}^{n_u n_y \times 1}$ ,  $\mathbf{G}_\delta^T \in \mathbb{R}^{n_u n_y \times n_u n_y}$ ,  $\mathbf{F}_\delta \in \mathbb{R}^{n_u n_y \times n_u n_y}$ .

#### 4. Globally optimal MIQP formulations (Problem 2)

We here consider the optimal measurement selection of finding the optimal  $\mathbf{H}$  with some zero columns (Problem 2). To address the measurement selection, we introduce a binary variable  $\sigma_j \in \{0, 1\}$  to complement  $j$ th measurement ( $j$ th column in  $\mathbf{H}$ ). If measurement  $j$  is present in the selected measurements, then  $\sigma_j = 1$  and  $j$ th column in  $\mathbf{H}$  may have non-zero elements, otherwise  $\sigma_j = 0$  and  $j$ th column in  $\mathbf{H}$  has only zero elements. The binary variables column vector for  $n_y$  candidate measurements is denoted as  $\boldsymbol{\sigma}_\delta = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_{n_y}]^T$ . The restrictions on elements in  $\mathbf{H}$  based on the presence or not of the  $j$ th candidate measurement are incorporated as mixed integer constraints. Overall, the idea in optimal measurement selection is to use the quadratic programming formulation in Theorem 2, and add additional mixed integer constraints to deal with the measurement selection.

##### 4.1. Optimal measurement selection

The mixed integer constraints on the columns in  $\mathbf{H}$  are formulated using the standard big- $m$  approach used in MIQP formulations (27c) [26] and are added to (26). The constraints on the binary variables can be written in the form

$$\mathbf{P}\boldsymbol{\sigma}_\delta = \mathbf{s}$$

For example, in order to select  $n$  optimal measurements out of  $n_y$  measurements, we have  $\sum_j \sigma_j = n$ , which can be written in this form with  $\mathbf{P} = \mathbf{1}_{1 \times n_y}^T$ , and  $\mathbf{s} = n$ , where  $\mathbf{1}$  is a column vector of ones.

Starting from the vectorized formulation in (26), we then have the important result that the generalized MIQP problem in the decision variables  $\mathbf{h}_\delta$  and  $\boldsymbol{\sigma}_\delta$  with big- $m$  constraints becomes

$$\begin{aligned} \min_{\mathbf{h}_\delta, \boldsymbol{\sigma}_\delta} \mathbf{h}_\delta^T \mathbf{F}_\delta \mathbf{h}_\delta \\ \text{s.t. } \mathbf{G}_\delta^T \mathbf{h}_\delta = \mathbf{j}_\delta \end{aligned} \quad (27a)$$

$$\mathbf{P}\boldsymbol{\sigma}_\delta = \mathbf{s} \quad (27b)$$

$$\begin{bmatrix} -m \\ -m \\ \vdots \\ -m \end{bmatrix} \boldsymbol{\sigma}_j \leq \begin{bmatrix} h_{1j} \\ h_{2j} \\ \vdots \\ h_{n_u j} \end{bmatrix} \leq \begin{bmatrix} m \\ m \\ \vdots \\ m \end{bmatrix} \boldsymbol{\sigma}_j, \forall j \in 1, 2, \dots, n_y \quad (27c)$$

where  $\mathbf{h}_\delta = [h_{11} \dots h_{1n_y} \ h_{21} \dots h_{2n_y} \dots h_{n_u 1} \dots h_{n_u n_y}]^T \in \mathbb{R}^{n_u n_y \times 1}$ ;  $\boldsymbol{\sigma}_\delta = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_{n_y}]^T$ ;  $\sigma_j \in \{0, 1\}$ . The dimension of matrix  $\mathbf{P}$  varies based on the integer constraints we impose, if we impose  $k$  number of integer constraints then  $\mathbf{P}$  will have a dimension of  $k \times n_y$ . The constraints in (27c) is the standard big- $m$  approach that we used to make the  $j$ th column of  $\mathbf{H}$  zero when  $\sigma_j = 0$  and at the same time to bound the decision variables in  $\mathbf{H}$ . The  $m$  value should be chosen small to reduce the computational time, but it should be sufficiently large to avoid that it becomes an active constraint. Selecting an appropriate  $m$  is problem dependent and appropriate selection of  $m$  can become an iterative method and can increase the computational intensiveness of the big- $m$  based MIQP formulations. In such cases, one can use indicator constraints in MIQP problem to set the columns in  $\mathbf{H}$  directly to zero, when  $\sigma_j = 0$ . This can be done by replacing the constraints in (27c) with indicator constraints as

$$\text{indicator constraints: } \sigma_j = 0 \Rightarrow \begin{bmatrix} h_{1j} \\ h_{2j} \\ \vdots \\ h_{n_u j} \end{bmatrix} = \mathbf{0}_{n_u \times 1}, \forall j \in 1, 2, \dots, n_y \quad (28)$$

where  $\mathbf{0}$  is a column vector of zeros. For MIQP, theoretically indicator constraint approach (28) would be faster than using big- $m$  approach (27c). This is because in MIQP, indicator constraint approach (28) solves an equality constrained QP at each node, whereas big- $m$  approach (27c) solves an inequality constrained QP.

For the solution of the MIQP problem with (27c) or (28), Theorem 2 applies. This statement is proved as follows: At each node in the MIQP search tree, we could use Theorem 2. This will preserve the loss ordering between different nodes in the MIQP search tree, because in Theorem 2, meeting the constraint  $\mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2}$  implies  $\mathbf{J}_{uu}^{1/2}(\mathbf{H}\mathbf{G}^y)^{-1} = \mathbf{I}$  and the loss value in (22) is equal to  $\|\mathbf{H}\mathbf{Y}\|_F$ .

##### 4.2. Specific cases

We consider five specific cases of Problem 2 and show how they can be solved using the MIQP formulation in (27). The integer constraint in (27b) is modified for each case. Note that Cases 2.1, 2.2 and 2.3 can alternatively be solved using the branch and bound approaches [17]. However, Cases 2.4 and 2.5 can only be solved using our MIQP formulation.

**Case 2.1 Given subset of measurements.** For example, assume we have two inputs and 5 measurements of which we will not use measurements 1 and 3, then  $\mathbf{H} = \begin{bmatrix} 0 & h_{12} & 0 & h_{14} & h_{15} \\ 0 & h_{22} & 0 & h_{24} & h_{25} \end{bmatrix}$ . The resulting constraints can be written in the form in (27b) with

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This is a very simple case, and we may use Theorem 2, which implies that  $\mathbf{J}_{uu}$  is not needed. The fact that Theorem 2 hold is quite obvious since it corresponds to simply deleting some

measurements (deleting rows in  $\mathbf{G}^y$  and  $\mathbf{Y}$ ), and keeping  $\mathbf{H}$  full for the remaining measurements.

**Case 2.2 Optimal subset of measurements.** Here the objective is to select a certain number ( $n$ ) of measurements (i.e.  $n_y - n$  columns in  $\mathbf{H}$  are zero). The constraint in the binary variables is

$$\sum_{j=1}^{n_y} \sigma_j = n \quad (29)$$

which can be written in the form in (27b) with

$$\mathbf{P} = \mathbf{1}_{1 \times n_y}^T, \quad \mathbf{s} = n$$

where  $\mathbf{1}$  is a column vector of ones.

