Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Optimal Operation by Controlling the Gradient to Zero

Johannes Jäschke Sigurd Skogestad

Department of Chemical Engineering, Norwegian University of Science and Technology, NTNU, Trondheim, Norway (e-mail: {jaschke}{skoge}@chemeng.ntnu.no)

Abstract: From an optimization point of view, the gradient is the key variable which gives information about the optimality of a process. In this paper we present how the gradient is related to the loss from optimality, and show how determining a good set of controlled variables can be considered as *weighted* approximation of the gradient. We show that even if there are setpoint changes for the controlled variables, this can still be considered as approximating the gradient.

Keywords: Self-optimizing control, Gradient control, Optimizing control, Control structure selection

1. INTRODUCTION

The overall objective of process operation is to minimize the cost J (or equivalently to maximize the profit $P = -J$) subject to given constraints. However, when using control, the objective is to keep selected controlled variables $\mathbf{c}$ at their optimal setpoints,

$$\mathbf{c}(\mathbf{y}) = \mathbf{c}_s. \tag{1}$$

With respect to these two goals, Morari et al. [1980] stated

> "...our main objective is to translate the economic objectives into process control objectives. In other words we want to find a function $\mathbf{c}$ of the process variables [...] which when held constant leads automatically to the optimal adjustment of the manipulated variables, and with it, the optimal operating conditions."

However, they do not give a systematic method for finding the controlled variables, nor do they mention that for the unconstrained case, the obvious approach to get consistency between economic and process control objectives is to select the gradient as the controlled variable. That is, to select

$$\mathbf{c} = J_{\mathbf{u}}(\mathbf{u}, \mathbf{d}), \tag{2}$$

and keep the setpoint constant at zero, $\mathbf{c}_s = 0$. Here $\mathbf{u}$ are the unconstrained degrees of freedom, $\mathbf{d}$ are unmeasured disturbances, and $J_{\mathbf{u}}(\mathbf{u}, \mathbf{d}) = \partial J(\mathbf{u}, \mathbf{d})/\partial \mathbf{u}$ is the gradient. Irrespective of the disturbance, the optimal value of $J_{\mathbf{u}}$ is zero, (Figure 1). This was proposed by Halvorsen and Skogestad [1997a], who write that the ideal controlled variable would be

$$\mathbf{c} = \mathbf{c}_1 J_{\mathbf{u}} + \mathbf{c}_0, \tag{3}$$

where $\mathbf{c}_0$ and $\mathbf{c}_1$ are constants. The idea has also been proposed by Halvorsen and Skogestad [1997b, 1999], Bonvin et al. [2001], Cao [2003, 2005], Srinivasan et al. [2008], and intuitively it seems to be an excellent idea. The elements

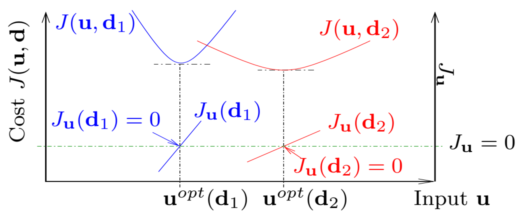

Fig. 1. Cost and gradients for different disturbances $\mathbf{d}$

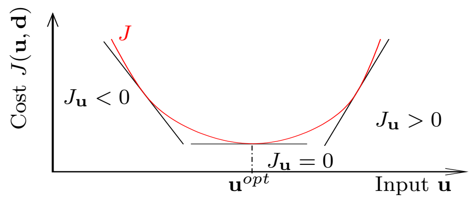

Fig. 2. Cost and gradient values

of the gradient change sign when moving from one side of the optimum to another side (Figure 2), thus, it is well suited for feedback control.

However, in practice, we rarely have a measurement of the gradient and it is often not clearly defined what it means to control the gradient to zero. The gradient is a vector, and in many practical cases is not possible to control all elements exactly to zero. What should we do in these cases? A first attempt to answer this question will be to find a control structure, which minimizes the norm of the gradient. This is a good start, however, it is important to keep in mind that our ultimate goal is to minimize the cost J, so this original criterion has to be applied to evaluate the possible control structures.

The starting point is to write the controlled variables as a function of measurements $\mathbf{y}$,

$$\mathbf{c} = \mathbf{H}\mathbf{y}, \tag{4}$$

which is controlled to zero, $\mathbf{c} = \mathbf{H}\mathbf{y} = 0$. Since the gradient is optimally at zero, we can consider $\mathbf{c} = \mathbf{H}\mathbf{y}$ as an approximation of the gradient. If the approximation is exact, $\mathbf{H}\mathbf{y} = J_{\mathbf{u}}$ then we will have optimal operation whenever $\mathbf{c} = 0$, provided convexity. If it is not possible to control the gradient (because of e.g. unmeasured disturbances, noise and missing measurements), there will be some loss associated to the chosen control structure. To evaluate the performance of the chosen control policy, we use the original cost function and define the loss from optimality:

$$L = J(\mathbf{u}, \mathbf{d}) - J(\mathbf{u}^{opt}, \mathbf{d}) \tag{5}$$

Considering the problem of selecting the best control structures, there are two important questions, which we would like to address in this paper:

Q1. Does a $\mathbf{H}$ which minimizes $||J_{\mathbf{u}}(\mathbf{H}\mathbf{y} = 0)||_2$ also minimize $L(\mathbf{H}\mathbf{y} = 0)$?
Q2. If not, is the difference significant?

