

# SELF-OPTIMIZING CONTROL: A DISTILLATION CASE STUDY

Sigurd Skogestad \*

\* Department of Chemical Engineering,  
Norwegian University of Science and Technology,  
N-7491 Trondheim, Norway  
Email: skoge@chembio.ntnu.no; Fax: +47-7359-4080

**Abstract:** The following important question is frequently overlooked: Which variables should we control? The answer is related to finding a simple and robust way of implementing the optimal operating strategy. It is argued that the controlled variables should be those that give “self-optimizing control”, which is when acceptable operation is achieved with constant setpoints for the controlled variables. The idea is applied to propane-propylene distillation case study.

**Keywords:** Chemical process control, control structure design, plantwide control

## 1. INTRODUCTION

If we consider the control system in a chemical plant, then we find that it is structured hierarchically into several layers, each operating on a different time scale. Typically, layers include scheduling (weeks), site-wide optimization (day), local optimization (hour), supervisory/predictive control (minutes) and regulatory control (seconds); see Figure 1. The task in each layer, e.g. optimization or control, may be performed by a computer algorithm or it may be performed manually (by an operator or engineer). The layers are interconnected through the controlled variables. More precisely,

the **controlled variables** ( $c$ ) are the (internal) variables that link two layers in a control hierarchy, whereby the upper layer computes the setpoint ( $c_s$ ) to be implemented by the lower layer.

Which should these internal controlled variables  $c$  be? That is, what should we control? More generally, the issue of selecting controlled variables is the first subtask in the **control structure design** problem (Foss, 1973); (Morari, 1982); (Skogestad and Postlethwaite, 1996) :

- (1) *Selection of controlled variables  $c$*  (variables with setpoints  $c_s$ )
- (2) *Selection of manipulated variables*
- (3) *Selection of measurements  $v$*  (for control purposes including stabilization)
- (4) *Selection of a control configuration* (structure of the controller that interconnects measurements/setpoints and manipulated variables)
- (5) *Selection of controller type* (control law specification, e.g., PID, decoupler, LQG, etc.).

Note that these *structural decisions* need to be made before we can start the actual design the controller. In most cases the control structure is designed by a mixture between a top-down consideration of control objectives and which degrees of freedom are available to meet these (tasks 1 and 2), combined with a bottom-up design of the control system, starting with the stabilization of the process (tasks 3,4 and 5). Even though control engineering is well developed in terms of providing optimal control algorithms, it is clear that most of the existing theories provide little help when it comes to making structural decisions.

The method presented in this paper for selecting controlled variables follows the ideas of Morari *et al.* (1980) and Skogestad and Postlethwaite (1996) and is

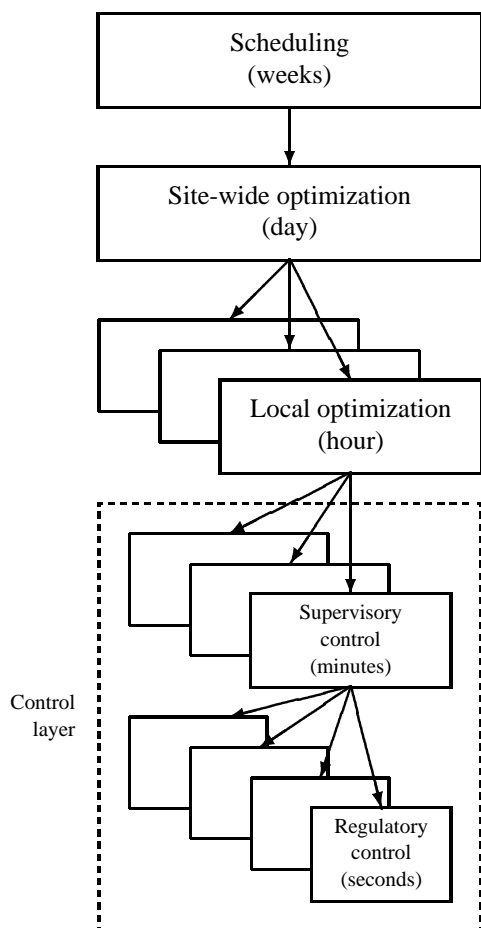


Fig. 1. Typical control hierarchy in a chemical plant.

very simple. The basis is to define mathematically the quality of operation in terms of a scalar cost function  $J$  to be minimized. To achieve truly optimal operation we would need a perfect model, we would need to measure all disturbances, and we would need to solve the resulting dynamic optimization problem on-line. This is unrealistic in most cases, and the question is if it is possible to find a simpler implementation which still operates satisfactorily (with an acceptable loss). More precisely, the **loss**  $L$  is defined as the difference between the actual value of the cost function obtained with a specific control strategy, and the truly optimal value of the cost function, i.e.  $L = J - J_{opt}$ . The simplest operation would result if we could select controlled variables such that we obtained acceptable operation with constant setpoints, thus effectively turning the complex optimization problem into a simple feedback problem and achieve “self-optimizing control”:

**Self-optimizing control** is when we can achieve an acceptable loss with constant setpoint values for the controlled variables

(The reader is probably familiar with the term self-regulation, which is when acceptable dynamic control performance can be obtained with constant manipulated inputs. Self-optimizing control is a direct gen-

eralization to the case where we can achieve acceptable (economic) performance with constant controlled variables.)

A simple example of self-optimizing control is the process of baking a cake, where the operation is indirectly kept close to its optimum (“a well-baked cake”) by controlling the oven temperature and baking time at the setpoints given in the cook book (which in this case is the “optimizer”).