**Case 2.3 Best individual measurements for decentralized control.** This is the case where we want to select  $n = n_c$  measurements, which is the minimum feasible number of measurements, if we want offset free control of  $\mathbf{c} = \mathbf{H}\mathbf{y}$ . For example, one candidate  $\mathbf{H}$  is

$$\mathbf{H} = \begin{bmatrix} h_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{24} & 0 \end{bmatrix} \quad (30)$$

The constraints to be used in (27b) are  $\sum_{j=1}^{n_y} \sigma_j = n_u = n_c$  and in addition the off diagonal elements for the selected  $n_c$  measurements should be zero (for this candidate  $\mathbf{H}$  the selected measurements are 1, 4 and the off-diagonal elements  $h_{21}$  and  $h_{14}$  are zero).

Fortunately, Theorem 2 which requires  $\mathbf{H}$  to be a full matrix may be used at each node in the MIQP, because the last restriction (off-diagonal elements are zero) may be omitted. The reason is that we can first find the optimal measurement subset for this selected  $n_c$

measurements, for example,  $\mathbf{H} = \begin{bmatrix} h_{11} & 0 & 0 & h_{14} & 0 \\ h_{21} & 0 & 0 & h_{24} & 0 \end{bmatrix}$ , and we can then use the extra degrees of freedom  $\mathbf{D}$  to make the off diagonal elements in  $\mathbf{H}$  zero.

To prove this, let  $\mathbf{H}_{nc}$  be the optimal solution for the best  $n_c$  measurements combination matrix, for example,  $\mathbf{H}_{nc} = \begin{bmatrix} h_{11} & h_{14} \\ h_{21} & h_{24} \end{bmatrix}$ .

The objective function is unaffected by  $\mathbf{D}$ , so as in the proof of Theorem 2 we choose  $\mathbf{D} = \mathbf{H}_{nc}^{-1}$ , to arrive at a diagonal  $\mathbf{H}$  as in (30).

**Case 2.4 Restriction on measurements from different process sections.** For example, consider a process with  $n_s$  sections with  $n_{y_k}$  measurements in section  $k$  (i.e. the total number of available measurements is  $n_y = \sum_{k=1}^{n_s} n_{y_k}$ ). If we want to select  $r_k$  measurements from each section  $k$ , the constraints (27b) become

$$\sum_{j=1}^{n_{y_k}} \sigma_{k-1} = r_k, \quad \forall k \in 1, 2, \dots, n_s \quad (31)$$

and Theorem 2 applies for the MIQP formulation.

**Case 2.5 Adding extra measurements to a given set of measurements.** This case may be very important in practice. For example, consider a process with  $n_y = 5$  measurements, where we have decided to use the measurements  $\{2, 3\}$ , and in addition want 2 other measurements (total 4 measurements). These constraints can be written

$$\sigma_j = 1, \quad \forall j = 2, 3$$

$$\sum_{j=1}^{n_y} \sigma_j = 4 \quad (32)$$

which can be written in the form (27b) with

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$$

and Theorem 2 applies at each MIQP node.

All of the above five cases belong to the optimal measurement selection (Problem 2) and can be easily solved using MIQP formulations. This is discussed in more detail for the examples below. Note that the Cases 2.4 and 2.5 cannot be dealt by BAB methods [17], at least not without changing the algorithms.

## 5. Examples (Problem 2)

### 5.1. Example 1: measurement selection for toy problem (Case 2.2)

To illustrate the problem formulation for (27) for Case 2.2, consider a “toy problem” from [18] which has two inputs  $\mathbf{u} = [u_1 \ u_2]^T$ , one disturbance  $d$  and two measured outputs  $\mathbf{z} = [z_1 \ z_2]^T$ . The cost function is

$$J = (z_1 - z_2)^2 + (z_1 - d)^2$$

where the outputs depend linearly on  $\mathbf{u}$ ,  $d$  as

$$\mathbf{z} = \mathbf{G}^z \mathbf{u} + \mathbf{G}_d^z d$$

with  $\mathbf{G}^z = \begin{bmatrix} 11 & 10 \\ 10 & 9 \end{bmatrix}$ ;  $\mathbf{G}_d^z = \begin{bmatrix} 10 \\ 9 \end{bmatrix}$ . The disturbances are of magnitude 1 and the measurements noise is at magnitude 0.01.

At the optimal point we have  $z_1 = z_2 = d$  and  $J_{\text{opt}}(d) = 0$ . Both the inputs and outputs are included in the candidate set of measurements

$$\mathbf{y} = \begin{bmatrix} z_1 \\ z_2 \\ u_1 \\ u_2 \end{bmatrix}$$

and we have  $n_y = 4$ ,  $n_u = 2$ . This gives

$$\mathbf{G}^y = \begin{bmatrix} 11 & 10 \\ 10 & 9 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{G}_d^y = \begin{bmatrix} 10 \\ 9 \\ 0 \\ 0 \end{bmatrix}$$

Furthermore,

$$\mathbf{J}_{uu} = \begin{bmatrix} 244 & 222 \\ 222 & 202 \end{bmatrix}, \quad \mathbf{J}_{ud} = \begin{bmatrix} 198 \\ 180 \end{bmatrix}$$

$$\mathbf{W}_d = 1, \quad \mathbf{W}_{n^y} = 0.01 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and  $\mathbf{J}_{uu}^{1/2} = \begin{bmatrix} 11.59 & 10.46 \\ 10.46 & 9.62 \end{bmatrix}$ . The resulting sensitivity matrix is

$$\mathbf{Y} = [\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}] = \begin{bmatrix} -1 & 0.01 & 0 & 0 & 0 \\ -1 & 0 & 0.01 & 0 & 0 \\ 9 & 0 & 0 & 0.01 & 0 \\ -9 & 0 & 0 & 0 & 0.01 \end{bmatrix}$$

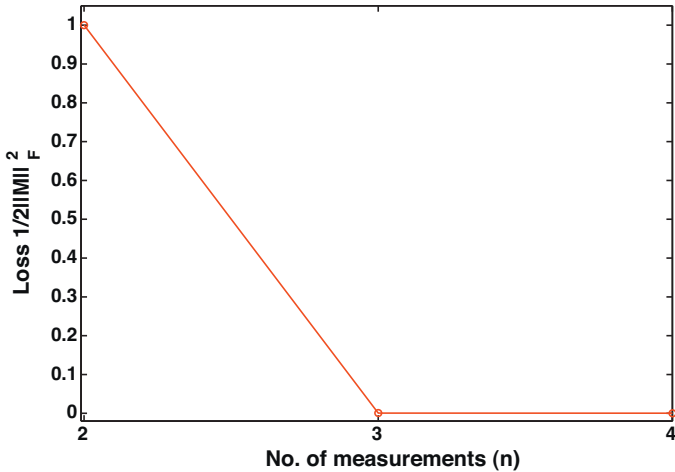


Fig. 4. The loss vs the number of included measurements ( $n$ ) for “toy problem”.