In terms of **Q1.** we show in Theorem 1 that minimizing the norm of the gradient is not quite the same as minimizing the loss L.

In terms of **Q2.** we show that it is important in the case, when we have structural constraints on $\mathbf{H}$. That is, we have control structures involving different measurements,

Another contribution of this paper is an extremely simple derivation of the null space method [Alstad and Skogestad, 2007].

Furthermore, we show we show how setpoint changes of the controlled variables can be seen in the context of minimizing the loss or approximating the gradient.

This paper is structured such that the next section presents our main result, a derivation of the expression for the economic loss based on the gradient. In Section 3 we describe how this interpretation is connected to existing methods, and its importance. Section 4 discusses how the case of varying setpoints can be treated in this framework. After presenting a distillation case study in Section 5, we close the paper with a discussion and conclusions.

2. DERIVATION OF THE LOSS EXPRESSION USING THE GRADIENT

2.1 Preliminaries

Consider the feedback system in Figure 3, where the variables $\mathbf{c}$ and $\mathbf{c}_s$ denote the vector valued controlled variable and its setpoint, and where the variables $\mathbf{n}^{\mathbf{y}}, \mathbf{n}^c$ denote the noise and the steady state control error, respectively. The noisy measurements are $\mathbf{y}_m = \mathbf{y} + \mathbf{n}^{\mathbf{y}}$, and we assume that the controllers have integral action so that there is no steady state error, $\mathbf{n}^c = 0$; then at steady state $\mathbf{c}_s = \mathbf{c} + \mathbf{n}$, where $\mathbf{n} = \mathbf{H}\mathbf{n}^{\mathbf{y}}$.

After all active constraints are satisfied (controlled), the remaining unconstrained problem can be formulated as

$$\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d}) \tag{6}$$

which in the neighbourhood of the optimal point can be approximated by a quadratic problem

$$\min_{\mathbf{u}} \frac{1}{2} \left[\mathbf{u}^{\mathrm{T}} \ \mathbf{d}^{\mathrm{T}} \right] \begin{bmatrix} \mathbf{J_{uu}} & \mathbf{J_{ud}} \\ \mathbf{J_{du}} & \mathbf{J_{dd}} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} \tag{7}$$

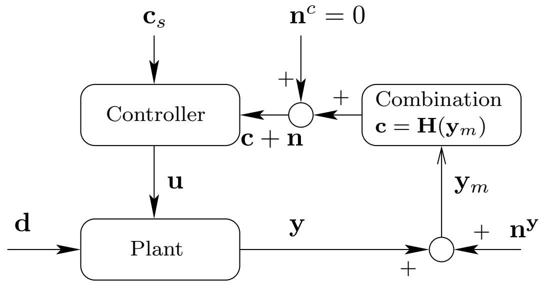

Fig. 3. Control structure (with integral action, $\mathbf{c}_s = \mathbf{c} + \mathbf{n}$ at steady state, $\mathbf{n}^c = 0$)

where we assume that $\mathbf{J_{uu}} > 0$. For small deviations around the nominal optimum, the plant can be described by the linear model

$$\mathbf{y} = \mathbf{G}^{\mathbf{y}}\mathbf{u} + \mathbf{G}^{\mathbf{y}}_{\mathbf{d}}\mathbf{d} = \tilde{\mathbf{G}}^{\mathbf{y}} \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix}, \tag{8}$$

where $\mathbf{G}^{\mathbf{y}}$ and $\mathbf{G}^{\mathbf{y}}_{\mathbf{d}}$ are the steady state gain matrices from $\mathbf{u}$ and $\mathbf{d}$ to the outputs $\mathbf{y}$. Our goal is to find controlled variables of the form

$$\mathbf{c} = \mathbf{H}\mathbf{y}_m, \tag{9}$$

which, when controlled to zero, yield optimal or near optimal operation.

2.2 Approximating the gradient

The first order accurate expression for gradient of (6) is

$$J_{\mathbf{u}}(\mathbf{u}, \mathbf{d}) = \left[\mathbf{J_{uu}} \ \mathbf{J_{ud}} \right] \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} = \mathbf{J_{uu}}\mathbf{u} + \mathbf{J_{ud}}\mathbf{d} \tag{10}$$

Assuming that (7) matches the real plant, the necessary condition for optimality is

$$J_{\mathbf{u}}(\mathbf{u}, \mathbf{d}) = \left[\mathbf{J_{uu}} \ \mathbf{J_{ud}} \right] \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} = 0. \tag{11}$$

As mentioned above, the ideal controlled variable is the gradient, $\mathbf{c} = J_{\mathbf{u}}(\mathbf{u}, \mathbf{d})$. When it is known exactly, using it as a controlled variable is the best choice and works fine. In practice, however, the gradient has to be somehow estimated using measurement information. Then the controlled variable becomes

$$\mathbf{c} = \hat{J}_{\mathbf{u}}. \tag{12}$$

Obtaining the gradient estimate $\hat{J}_{\mathbf{u}}$ can be done in several ways, such as e.g. black box modelling or estimating the gradient using statistical methods. In the case of zero-mean noise, the effects may cancel out, but if there is a constant non-zero offset, the noise can deteriorate performance severely, thus we have to include the noise in the analysis, too.