After vectorization (see Appendix A) we generate the matrices in (26). The resulting matrices to be used in MIQP problem (27) are

$$\mathbf{F}_\delta = \begin{bmatrix} 2 & 2 & -18 & 18 & 0 & 0 & 0 & 0 \\ 2 & 2 & -18 & 18 & 0 & 0 & 0 & 0 \\ -18 & -18 & 162 & -162 & 0 & 0 & 0 & 0 \\ 18 & 18 & -162 & 162 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 & -18 & 18 \\ 0 & 0 & 0 & 0 & 2 & 2 & -18 & 18 \\ 0 & 0 & 0 & 0 & -18 & -18 & 162 & -162 \\ 0 & 0 & 0 & 0 & 18 & 18 & -162 & 162 \end{bmatrix} \in \mathbb{R}^{8 \times 8}$$

$$\mathbf{G}_\delta^y = \begin{bmatrix} 11 & 10 & 1 & 0 & 0 & 0 & 0 & 0 \\ 10 & 9 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11 & 10 & 1 & 0 \\ 0 & 0 & 0 & 0 & 10 & 9 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 8}, \quad \mathbf{j}_\delta = \begin{bmatrix} 11.59 \\ 10.46 \\ 10.46 \\ 9.62 \end{bmatrix} \in \mathbb{R}^{4 \times 1}$$

To obtain the optimal  $n < 4$  measurement subset the constraint (27b) is

$$\sum_{j=1}^{n_y} \sigma_j = n$$

We used  $m=120$  for the big-m in (27) and with  $n=3$  we find by solving MIQP problem that the optimal solution is  $\mathbf{H} = \begin{bmatrix} 1.02 & 0 & 0.40 & 0.28 \\ 0.76 & 0 & 2.06 & 1.98 \end{bmatrix}$ , that is measurement 2 is not used. We can always choose the degrees of freedom in the matrix  $\mathbf{D}$ , for example, to have identity in measurements 1 and 3 to get, for example,  $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & -0.11 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ . The minimized loss (20) as a function of the number of measurements  $n$  is shown in Fig. 4. As expected, the loss is reduced as we use more measurements, but the reduction in loss is very small when we increase the number of measurements from 3 to 4. Based on Fig. 4, we conclude that using CVs as a combination of a 3 measurement subset is the best for this toy problem.

### 5.2. Example 2: measurement selection for evaporator process (Case 2.2)

The main purpose of this example is to evaluate the MIQP method (27) for Case 2.2 on a simple but realistic process. We consider the evaporator example of [27] (Fig. 5) as modified by

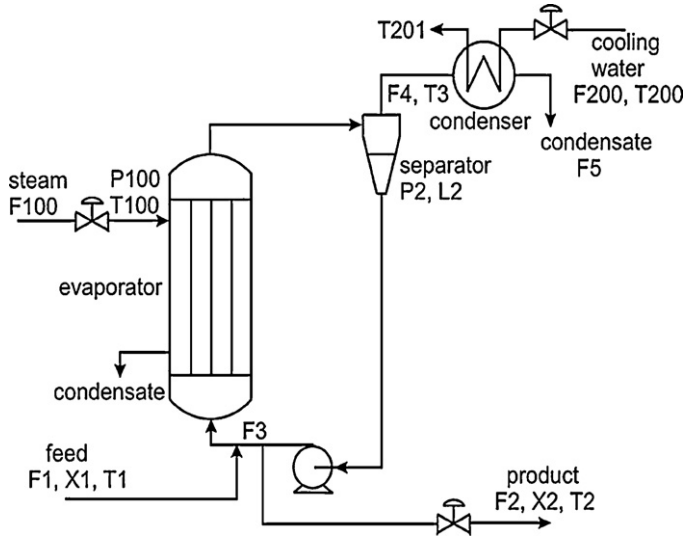


Fig. 5. Evaporator process.

[24]. The process has 2 steady-state degrees of freedom (inputs), 10 candidate measurements and 3 disturbances.

$$\mathbf{u} = [F_{200} \quad F_1]^T$$

$$\mathbf{y} = [P_2 \quad T_2 \quad T_3 \quad F_2 \quad F_{100} \quad T_{201} \quad F_3 \quad F_5 \quad F_{200} \quad F_1]^T$$

$$\mathbf{d} = [X_1 \quad T_1 \quad T_{200}]^T$$

Note that we as usual have included the inputs in the candidate measurements. The economic objective is to maximize the operating profit [\$/h], formulated as minimization of the negative profit [24].

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) + 0.2F_1 - 4800F_2 \quad (33)$$

The objective in self-optimizing control is to find optimal CVs that minimize the loss,  $L = J - J_{\text{opt}}(\mathbf{d})$ , in presence of disturbances and implementation errors. We formulated the problem (27) for this evaporator example and solved the MIQP to find the optimal CVs as the combinations of the best measurement subset size from 2 to 10. The YALMIP toolbox [28] is used to solve the MIQP problem with  $m=200$  in the big-m constraints in (27). To compare, the same problem was also solved by the downwards branch and bound (Downwards BAB) method and the partial bidirectional branch bound (PB<sup>3</sup>) method [22]. The three methods gave the same results and the loss as a function of the number of measurements ( $n$ ) used is shown in Fig. 6. The corresponding optimal measurements sets for the 9 subsets are given in Table 1. We note that  $F_{200}$  is included in all cases. From Fig. 6, we see that the loss decreases rapidly when the number of measurements is increased from 2 to 3, but from 3 measurements and on the loss decrease is smaller. Based on Fig. 6, Table 1 and acceptable loss CVs can be found as combinations of optimal measurement subsets for this 10-measurement evaporator example.

The average computational times (CPU time) using a Windows XP SP2 notebook with Intel® Core™ Duo Processor T7250 (2.00 GHz, 2M Cache, 800 MHz FSB) using MATLAB® R2009a for the MIQP, Downwards BAB, PB<sup>3</sup> methods and in addition the exhaustive search method are also tabulated in Table 1. Note that the exhaustive search was not actually performed and the given CPU time is an estimate based on assuming 0.001 s for each evaluation.

From Table 1, it can be seen that the MIQP method finds the optimal solution about one order of magnitude faster than the exhaustive search method, whereas the PB<sup>3</sup> and Downwards BAB methods are even one order of magnitude faster than MIQP. In



**Table 1**

Evaporator example: optimal measurement sets as a function of the number of measurements with associated losses and computational times.

No. meas, $n$	Optimal measurements	Loss <sup>a</sup> , $\frac{1}{2} \ \mathbf{M}\ _F^2$	CPU time (s)			
			MIQP	Downwards BAB	PB <sup>3</sup>	Exhaustive
2	[ $F_3$ $F_{200}$ ]	56.0260	0.0235	0.0028	0.0023	0.045
3	[ $F_2$ $F_{100}$ $F_{200}$ ]	11.7014	0.0350	0.0013	0.0028	0.12
4	[ $F_2$ $T_{201}$ $F_3$ $F_{200}$ ]	9.4807	0.0400	0.0016	0.0025	0.21
5	[ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_{200}$ ]	8.0960	0.0219	0.0011	0.0014	0.252
6	[ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_5$ $F_{200}$ ]	7.7127	0.0204	0.0016	0.0017	0.21
7	[ $P_2$ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_5$ $F_{200}$ ]	7.5971	0.0289	0.0009	0.0016	0.12
8	[ $P_2$ $T_2$ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_5$ $F_{200}$ ]	7.5756	0.0147	0.0005	0.0009	0.045
9	[ $P_2$ $T_2$ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_5$ $F_{200}$ $F_1$ ]	7.5617	0.0110	0.0008	0.0009	0.01
10	[ $P_2$ $T_2$ $T_3$ $F_2$ $F_{100}$ $T_{201}$ $F_3$ $F_5$ $F_{200}$ $F_1$ ]	7.5499	0.0008	0.0011	0.0009	0.001

<sup>a</sup> The results are the same as in [24], but the loss given in [24] is a factor  $3(n+n_d)$  smaller, see Section 6.2.

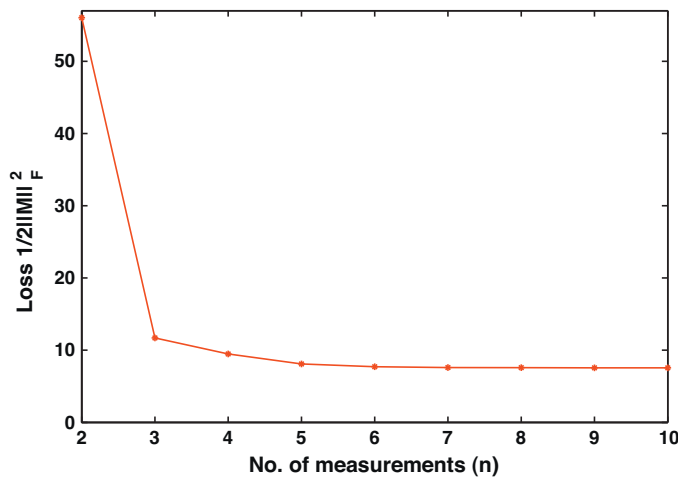


Fig. 6. Evaporator: loss vs the number of included measurements ( $n$ ).

conclusion, even though the MIQP method is not as fast as that of Downwards BAB and PB<sup>3</sup> methods; it is still acceptable as the optimal CVs selection is performed off-line. The advantage of MIQP method is that the method is simple, intuitive and can easily incorporate structural constraints which cannot be included with the BAB methods. This is considered in the next example.

5.3. Example 3: evaporator process with structural constraints (Case 2.4)

This example considers optimal measurement selection using MIQP formulations with the additional restrictions (31). As above, there are 3 temperature measurements, 6 flow measurements and 1 pressure measurement. The task is to use only 5 out of 10 measurements, more specifically, we want to use 1 pressure (among 1), 2 temperatures (among 3) and 2 flows (among 6). These constraints can easily be incorporated in the MIQP formulations, whereas these cannot be incorporated directly in the Downwards BAB and PB<sup>3</sup> methods. For the constraint (27b) we have

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

The optimal loss with these structural constraints is 12.9096 and the optimal measurement set is [ $F_2$   $F_{100}$   $T_{201}$   $T_2$   $P_2$ ]. To compare the loss with five measurements without any structural requirements is 8.0960 and the optimal measurements are [ $F_2$   $F_{100}$   $F_3$   $F_{200}$   $T_{201}$ ].

5.4. Example 4: measurement selection for distillation column (Case 2.2)

This example is included to apply the MIQP (27) formulations on a case with a large number of measurements and to highlight the computational effectiveness of the developed methods over the exhaustive search methods. We also include the computational effectiveness of both big-m approach (27c) and indicator constraint approach (28) for MIQP (27). We consider indirect composition control of a binary distillation column with 41 stages [29,30] and reflux ( $L$ ) and boil-up ( $V$ ) as the remaining unconstrained steady state degrees of freedom ( $\mathbf{u}$ ). The considered disturbances are in feed flow rate ( $F$ ), feed composition ( $z_F$ ) and liquid fraction ( $q_F$ ), which can vary between  $1 \pm 0.2$ ,  $0.5 \pm 0.1$  and  $1 \pm 0.1$ , respectively. As online composition measurements are assumed unavailable, we use stage temperatures inside the column to control the compositions indirectly. The boiling points difference between light key component ( $L$ ) and heavy key component ( $H$ ) is  $10^\circ\text{C}$ . We assume constant relative volatility of the components, constant pressure, no vapor hold up, equilibrium on each stage and constant molar flow rate. Under these assumptions only mass and component balances are included in this binary distillation column model and temperatures are approximated as linear functions of mole fractions. The temperature  $T_i$  ( $^\circ\text{C}$ ) on stage  $i$  is calculated as a simple linear function of the liquid composition  $x_i$  on each stage [29] (Fig. 7).

$$T_i = 0x_i + 10(1 - x_i) \tag{34}$$

The candidate measurements are the 41 stage temperatures which are measured with an accuracy of  $\pm 0.5^\circ\text{C}$ . Note that we do not include the inputs (flows  $L$  and  $V$ ) in the candidate measurements for this example because we would like to use only temperature combinations for control. The cost function  $J$  for the indirect composition control problem is the relative steady-state composition deviation,

$$J = \left( \frac{x_{top}^H - x_{top,s}^H}{x_{top,s}^H} \right)^2 + \left( \frac{x_{btm}^L - x_{btm,s}^L}{x_{btm,s}^L} \right)^2 \tag{35}$$

where  $x_{top}^H$  and  $x_{btm}^L$  denote the heavy key component ( $H$ ) composition in top product and light key component ( $L$ ) composition in bottom product and  $x_{top}^H = x_{btm}^L = 0.01$  (99% purity). The specification or set point value is denoted with subscript 's' [30]. This cost can be written in the general form in (3).

The MIQP formulation described in Case 2.2 in Section 4 is used to find 2 CVs as the optimal subset combinations of 2 to 41 stage temperatures. An MIQP is set up for this distillation column with the choice  $m = 2$  for the big-m constraints in Eq. (27). To obviate the need to select an appropriate  $m$ , another MIQP is set up by replacing big-m constraints (27c) with indicator constraint approach (28). The constraint in (27b) becomes  $\sum_{j=1}^{n_y} \sigma_j = n$ , where  $n$  varies from

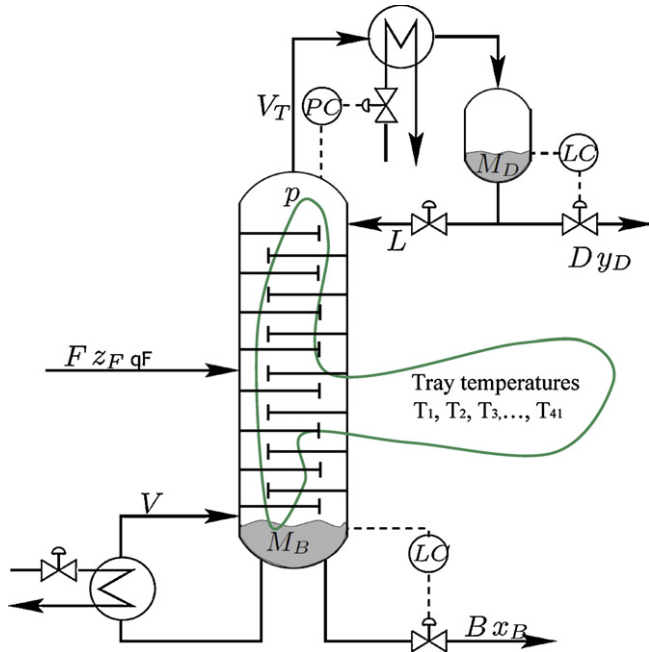


Fig. 7. Distillation column using LV-configuration.