A first approach would be to find a controlled variable $\mathbf{c} = \mathbf{H}\mathbf{y}_m$ which minimizes the worst case gradient norm, e.g. to select $\mathbf{H}$ as

$$\mathbf{H} = \arg \left(\min_{\mathbf{H}} \max_{\mathbf{d}} ||J_{\mathbf{u}} - \mathbf{H}\mathbf{y}_m||_2 \right). \tag{13}$$

In the non-ideal case, when $\mathbf{H}\mathbf{y}_m \neq J_{\mathbf{u}}$, controlling $\mathbf{H}\mathbf{y}_m$ to zero will result in a gradient which has nonzero elements, and therefore has nonzero norm,

$$||J_{\mathbf{u}}(\mathbf{H}\mathbf{y}_m = 0)||_2 \neq 0. \tag{14}$$

The norm of the gradient may seem a good criterion to evaluate suboptimality, however it does not truly reflect

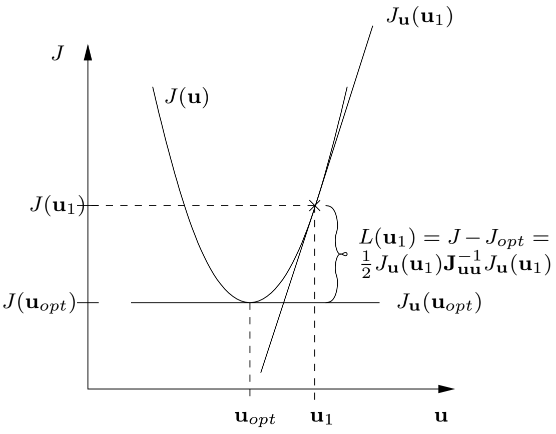

Fig. 4. Loss L imposed by non-optimal operation

the performance in terms of the original cost function. To quantify the suboptimality, we consider the loss L, which is defined as the difference between the actual cost and the optimal cost for a given disturbance $\mathbf{d}$,

$$L = J(\mathbf{u}, \mathbf{d}) - J(\mathbf{u}_{opt}, \mathbf{d}). \tag{15}$$

Note that we are considering the loss with respect to the truly optimal instead of the cost. The loss has the properties of a weighted norm.

Theorem 1. The local economic loss can be expressed to first order in terms of the current gradient value as

$$L = \frac{1}{2} \left|\left| \mathbf{J}_{\mathbf{uu}}^{-1/2} J_{\mathbf{u}} \right|\right|_2^2 \tag{16}$$

Proof From Halvorsen et al. [2003] it is known that the loss can be written as

$$L = \frac{1}{2}(\mathbf{u} - \mathbf{u}^{opt})^{\mathrm{T}} \mathbf{J}_{\mathbf{uu}} (\mathbf{u} - \mathbf{u}^{opt}). \tag{17}$$

Solving $J_{\mathbf{u}} = 0$ (11) for $\mathbf{u}^{opt} = -\mathbf{J}_{\mathbf{uu}}^{-1}\mathbf{J}_{\mathbf{ud}}\mathbf{d}$ and inserting into (17) yields (note that $\mathbf{J}_{\mathbf{uu}}$ is symmetric):

$$\begin{aligned} L &= \frac{1}{2}(\mathbf{u} - \mathbf{J}_{\mathbf{uu}}^{-1}\mathbf{J}_{\mathbf{ud}}\mathbf{d})^{\mathrm{T}} \mathbf{J}_{\mathbf{uu}} (\mathbf{u} - \mathbf{J}_{\mathbf{uu}}^{-1}\mathbf{J}_{\mathbf{ud}}\mathbf{d}) \\ &= \frac{1}{2}(\mathbf{u}^{\mathrm{T}} - \mathbf{d}^{\mathrm{T}}\mathbf{J}_{\mathbf{ud}}^{\mathrm{T}}\mathbf{J}_{\mathbf{uu}}^{-T}) \mathbf{J}_{\mathbf{uu}} (\mathbf{u} - \mathbf{J}_{\mathbf{uu}}^{-1}\mathbf{J}_{\mathbf{ud}}\mathbf{d}) \\ &= \frac{1}{2}(\mathbf{u}^{\mathrm{T}}\mathbf{J}_{\mathbf{uu}}^{\mathrm{T}} - \mathbf{d}^{\mathrm{T}}\mathbf{J}_{\mathbf{ud}}^{\mathrm{T}}) \mathbf{J}_{\mathbf{uu}}^{-1} (\mathbf{J}_{\mathbf{uu}}\mathbf{u} - \mathbf{J}_{\mathbf{ud}}\mathbf{d}) \\ &= \frac{1}{2} J_{\mathbf{u}}^{\mathrm{T}} \mathbf{J}_{\mathbf{uu}}^{-1} J_{\mathbf{u}} = \frac{1}{2} \left|\left| \mathbf{J}_{\mathbf{uu}}^{-1/2} J_{\mathbf{u}} \right|\right|_2^2 . \qquad \square \end{aligned} \tag{18}$$

At the optimum, $\mathbf{u}_{opt}$, the gradient $J_{\mathbf{u}} = 0$, and the loss $L = 0$. Around the optimum $J_{\mathbf{u}} \neq 0$, the loss L is equal to the norm the weighted gradient, where the weight factor is $\mathbf{J}_{\mathbf{uu}}^{-1/2}$, Figure 4,

Remark 1. (Effect of constraints). The above analysis is locally valid for a system where all active constraints are known and have been satisfied, $g(\mathbf{u}, \mathbf{d}) = 0$. If an active constraint is not satisfied exactly, $g(\mathbf{u}, \mathbf{d}) = \epsilon$, then the effect on the objective function will be given by the corresponding Lagrangian multiplier [Nocedal and Wright, 2006]:

$$\lambda = \partial J / \partial \epsilon \tag{19}$$

A perturbation of the constraints ϵ has therefore a first order effect on the cost function, while from (16), a small change in $J_{\mathbf{u}}$ has a second order effect on the cost. From an economic point of view, tight control of the active constraints will generally be more important than tight control of the unconstrained variable $\mathbf{c}$.