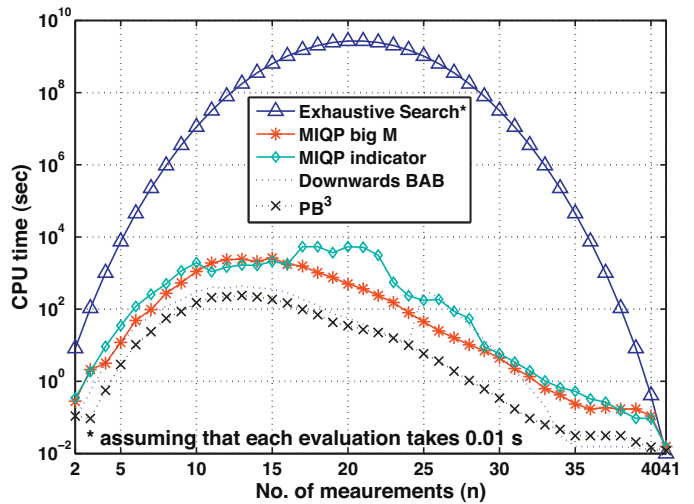


Fig. 9. Distillation column: CPU time requirement for computations in Fig. 8.

the  $\mathbf{D}$  matrix to make, for example,  $\mathbf{H} = \mathbf{I}$  (identity). For the case with 3 and 4 measurements, we choose to use the degrees of freedom in  $\mathbf{D}$  to make selected elements in  $\mathbf{H}$  equal to 1.

The same problem was also solved by the downwards branch and bound and partial bidirectional branch bound methods [22]. The computational times (CPU time) taken by MIQP with big-m approach, MIQP with indicator constraint approach, Downward BAB and  $\text{PB}^3$  methods and also the exhaustive search method are compared in Fig. 9. Note that exhaustive search is not performed and instead we give an estimate assuming 0.01 s for each evaluation. From Fig. 9, it can be seen that the MIQP finds the optimal solution 6 orders of magnitude faster than the exhaustive search methods. Contrary to theory, MIQP with indicator constraints take slightly higher computational times than MIQP with big-m approach, this could be due to the branching strategy used in CPLEX solver resulting in exploration of higher number of nodes. On an average, the MIQP with the big-m or indicator constraint approaches is about 1 order of magnitude slower than the  $\text{PB}^3$  and Downwards BAB methods. The MIQP method is relatively quick for measurement subset sizes between 25 and 41, but slower for subset sizes from 10 to 23. This is reasonable because subset sizes from 10 to 23 have a very high number of possibilities  $\left( \binom{41}{10} \text{ to } \binom{41}{23} \right)$ . In conclusion, even though the MIQP methods are not as computationally attractive as Downwards BAB and  $\text{PB}^3$  methods, the differences are not excessive.

5.5. Example 5: measurement selection for Kaibel column (Cases 2.4 and 2.5)

The Kaibel column example is included to show the optimal measurement selection using MIQP formulations with additional restrictions as given in (31) and (32). The 4-product Kaibel column shown in Fig. 10 has high energy saving potential [31], but presents a difficult control problem. The given 4-product Kaibel column arrangement separates a mixture of methanol (A), ethanol (B), propanol (C), butanol (D) into almost pure components. The economic objective function  $J$  is to minimize the impurities in the products.

$$J = D(1 - x_{A,D}) + S_1(1 - x_{B,S_1}) + S_2(1 - x_{C,S_2}) + B(1 - x_{D,B}) \quad (36)$$

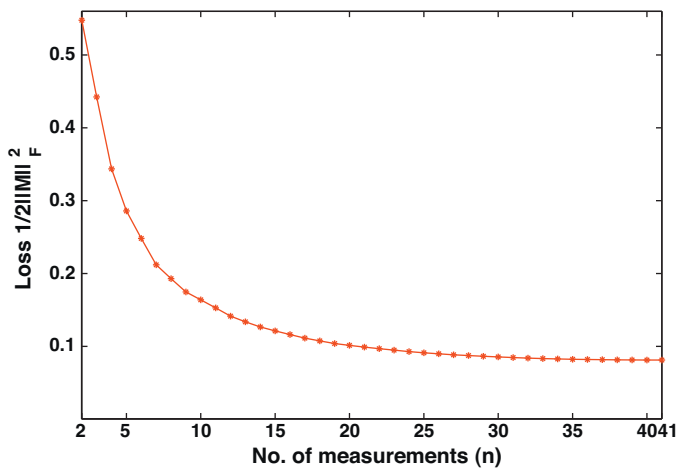


Fig. 8. Distillation column: loss vs the number of included measurements (n).

Table 2  
Distillation column example: optimal measurements and optimal controlled variables with loss.

No. meas, n	c's as combinations of measurements	Loss, $\frac{1}{2} \ \mathbf{M}\ _F^2$
2	$c_1 = T_{12}$ $c_2 = T_{30}$	0.5477
3	$c_1 = T_{12} + 0.0446T_{31}$ $c_2 = T_{30} + 1.0216T_{31}$	0.4425
4	$c_1 = 1.0316T_{11} + T_{12} + 0.0993T_{31}$ $c_2 = 0.0891T_{11} + T_{30} + 1.0263T_{31}$	0.3436
41	$c_1 = f(T_1, T_2, \dots, T_{41})$ $c_2 = f(T_1, T_2, \dots, T_{41})$	0.0813

2 to 41. The IBM ILOG Optimizer CPLEX solver is used to solve the MIQP problem. The minimized loss function with the number of measurements is shown in Fig. 8.

The optimal controlled variables (measurement combination matrix  $\mathbf{H}$ ) for the cases with 2, 3, 4 and 41 measurements are shown in Table 2. For the case with 2 measurements, we just give the measurement, and not the combination, because we can always choose

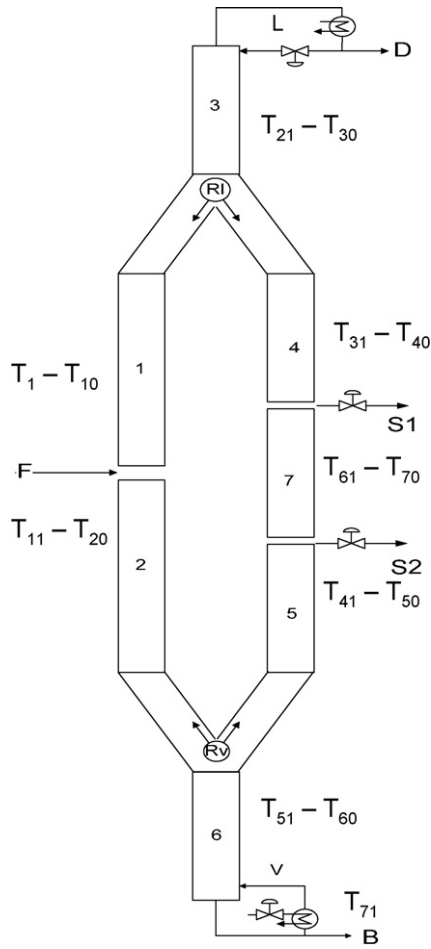


Fig. 10. The 4-product Kaibel column.

where  $D$ ,  $S_1$ ,  $S_2$  and  $B$  are the distillate, side product 1, side product 2 and bottom flow rates (mol/min) respectively.  $x_{ij}$  is mole fraction of component  $i$  in product  $j$ .

The Kaibel column has 4 inputs ( $L$ ,  $S_1$ ,  $S_2$ ,  $R_L$ ) and 71 temperature measurements (7 sections with each section having 10 tray temperatures plus 1 temperature for reboiler), which we included as the candidate measurements ( $\mathbf{y}$ ) and are measured with an accuracy of  $\pm 0.1^\circ\text{C}$ . The considered disturbances are in vapor boil up ( $V$ ), vapor split ( $R_V$ ), feed flow rate ( $F$ ), mole fraction of A in feed stream ( $z_A$ ), mole fraction of B in feed stream ( $z_B$ ), mole fraction of C in feed stream ( $z_C$ ), liquid fraction of the feed stream ( $q_F$ ), which vary between  $3 \pm 0.25$ ,  $0.4 \pm 0.1$ ,  $1 \pm 0.25$ ,  $0.25 \pm 0.05$ ,  $0.25 \pm 0.05$ ,  $0.25 \pm 0.05$ ,  $0.9 \pm 0.05$ , respectively. The reader is referred to [32] for further details on this example.

We consider the selection of the control variables as individual measurements or combinations of a measurement subset with measurements from specified sections of the column as structural constraints. Such structural constraints may be important for dynamic reasons, for example, at least one temperature in the prefractionator should be used in the regulatory layer [32]. The 4-product Kaibel column is divided into 4 segments with 20, 20, 10 and 21 measurements, respectively. The measurements in the four segments are  $T_1 - T_{20}$ ,  $T_{21} - T_{40}$ ,  $T_{61} - T_{70}$  and  $T_{41} - T_{60}$  plus  $T_{71}$ , respectively (Fig. 10). Note that segment 4 includes reboiler temperature  $T_{71}$ . The candidate measurements  $\mathbf{y}$  and given inputs  $\mathbf{u}$  are

$$\mathbf{y} = [T_1 \ T_2 \ T_3 \ \dots \ T_{71}]^T$$

$$\mathbf{u} = [L \ S_1 \ S_2 \ R_L]^T$$

Table 3

Kaibel column: optimal measurement sets and loss using optimal combination of these measurements.

Case	No. meas. $n$	Optimal measurements	Loss, $\frac{1}{2} \ \mathbf{M}\ _F^2$	CPU time (min)
(i)	4	$[T_{12} \ T_{40} \ T_{51} \ T_{66}]$	11.6589	34.23
(i)	5	$[T_{12} \ T_{51} \ T_{62} \ T_{65} \ T_{66}]$	2.9700	120
(i)	6	$[T_{12} \ T_{20} \ T_{23} \ T_{57} \ T_{60} \ T_{64}]$	1.0140	120
(i)	71	$[T_1 \ T_2 \ \dots \ T_{71}]$	0.0101	0.0007
(ii)	4 <sup>a</sup>	$[T_{12} \ T_{40} \ T_{51} \ T_{66}]$	11.6589	1.19
(iii)	4 <sup>c</sup>	$[T_{12} \ T_{25} \ T_{45} \ T_{62}]$	1328.6691	0.0005
(iii)	5 <sup>b</sup>	$[T_{12} \ T_{25} \ T_{45} \ T_{62} \ T_{69}]$	65.7180	0.096
(iii)	6 <sup>b</sup>	$[T_{12} \ T_{25} \ T_{45} \ T_{55} \ T_{62} \ T_{71}]$	3.5646	0.19
(iii)	7 <sup>b</sup>	$[T_{12} \ T_{25} \ T_{45} \ T_{51} \ T_{62} \ T_{65} \ T_{67}]$	0.9450	2.21

<sup>a</sup> (ii) Case 2.4.

<sup>b</sup> (iii) Case 2.5.

<sup>c</sup> Given non-optimal measurement set.

We formulate an MIQP using (27) to find four CVs for the following three cases:

- (i) Optimal combinations of 4, 5, 6 and 71 measurements with no constraint on sections (Case 2.2).
- (ii) Single measurements from each of the four segments (Case 2.4).
- (iii) Including extra measurements to a given set of measurements (Case 2.5). In this case,  $\{T_{12}, T_{25}, T_{45}, T_{62}\}$  are taken as the given set of measurements, which could have been selected based on considerations for stabilizing the column profiles.

The constraint for (i) is

$$\sum_{j=1}^{n_y} \sigma_j = n$$

for  $n = 4, 5, 6$  and 71. This can alternatively be written in the general form in (27b) with

$$\mathbf{P} = \mathbf{1}_{1 \times n_y}, \mathbf{s} = n$$

where  $\mathbf{1}$  is a column vector of ones and  $n$  is 4, 5, 6 and 71.

The constraints for (ii) can be written in the general form (27b) with

$$\mathbf{P} = \begin{bmatrix} \mathbf{1}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 10}^T & 0 \\ \mathbf{0}_{1 \times 20}^T & \mathbf{1}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 10}^T & 0 \\ \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{1}_{1 \times 20}^T & \mathbf{0}_{1 \times 10}^T & 1 \\ \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{0}_{1 \times 20}^T & \mathbf{1}_{1 \times 10}^T & 0 \end{bmatrix}, \mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

where  $\mathbf{1}$  is a column vector of ones and  $\mathbf{0}$  is a column vector of zeros.

(iii) We consider including 1, 2 and 3 extra measurements to the given set  $\{T_{12}, T_{25}, T_{45}, T_{62}\}$ . The constraints for this case are

$$\sigma_j = 1, \quad \forall j = 12, 25, 45, 62$$

$$\sum_{j=1}^{71} \sigma_j = n$$

where  $n = 5, 6$  or 7. The optimal measurements sets for Cases (i), (ii), (iii) together with the loss and computational times are reported in Table 3. Note that for Case (i) with 5, 6 measurements, the reported solutions are not optimal solutions as the computational time required for these cases exceeded the set maximum computational time limit of 120 min. The measurements sets for  $n = 4$  are the same for (i) and (ii) because it happens that the optimal measurements in Case (i) have the desired distribution. However, the computational time is about 30 times higher for Case (i) as the number of possibilities is higher in (i) than in (ii). For Case (iii), the

loss decreases as we add 1, 2, and 3 extra measurements to the given set.