3. MINIMIZING THE GRADIENT VS. MINIMIZING THE LOSS

Theorem 1 shows that a controlled variable which minimizes $||J_{\mathbf{u}}||_2$, does not necessarily minimize the loss L. One case, where $\mathbf{J}_{\mathbf{uu}}^{-1/2}$ has no effect is, when it is orthogonal, $\mathbf{J}_{\mathbf{uu}}^{-1/2} = \mathbf{J}_{\mathbf{uu}}^{-1/2^{\mathrm{T}}}$, or scalar. In the next sections, we examine in which further cases an $\mathbf{H}$ which minimizes $||J_{\mathbf{u}}||_2$ is the same that minimizes the loss L.

3.1 Enough measurements, no noise, full $\mathbf{H}$*: same* $\mathbf{H}$

If it is possible to have zero loss (no noise and sufficient measurements), optimal operation corresponds to $J_{\mathbf{u}} = 0$. Then, $\mathbf{J}_{\mathbf{uu}}^{-1/2}$ has no effect. Assume that $\mathbf{y}$ contains all available information, then we require that

$$J_{\mathbf{u}} = \mathbf{H}\mathbf{y}. \tag{20}$$

Theorem 2. (Null space method, no noise). Given a linear model as in (8), with a sufficient number of independent measurements ($n_{\mathbf{y}} \geq n_{\mathbf{u}} + n_{\mathbf{d}}$) and no noise ($\mathbf{n}^{\mathbf{y}} = 0$), selecting

$$\mathbf{H} = [\, \mathbf{J}_{\mathbf{uu}} \;\, \mathbf{J}_{\mathbf{ud}} \,] \left[\tilde{\mathbf{G}}^{\mathbf{y}} \right]^{-1} \tag{21}$$

and controlling $\mathbf{H}\mathbf{y} = 0$ gives zero loss from optimal operation. Here, $\tilde{\mathbf{G}}^{\mathbf{y}}$ is the gain matrix of any subset of $n_u + n_d$ measurements.

Proof The gradient from (10) is:

$$J_{\mathbf{u}} = [\, \mathbf{J}_{\mathbf{uu}} \;\, \mathbf{J}_{\mathbf{ud}} \,] \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} \tag{22}$$

We want to eliminate the variables $[\mathbf{u}, \mathbf{d}]^{\mathrm{T}}$ using the available measurements,

$$\mathbf{y} = \tilde{\mathbf{G}}^{\mathbf{y}} \begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix}. \tag{23}$$

Solving for $[\mathbf{u}^{\mathrm{T}}, \; \mathbf{d}^{\mathrm{T}}]^{\mathrm{T}}$,

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{d} \end{bmatrix} = \left[\tilde{\mathbf{G}}^{\mathbf{y}} \right]^{-1} \mathbf{y}, \tag{24}$$

and inserting into (22) gives:

$$\begin{aligned} J_{\mathbf{u}} &= [\, \mathbf{J}_{\mathbf{uu}} \;\, \mathbf{J}_{\mathbf{ud}} \,] \left[\tilde{\mathbf{G}}^{\mathbf{y}} \right]^{-1} \mathbf{y} \\ &= \mathbf{H}\mathbf{y} \end{aligned} \tag{25}$$

Controlling $\mathbf{c} = \mathbf{H}\mathbf{y} = 0$ results in zero loss. $\square$

- This is a new derivation of the null space method reported in [Alstad and Skogestad, 2007]. It shows that the optimal controlled variable found by self-optimizing control is identical to the gradient,

$$\mathbf{c} = J_{\mathbf{u}} = \mathbf{H}\mathbf{y}. \tag{26}$$

3.2 Enough measurements, noise, full $\mathbf{H}$*: same* $\mathbf{H}$

The case of finding a controlled variable combination, which minimimizes the loss in presence of sufficient (noisy) measurements and a full $\mathbf{H}$ matrix is addressed in the "exact local method" Alstad et al. [2009]. First, we scale the disturbances and the noise, such that $\mathbf{d} = \mathbf{W}_{\mathbf{d}}\mathbf{d}'$ and

$\mathbf{n^y} = \mathbf{W_{n^y}} \mathbf{n^{y\prime}}$ where $\left|\left| \left[\mathbf{d}^{\prime\mathrm{T}} \;\; \mathbf{n}^{\mathbf{y}\mathrm{T}\prime} \right] \right|\right|_2 \leq 1$ and $\mathbf{W_d}$ and $\mathbf{W_{n^y}}$ are diagonal scaling matrices of appropriate sizes. Then we:

(1) Express L as a function of $\mathbf{H}$, $\mathbf{d}$ and $\mathbf{n^y}$ (assuming $\mathbf{c} = \mathbf{H}(\mathbf{y} + \mathbf{n^y}) = 0$)
(2) Then find an expression for the worst-case loss $L(\mathbf{H})$ (worst-case w.r.t. $\mathbf{d}$ and $\mathbf{n^y}$); which is the maximum singular value $\bar{\sigma}(\mathbf{M})$. Here
$$\mathbf{M} = \mathbf{J_{uu}^{1/2}} (\mathbf{H}\mathbf{G^y})^{-1} \mathbf{H}\mathbf{Y} \tag{27}$$
and
$$\mathbf{Y} = [\mathbf{F}\mathbf{W_d} \quad \mathbf{W_{n^y}}], \tag{28}$$
where
$$\mathbf{F} = \frac{\partial \mathbf{y}^{opt}}{\partial \mathbf{d}} \tag{29}$$
is the optimal measurement sensitivity matrix, see Halvorsen et al. [2003]. (Kariwala et al. [2008] have shown that the average loss is given by $||\mathbf{M}||_F$, where $||\cdot||_F$ denotes the Frobenius norm)
(3) Find a convex problem formulation for finding $\mathbf{H}$ (see Alstad et al. [2009]).

The convex problem for finding an $\mathbf{H}$ which minimizes the average and worst case loss for a given set of disturbances is [Alstad et al., 2009]:
$$\begin{aligned} &\min_{\mathbf{H}} ||\mathbf{H}[\mathbf{F}\mathbf{W_d} \quad \mathbf{W_{n^y}}]||_F \\ &\text{subject to } \mathbf{H}\mathbf{G^y} = \mathbf{Q}, \end{aligned} \tag{30}$$
where $\mathbf{Q}$ is any nonsingular $n_\mathbf{u} \times n_\mathbf{u}$ matrix, and $\mathbf{F}$ as defined in (29). Here, too, $\mathbf{J_{uu}}$ is not needed for fining the best measurement combination. However, if we want to know the actual worst case loss, we need $\mathbf{J_{uu}}$ in (27).

- In the case of no structural constraints on $\mathbf{H}$, it is found that $\mathbf{J_{uu}}$ is not needed for finding the minimum. That is, a controlled variable which minimizes $||J_\mathbf{u}||_2$ minimizes also the loss $L = \frac{1}{2} \left|\left| \mathbf{J_{uu}^{-1/2}} J_\mathbf{u} \right|\right|_2^2$.

3.3 Structural constraints on $\mathbf{H}$*: not the same* $\mathbf{H}$

In the above cases, we used all measurements $\mathbf{y}$ to generate the controlled variables as linear combinations of all measurements. In practice however, there are often structural constraints on the controlled variables. Examples for structural constraints include controlling single measurements, or using only two measurements from the rectifier section and two measurements from the stripping section of a distillation column. When we have to decide between two or more controlled structures, the norm of the gradient (if it is nonzero) does no longer give accurate information about what controlled variable is best. To be able to make a good decision in these cases, we need to consider norm of the weighted gradient $L = \frac{1}{2} \left|\left| \mathbf{J_{uu}^{-1/2}} J_\mathbf{u} \right|\right|_2^2$.

Consider a process with
$$\mathbf{J_{uu}} = \begin{bmatrix} 244 & 222 \\ 222 & 202 \end{bmatrix} \text{ and } \mathbf{J_{ud}} = \begin{bmatrix} 10 \\ 10 \end{bmatrix} \tag{31}$$
and assume that we have the choice between the two controlled variables
$$\mathbf{c}_1 = \begin{bmatrix} 1 & 0.275 \\ 2.78 & 2 \end{bmatrix} \mathbf{u} + \begin{bmatrix} 10 \\ 10 \end{bmatrix} d \text{ and } \mathbf{c}_2 = \begin{bmatrix} 5 & -4.9130 \\ 2.1826 & 2 \end{bmatrix} \mathbf{u} + \begin{bmatrix} 10 \\ 0 \end{bmatrix} d. \tag{32}$$

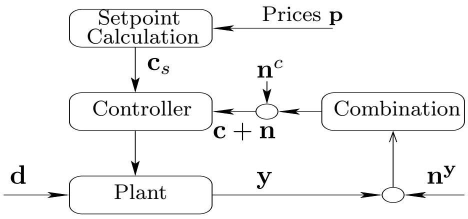

Fig. 5. Feedback control structure with setpoint calculation

For a disturbance $d = 1$, the corresponding gradients are
$$J_\mathbf{u}(\mathbf{c}_1 = 0) = [\,1\;1\,]^\mathrm{T} \text{ and } J_\mathbf{u}(\mathbf{c}_2 = 0) = [\,-1\;1\,]^\mathrm{T}. \tag{33}$$
The norm of the gradient is in both cases $||J_\mathbf{u}(\mathbf{c}_1 = 0)||_2 = ||J_\mathbf{u}(\mathbf{c}_2 = 0)||_2 = \sqrt{2}$. This indicates that the two controlled variables give equivalent performance. However, if we consider the loss imposed by the two different control structures, we have that
$$L(\mathbf{c}_1) = \frac{1}{2} \left|\left| \mathbf{J_{uu}^{-1/2}} J_\mathbf{u}(\mathbf{c}_1) \right|\right|_2^2 = 0.25 \tag{34}$$
and
$$L(\mathbf{c}_2) = \frac{1}{2} \left|\left| \mathbf{J_{uu}^{-1/2}} J_\mathbf{u}(\mathbf{c}_2) \right|\right|_2^2 = 0.111.25 \tag{35}$$

- When we want to compare sets of controlled variables with each other, we need to examine the loss, as the gradient does not give sufficient information.
- When searching for the best linear combination for a given set of measurements, it is sufficient to consider the gradient.

4. VARYING SETPOINTS FOR THE CONTROLLED VARIABLES

Many processes are operated such that the setpoints of the controlled variables are changed, when for example product prices $\mathbf{p}$ and specifications change, Figure 5. To handle this in the framework above, we consider the reason for the setpoint change as a *measured disturbance*. The relationship between the measurements $\mathbf{y}$ and the controlled variables is
$$\mathbf{c} = \mathbf{H}\mathbf{y}, \tag{36}$$
and the relationship between the setpoint change and the prices $\mathbf{p}$ is
$$\mathbf{c}_{set} = \mathbf{H}_{set}\mathbf{p}. \tag{37}$$
We define $\hat{J}_\mathbf{u} = \mathbf{c} - \mathbf{c}_{set}$
$$\begin{aligned} \hat{J}_\mathbf{u} &= \mathbf{H}\mathbf{y} - \mathbf{H}_{set}\mathbf{p} = [\,\mathbf{H}\; -\mathbf{H}_{set}\,] \begin{bmatrix} \mathbf{y} \\ \mathbf{p} \end{bmatrix} \\ &= \mathbf{H}_{aug}\mathbf{y}_{aug} \end{aligned} \tag{38}$$
The gain matrices are augmented according to
$$\mathbf{G}^\mathbf{y}_{aug} = \begin{bmatrix} \mathbf{G}^\mathbf{y}_{n_y \times n_u} \\ 0_{n_p \times n_u}, \end{bmatrix}, \quad \mathbf{G}^\mathbf{y}_{aug,\mathbf{d}} = \begin{bmatrix} \mathbf{G^y} & 0_{n_p \times n_p} \\ 0_{n_p \times n_d} & \mathbf{I}_{n_p \times n_p} \end{bmatrix}, \tag{39}$$
and the scaling matrices according to
$$\mathbf{W}_{\mathbf{d},aug} = \begin{bmatrix} \mathbf{W_d} & 0_{n_d \times n_p} \\ 0_{n_p \times n_d} & \mathbf{W_p}_{n_p \times n_p} \end{bmatrix}, \tag{40}$$
$$\mathbf{W}_{\mathbf{n},aug} = \begin{bmatrix} \mathbf{W_n} & 0_{n_y \times n_p} \\ 0_{n_p \times n_y} & \mathbf{W_{np}}_{n_p \times n_p} \end{bmatrix}. \tag{41}$$
Here, $\mathbf{W_p}$ and $\mathbf{W_{np}}$ are diagonal matrices with the expected price variations and uncertainties, respectively.

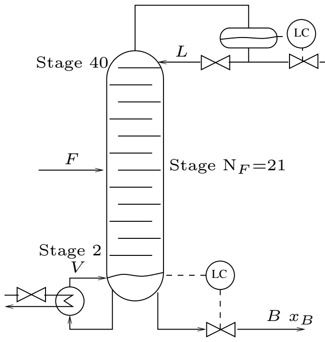

Fig. 6. Distillation column

If the prices are known exactly, $\mathbf{W}_{\mathbf{np},aug} = 0$. The sensitivity matrix $\mathbf{F}_{aug}$ may be found by reoptimization or by evaluating $\mathbf{F}_{aug} = \mathbf{G}^{\mathbf{y}}_{aug,d} - \mathbf{G}^{\mathbf{y}}_{aug}\mathbf{J}^{-1}_{\mathbf{uu}}\mathbf{J}_{\mathbf{ud},aug}$, [Alstad and Skogestad, 2007]. After the problem has been formulated, the optimal $\mathbf{H}_{aug}$, which minimizes the loss is found by solving

$$\begin{aligned} &\min_{\mathbf{H}_{aug}} ||\mathbf{H}_{aug}[\mathbf{F}_{aug}\mathbf{W}_{\mathbf{d},aug} \quad \mathbf{W}_{\mathbf{n},aug}]||_F \\ &\text{subject to } \mathbf{H}_{aug}\mathbf{G}^{\mathbf{y}} = \mathbf{Q} \end{aligned} \tag{42}$$

De-partitioning $\mathbf{H}_{aug} = [\mathbf{H} \quad -\mathbf{H}_{set}]$, the controlled variables and the setpoint updates are

$$\mathbf{c} = \mathbf{H}\mathbf{y} \text{ and } \mathbf{c}_{set} = \mathbf{H}_{set}\mathbf{p}.$$

5. DISTILLATION CASE STUDY

5.1 Problem description and setup

A binary distillation column is used to demonstrate the results. The column model is taken from Skogestad [1997]. It is controlled in the LV configuration and has 41 stages, Figure 6. We assume that the temperatures on stage 9, 16, 24, and 33 are measured and that they can be used for control, i.e. $\mathbf{y} = [T_9, T_{16}, T_{24}, T_{33}]^{\mathrm{T}}$. The temperatures are calculated as a linear function of the liquid composition for the respective stages i,

$$T_i = 10(1 - x_i) \tag{43}$$

This corresponds to a pure product boiling point difference of 10°C. In order to be able to sell the top product, a purity of 99% is required for the distillate D. This is considered an active constraint, and is controlled to its setpoint using the liquid reflux L. The remaining degree of freedom $\mathbf{u}$ (the boilup V) can be used to maximize the profit which is the same as minimizing the difference between the costs for the feed and evaporation, and the profit from selling the purified products:

$$J = -(p_D D + p_B B - p_V V - p_F F) \tag{44}$$

Assuming the price for the feed is equal to the price of the bottom product, $p_F = p_B$, and introducing the overall mass balance, the cost function can be simplified to

$$J = p_V\left(\frac{p_F - p_D}{p_V}D + V\right) = p_V(p'D + V). \tag{45}$$

The only parameter which affects the minimum is the relative price difference of the feed and the distillate, $p' = (p_F - p_D)/p_V$, We assume $p' = -64$ currency units.