**6. Discussion**

6.1. Structured  $\mathbf{H}$  with specified zero elements (Problem 3)

Unfortunately, the convex formulation in Theorem 2 used in the above Examples, does not generally apply when specified elements in  $\mathbf{H}$  are zero. Some examples are

- (I) **Decentralized structure.** This is the case, where we want to combine measurements from a individual unit/section alone in a plant, so the measurement sets are disjoint. This can be viewed as selecting CVs for individual units/sections in the plant. As an example, consider a process with 2 inputs (degrees of freedom) and 5 measurements with 2 disjoint measurement sets  $\{1, 2, 3\}, \{4, 5\}$ ; the structure is

$$\mathbf{H}_I = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 & 0 \\ 0 & 0 & 0 & h_{24} & h_{25} \end{bmatrix}$$

- (II) **Triangular structure.** More generally,  $\mathbf{H}$  may have a triangular structure. As an example, consider a process with 2 degrees of freedom and 5 measurements with partially disjoint measurement sets as  $\{1, 2, 3, 4, 5\}$  for one CV and  $\{4, 5\}$  for another CV, the structure is

$$\mathbf{H}_{II} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} \\ 0 & 0 & 0 & h_{34} & h_{35} \end{bmatrix}$$

Since Theorem 2 does **not** hold for these cases with specified structures, we need to solve non-convex problems. This is outside the scope of this paper, where convex formulations are considered.

6.2. Use of average loss  $\frac{1}{2} \|\mathbf{M}\|_F^2$

For the measurement selection problem, using an uniform distribution for  $\mathbf{d}'$  and  $\mathbf{n}^{y'}$  with  $\left\| \begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^{y'} \end{bmatrix} \right\|_2 \leq 1$  results in the average loss  $\hat{L}_{avg} = (1/(6(n_y + n_d))) \|\mathbf{M}\|_F^2$  [24]. Although this loss expression is mathematically correct, the use of a uniform distribution is not meaningful from an engineering point of view. Specifically, the reduction in the loss by the factor  $(n_y + n_d)$  is not meaningful. To illustrate this, note that we can add dummy measurements and thus set  $n_y$  to any number, and then choose to not use these dummy measurements when selecting  $\mathbf{c} = \mathbf{H}\mathbf{y}$ , simply by setting the corresponding columns in  $\mathbf{H}$  to zero. As the Frobenius norm of a matrix is the same if we add columns of zeros,  $\|\mathbf{M}\|_F$  will be unchanged, but  $n_y$  increases and the loss  $\hat{L}_{avg}$  decreases. Since the loss should not change by adding dummy measurements that we do not use, the use of uniform distribution of the two-norm is not physically meaningful. Hence, in this paper, we choose to use the more common normal distribution for  $\mathbf{d}'$  and  $\mathbf{n}^{y'}$  which gives the average loss (expected loss)  $L_{avg} = (1/2) \|\mathbf{M}\|_F^2$  in (20).

**7. Conclusions**

The problem of finding optimal CV measurement combinations that minimize the loss from optimal operation is solved. The optimal CV selection problem from self optimizing control framework is reformulated as a QP and the optimal CV selection for measurement subsets is formulated as an MIQP problem. The developed MIQP based method allows for additional structural constraints compared to the bidirectional branch and bound methods reported in

literature. The MIQP based method was found to use about 10 times more CPU time than the bidirectional branch and bound methods, but this is acceptable as the optimal CV selection problem is done offline. In addition, the MIQP method can be used on some problems where the branch and bound methods do not apply, as shown for the Kaibel column example.

**Appendix A.**

The vectorization procedure of convex optimization problem in decision matrix  $\mathbf{H}$  to convex optimization problem in  $\mathbf{h}_\delta$  is described [19]. We write

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n_y} \\ h_{21} & h_{22} & \dots & h_{2n_y} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_u1} & h_{n_u2} & \dots & h_{n_u n_y} \end{bmatrix} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \dots \quad \mathbf{h}_{n_y}] = \begin{bmatrix} \tilde{\mathbf{h}}_1^T \\ \tilde{\mathbf{h}}_2^T \\ \vdots \\ \tilde{\mathbf{h}}_{n_u}^T \end{bmatrix}$$

where

$$\mathbf{h}_j = j\text{th column of } \mathbf{H}, \quad \mathbf{h}_j \in \mathbb{R}^{n_u \times 1}$$

$$\tilde{\mathbf{h}}_j = j\text{th row of } \mathbf{H}, \quad \tilde{\mathbf{h}}_j \in \mathbb{R}^{n_y \times 1}$$

The transpose must be included because all vectors including  $\tilde{\mathbf{h}}_i$  are column vectors.

Similarly, let  $\mathbf{J}_{uu}^{1/2} = [\mathbf{j}_1 \quad \mathbf{j}_2 \quad \dots \quad \mathbf{j}_{n_u}]$ .

We further introduce the long vectors  $\mathbf{h}_\delta$  and  $\mathbf{j}_\delta$ ,

$$\mathbf{h}_\delta = \begin{bmatrix} \tilde{\mathbf{h}}_1 \\ \tilde{\mathbf{h}}_2 \\ \vdots \\ \tilde{\mathbf{h}}_{n_u} \end{bmatrix} = \begin{bmatrix} h_{11} \\ h_{12} \\ \vdots \\ h_{1n_y} \\ h_{21} \\ h_{22} \\ \vdots \\ h_{2n_y} \\ h_{n_u1} \\ h_{n_u2} \\ \vdots \\ h_{n_u n_y} \end{bmatrix} \in \mathbb{R}^{n_u n_y \times 1}$$

$\mathbf{j}_\delta^T = [\mathbf{j}_1^T \quad \mathbf{j}_2^T \quad \dots \quad \mathbf{j}_{n_u}^T] \in \mathbb{R}^{n_u n_u \times 1}$  and the large matrices

$$\mathbf{G}_\delta^T = \begin{bmatrix} \mathbf{G}^{y^T} & 0 & 0 & \dots \\ 0 & \mathbf{G}^{y^T} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & \dots & \mathbf{G}^{y^T} \end{bmatrix}, \quad \mathbf{Y}_\delta = \begin{bmatrix} \mathbf{Y} & 0 & 0 & \dots \\ 0 & \mathbf{Y} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & \dots & \mathbf{Y} \end{bmatrix}$$

Then,  $\mathbf{H}\mathbf{Y} = \begin{bmatrix} \tilde{\mathbf{h}}_1^T \mathbf{Y} \\ \tilde{\mathbf{h}}_2^T \mathbf{Y} \\ \vdots \\ \tilde{\mathbf{h}}_{n_u}^T \mathbf{Y} \end{bmatrix}$  and for the Frobenius norm the following equalities apply.