As disturbances, we consider the flow rate F, composition z and liquid fraction q of the feed. These disturbances are detectable through the measurement model in deviation variables:

$$\mathbf{y} = \mathbf{G}^{\mathbf{y}}\mathbf{u} + \mathbf{G}^{\mathbf{y}}_{\mathbf{d}}\mathbf{d} \tag{46}$$

In addition, we assume that the product prices change and are known. We use the prices to update the setpoint of $\mathbf{c}$ The self-optimizing controlled variable is selected as a linear combination of the four tray measurements. The augmented gain matrices and the augmented optimal sensitivity matrix are

$$\mathbf{G}^{\mathbf{y}}_{aug} = \begin{bmatrix} 1.71 \\ 3.22 \\ 1.36 \\ 0.20 \\ 0.00 \end{bmatrix} \quad \mathbf{G}^{\mathbf{y}}_{\mathbf{d},aug} = \begin{bmatrix} -8.14 & -3.67 & -1.50 & 0 \\ -12.28 & -11.09 & -2.75 & 0 \\ -4.36 & -9.81 & -1.44 & 0 \\ -0.65 & -1.45 & -0.21 & 0 \\ 0.00 & 0.00 & 0.00 & 1 \end{bmatrix} \tag{47}$$

$$\mathbf{F}_{aug} = \begin{bmatrix} -2.63 & -2.37 & -0.49 & -0.0055 \\ -1.90 & -8.65 & -0.84 & -0.0103 \\ 0.02 & -8.78 & -0.64 & -0.0044 \\ 0.01 & -1.30 & -0.09 & -0.0007 \\ 0.00 & 0.00 & 0.00 & 1.0000 \end{bmatrix} \tag{48}$$

The weighting matrix $\mathbf{W}_{\mathbf{n},aug}$ is chosen such that all temperature measurements have an uncertainty of 0.5°C, and the price uncertainty is zero. The expected variation in the disturbances is captured in

$$\mathbf{W}_{\mathbf{d},aug} = \mathrm{diag}([0.1,\ 0.1,\ 0.1,\ 6.4]),$$

which corresponds to 10% variation in every disturbance variable. The corresponding second derivatives are

$$\begin{aligned} \mathbf{J}_{\mathbf{uu}} &= 4.85 \\ \mathbf{J}_{\mathbf{ud}} &= [-15.64,\ -3.68,\ -2.87,\ 0.02]. \end{aligned} \tag{49}$$

This gives a controlled variable combination $\mathbf{c} = \mathbf{H}\mathbf{y}$ with

$$\mathbf{H} = [0.23,\ 0.69,\ -0.28,\ -0.04] \tag{50}$$

and the setpoint is updated using $\mathbf{c}_{set} = \mathbf{H}_{set}p'$ with

$$\mathbf{H}_{set} = 0.0071. \tag{51}$$

The first order loss from optimality estimate is calculated according to $L = \left|\left|\mathbf{J}^{-1/2}_{\mathbf{uu}} J_{\mathbf{u}}\right|\right|^2_2$ or alternatively according to Halvorsen et al. [2003] as

$$\left|\left|\mathbf{J}^{1/2}_{\mathbf{uu}}(\mathbf{H}_{aug}\mathbf{G}^{\mathbf{y}}_{aug})^{-1}\mathbf{H}_{aug}\left[\mathbf{F}_{aug}\mathbf{W}_{\mathbf{d},aug}\ \mathbf{W}_{\mathbf{n},aug}\right]\right|\right|^2_2, \tag{52}$$

and equals 1.4869 currency units.

5.2 Simulations

We consider disturbances in the flow rate, $\Delta F = 10\%$, the feed concentration, $\Delta z = 10\%$, the feed liquid fraction, $\Delta q = -10\%$, and the price, $\Delta p' = 10\%$. The disturbance scenario is given in Figure 7, and the resulting profit is plotted together with the inputs in Figure 8.

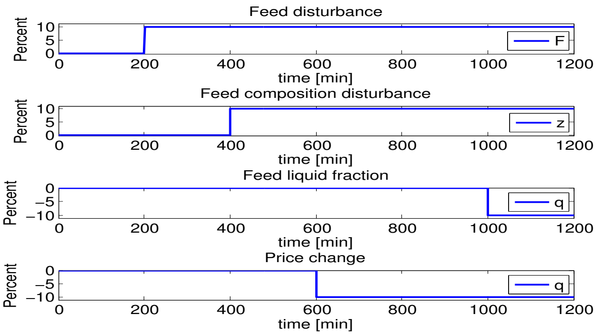

Fig. 7. Disturbance trajectories

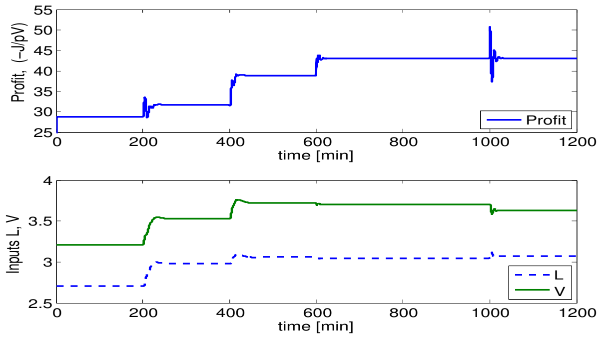

Fig. 8. Profit and inputs for the distillation column

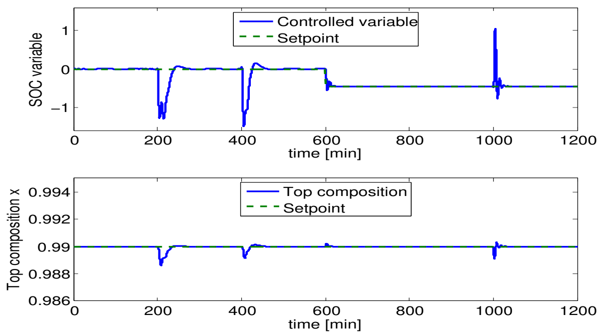

Fig. 9. Controlled variables: Self-optimizing controlled variable and top composition

In Figure 9, the controlled variables are given together with their setpoints. The self-optimizing controlled variable is nicely controlled back to the setpoint after a disturbance enters the process. As long as the prices are constant, the setpoint is zero. When the price ratio p' changes, the setpoint is adapted to the new value. The top composition is controlled well at its specification, as can be noted from the plotting scale.

6. DISCUSSION AND CONCLUSIONS

We have given a first order approximation of the loss as the weighted norm of the gradient, and we have shown that if all measurements are used, the weighting is not required. However, when selecting between different sets of controlled variables, we need to consider the weighted gradient, because neglecting the weighting can be seriously misleading. The previously published "exact local method" is basically indicating how close the norm of the weighted gradient is to zero when a particular set of controlled variables $\mathbf{c} = \mathbf{Hy}$ is used.

The key points are to weight the gradient when approximating it and to include noise in the analysis. Otherwise we may approximate the gradient well while still suffering from unnecessary economic loss. The controlled variables obtained by this method have a robustness against measurement noise. However, the underlying linear model and the cost function parameters are assumed to be locally exact, that is all the uncertainty is assumed to be taken care of in the measurement noise and disturbances.

Our analysis is based on the assumption that the active constraints do not to change. If the active constraints change, it is necessary to adjust the control structure to satisfy the new active set. However, if there are unconstrained degrees of freedom in the new active set, the above analysis can be reapplied.

The second part of this paper deals with disturbances which do not enter through the model. By considering them as additional measurements, this can be formulated in terms of minimizing the weighted gradient, and the techniques from self-optimizing control can be used to update the setpoints to ensure optimal operation for all considered process and price disturbances

REFERENCES

Vidar Alstad and Sigurd Skogestad. Null space method for selecting optimal measurement combinations as controlled variables. *Industrial & Engineering Chemistry Research*, 46:846–853, 2007.

Vidar Alstad, Sigurd Skogestad, and Eduardo Shigueo Hori. Optimal measurement combinations as controlled variables. *Journal of Process Control*, 19(1):138–148, 2009.

Dominique Bonvin, Bala Srinivasan, and David Ruppen. Dynamic Optimization in the Batch Chemical Industry. *in: Chemical Process Control-VI*, 2001.

Yi Cao. Self-optimizing control structure selection via differentiation. In *Proceedings of the European Control Conference*, pages 445–453, 2003.

Yi Cao. Direct and indirect gradient control for static optimisation. *International Journal of Automation and Computing*, 2:60–66, 2005.

Ivar J. Halvorsen and Sigurd Skogestad. Indirect on-line optimization through setpoint control. *AIChE 1997 Annual Meeting, Los Angeles; paper 194h*, 1997a.

Ivar J. Halvorsen and Sigurd Skogestad. Optimizing control of petlyuk distillation: Understanding the steady-state behavior. *Computers & Chemical Engineering*, 21 (Supplement 1):S249 – S254, 1997b.

Ivar J. Halvorsen and Sigurd Skogestad. Optimal operation of petlyuk distillation: steady-state behavior. *Journal of Process Control*, 9(5):407 – 424, 1999.

Ivar J. Halvorsen, Sigurd Skogestad, John C. Morud, and Vidar Alstad. Optimal selection of controlled variables. *Industrial & Engineering Chemistry Research*, 42(14): 3273–3284, 2003.

Vinay Kariwala, Yi Cao, and S. Janardhanan. Local self-optimizing control with average loss minimization. *Industrial & Engineering Chemistry Research*, 47:1150–1158, 2008.

Manfred Morari, George Stephanopoulos, and Yaman 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. analsysis of the optimizing control structures. *AIChE Journal*, 26(2):220–232, 1980.

Jorge Nocedal and Stephen Wright. *Numerical Optimization*. Springer, 2006.

Sigurd Skogestad. Dynamics and control of distillation columns: A tutorial introduction. *Chemical Engineering Research and Design*, 75(6):539 – 562, 1997.

Bala Srinivasan, Lorenz T. Biegler, and Dominique Bonvin. Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function. *Computers & Chemical Engineering*, 32:572–279, 2008.