$$\|\mathbf{H}\mathbf{Y}\|_F^2 = \left\| \begin{bmatrix} \tilde{\mathbf{h}}_1^T \mathbf{Y} \\ \tilde{\mathbf{h}}_2^T \mathbf{Y} \\ \vdots \\ \tilde{\mathbf{h}}_{n_u}^T \mathbf{Y} \end{bmatrix} \right\|_F^2 = \|\tilde{\mathbf{h}}_1^T \mathbf{Y} \quad \tilde{\mathbf{h}}_2^T \mathbf{Y} \quad \dots \quad \tilde{\mathbf{h}}_{n_u}^T \mathbf{Y}\|_F^2$$

$$= \|\mathbf{h}_\delta^T \mathbf{Y}_\delta\|_F = \|\mathbf{h}_\delta \mathbf{Y}_\delta^T\|_F = \mathbf{h}_\delta^T \underbrace{\mathbf{Y}_\delta \mathbf{Y}_\delta^T}_{\mathbf{F}_\delta} \mathbf{h}_\delta = \mathbf{h}_\delta^T \mathbf{F}_\delta \mathbf{h}_\delta$$

Because  $\mathbf{HG}^y = \mathbf{J}_{uu}^{1/2}$  where  $\mathbf{J}_{uu}^{1/2}$  is symmetric matrix, we have  $\mathbf{HG}^y = \mathbf{G}^{yT} \mathbf{H}^T = \mathbf{J}_{uu}^{1/2}$  and

$$[\mathbf{G}^{yT} \tilde{\mathbf{h}}_1 \quad \mathbf{G}^{yT} \tilde{\mathbf{h}}_2 \quad \dots \quad \mathbf{G}^{yT} \tilde{\mathbf{h}}_{n_u}] = [\mathbf{j}_1 \quad \mathbf{j}_2 \quad \mathbf{j}_{n_u}] \Rightarrow \mathbf{G}_\delta^T \mathbf{h}_\delta = \mathbf{j}_\delta$$

## References

- [1] A. Foss, Critique of chemical process control theory, *AIChE Journal* 19 (1973) 209–214.
- [2] S. Skogestad, Plantwide control: the search for the self-optimizing control structure, *Journal of Process Control* 10 (2000) 487–507.
- [3] M. van de Wal, A. de Jager, A review of methods for input/output selection, *Automatica* 37 (2001) 487–510.
- [4] L. Narraway, J. Perkins, G. Barton, Interaction between process design and process control: economic analysis of process dynamics, *Journal of Process Control* 1 (1991) 243–250.
- [5] L.T. Narraway, J.D. Perkins, Selection of process control structure based on linear dynamic economics, *Industrial & Engineering Chemistry Research* 32 (1993) 2681–2692.
- [6] 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 (1980) 220–232.
- [7] A.C. de Araújo, M. Govatsmark, S. Skogestad, Application of plantwide control to the HDA process. I. Steady-state optimization and self-optimizing control, *Control Engineering Practice* 15 (2007) 1222–1237.
- [8] S. Vasudevan, G.P. Rangaiah, N.V.S.N.M. Konda, W.H. Tay, Application and evaluation of three methodologies for plantwide control of the styrene monomer plant, *Industrial & Engineering Chemistry Research* 48 (2009) 10941–10961.
- [9] M. Panahi, S. Skogestad, Economically efficient operation of CO<sub>2</sub> capturing process. Part I. Self-optimizing procedure for selecting the best controlled variables, *Chemical Engineering and Processing: Process Intensification* 50 (2011) 247–253.
- [10] J.F. Forbes, T.E. Marlin, Design cost: a systematic approach to technology selection for model-based real-time optimization systems, in: *Fifth International Symposium on Process Systems Engineering*, Computers & Chemical Engineering 20 (1996) 717–734.
- [11] B. Srinivasan, D. Bonvin, E. Visser, S. Palanki, Dynamic optimization of batch processes. II. Role of measurements in handling uncertainty, *Computers & Chemical Engineering* 27 (2003) 27–44.
- [12] J.V. Kadam, W. Marquardt, B. Srinivasan, D. Bonvin, Optimal grade transition in industrial polymerization processes via NCO tracking, *AIChE Journal* 53 (2007) 627–639.
- [13] K.B. Ariyur, M. Krstic, *Real-Time Optimization by Extremum-Seeking Control*, Wiley-Interscience, 2003.
- [14] M. Guay, T. Zhang, Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties, *Automatica* 39 (2003) 1283–1293.
- [15] A. Kassidas, J. Patry, T. Marlin, Integrating process and controller models for the design of self-optimizing control, *Computers & Chemical Engineering* 24 (2000) 2589–2602.
- [16] S. Engell, Feedback control for optimal process operation, in: *Special Issue ADCHEM 2006 Symposium*, *Journal of Process Control* 17 (2007) 203–219.
- [17] V. Kariwala, Y. Cao, Bidirectional branch and bound for controlled variable selection. Part III. Local average loss minimization, *IEEE Transactions on Industrial Informatics* 6 (2010) 54–61.
- [18] I.J. Halvorsen, S. Skogestad, J.C. Morud, V. Alstad, Optimal selection of controlled variables, *Industrial & Engineering Chemistry Research* 42 (2003).
- [19] V. Alstad, S. Skogestad, E. Hori, Optimal measurement combinations as controlled variables, *Journal of Process Control* 19 (2009) 138–148.
- [20] V. Kariwala, Optimal measurement combination for local self-optimizing control, *Industrial & Engineering Chemistry Research* 46 (2007) 3629–3634.
- [21] S. Heldt, Dealing with structural constraints in self-optimizing control engineering, *Journal of Process Control* 20 (2010) 1049–1058.
- [22] V. Kariwala, Y. Cao, Bidirectional branch and bound for controlled variable selection. Part II. Exact local method for self-optimizing control, *Computers & Chemical Engineering* 33 (2009) 1402–1414.
- [23] S. Skogestad, I. Postlethwaite, *Multivariable Feedback Control*, 1st edition, Wiley, 1996.
- [24] V. Kariwala, Y. Cao, S. Janardhanan, Local self-optimizing control with average loss minimization, *Industrial & Engineering Chemistry Research* 47 (2008) 1150–1158.
- [25] V. Alstad, S. Skogestad, Null space method for selecting optimal measurement combinations as controlled variables, *Industrial & Engineering Chemistry Research* 46 (2007) 846–853.
- [26] J.N. Hooker, M.A. Osorio, Mixed logical-linear programming, *Discrete Applied Mathematics* 96–97 (1999) 395–442.
- [27] R.B. Newell, P. Lee, *Applied Process Control: A Case Study*, Prentice-Hall of Australia, New York/Sydney, 1989.
- [28] J. Lofberg, Yalmip: a toolbox for modeling and optimization in matlab, in: *2004 IEEE International Symposium on Computer Aided Control Systems Design*, 2004, pp. 284–289.
- [29] S. Skogestad, Dynamics and control of distillation columns: a tutorial introduction, *Chemical Engineering Research and Design* 75 (1997) 539–562, Distillation.
- [30] E.S. Hori, S. Skogestad, Selection of controlled variables: maximum gain rule and combination of measurements, *Industrial & Engineering Chemistry Research* 47 (2008) 9465–9471.
- [31] I.J. Halvorsen, S. Skogestad, Minimum energy consumption in multicomponent distillation. 3. More than three products and generalized petlyuk arrangements, *Industrial & Engineering Chemistry Research* 42 (2003) 616–629.
- [32] J. Strandberg, S. Skogestad, Stabilizing operation of a 4-product integrated kaibel column, *Institution of Chemical Engineers Symposium Series* 152 (2006) 636–647.