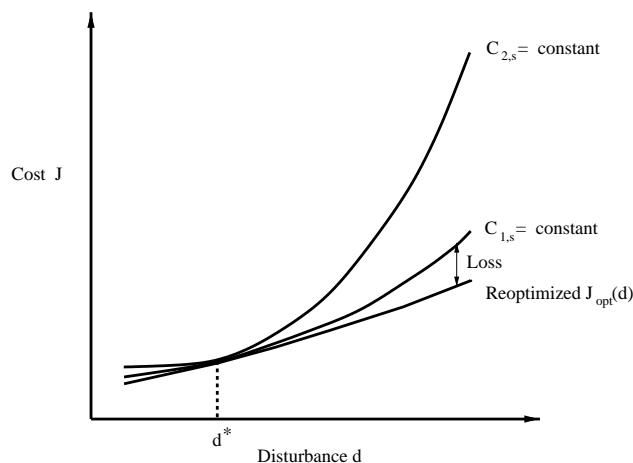


Fig. 2. Loss imposed by keeping constant setpoint for the controlled variable

The idea is further illustrated in Figure 2, where we see that there is a loss if we keep a constant setpoint rather than reoptimizing when a disturbance moves the process away from its nominally optimal operating point (denoted  $*$ ). For the case illustrated in the figure it is better (with a smaller loss) to keep the setpoint  $c_{1s}$  constant than to keep  $c_{2s}$  constant.

An additional concern with the constant setpoint policy is that there will always be an implementation error  $d_c = c - c_s$ , e.g. caused by measurement error. The implementation error may cause a large additional loss if the optimum surface is “sharp”. To be more specific, we may, as illustrated in Figure 3, distinguish between three classes of problems when it comes to the actual implementation:

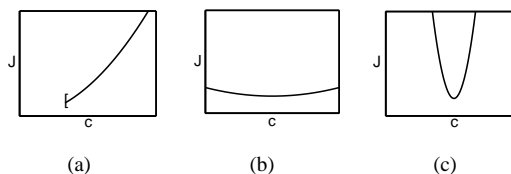


Fig. 3. Implementing the controlled variable

(a) *Constrained optimum: Implementation easy.* In the figure is shown the case where the minimum value of the cost  $J$  is obtained for  $c = c_{min}$ . In this case there is no loss imposed by keeping a constant  $c_s = c_{min}$ . In addition, implementation of an “active” constraint is usually easy, e.g., it is easy to keep a valve closed.

(b) *Unconstrained flat optimum: Implementation easy.*

In this case the cost is insensitive to value of the controlled variable  $c$ , and implementation is again easy.

(c) *Unconstrained sharp optimum: Implementation difficult.* The more difficult problems for implementation is when the cost (operation) is sensitive to value of the controlled variable  $c$ . In this case, we want to find another controlled variable  $c$  in which the optimum is flatter.

The latter unconstrained problems are the focus of this paper.

Inspired by the work of Findeisen (e.g. see Findeisen *et al.* (1980)), Morari *et al.* (1980) gave a clear description of what we here denote self-optimizing control, including a procedure for selecting controlled variables based on evaluating the loss. However, it seems that so far nobody has followed up on their ideas. One reason was probably that no good example was given in the paper.

The main objective of this paper is to demonstrate, with a few examples, that the issue of selecting controlled variables is very important and that the concept of self-optimizing control provides a useful tool.

## 2. THE OPTIMIZATION PROBLEM

The optimizing control problem can be formulated as

$$\min_u J(u, d) \quad (1)$$

subject to the inequality constraints

$$g(u, d) \leq 0 \quad (2)$$

where  $u$  are the  $N_u$  independent variables we can affect (“base set” for the optimization degrees of freedom), and  $d$  are independent variables we can not affect (disturbances). The solution to this problem (if a feasible solution exists) is  $u_{opt}(d)$ , where

$$\min_u J(u, d) = J(u_{opt}(d), d) = J_{opt}(d) \quad (3)$$

There are two main issues when it comes to optimizing control. The first is the mathematical and numerical problem of solving the optimization problem in (1) to obtain the optimal operating point. The optimization problem may be very large, with hundreds of thousands of equations and hundreds of degrees of freedom (e.g. for a complete ethylene plant), but with today's computers and optimization methods this problem is solvable, and it is indeed solved routinely today in some plants. The second issue, the focus of this paper, is how the optimal solution should be implemented in practice. Surprisingly, this issue has received much less attention. The objective is to, if possible,

Find a set of controlled variables  $c$ , which when specified at constant setpoints, give an acceptable loss ( $c$  replaces  $u$  as degrees of freedom, and a special case is to select  $c = u$ ).

The analysis in this paper is based on steady-state models and use of constant setpoints  $c_s$  at each steady-state (operating point). To analyze the effect of disturbances we may time-average various steady-states. The main justification for using a steady-state analysis is that the economic performance is primarily determined by steady-state considerations.

## 3. PROCEDURE FOR SELECTING CONTROLLED VARIABLES

**Step 1: Degree of freedom analysis.** Determine the number of degrees of freedom ( $N_u$ ) available for optimization, and identify a base set ( $u$ ) for the degrees of freedom.

**Step 2: Cost function and constraints.** Define the optimal operation problem by formulating a scalar cost function  $J$  to be minimized for optimal operation, and specify the constraints that need to be satisfied.

**Step 3: Identify the most important disturbances  $d$  (uncertainty).** These may be caused by

- Errors in the assumed (nominal) model used in the optimization
- Disturbances  $d-d^*$  (including parameter changes) that occur after the optimization
- Implementation errors ( $d_c$ ) for the controlled variables  $c$  (e.g. due to measurement error or poor control)

From this one must define the set of disturbances  $\mathcal{D}$  and set of implementation errors  $\mathcal{D}_c$  to be considered. Often it is a finite set of disturbance combinations, for example, consisting of the extreme values for the individual disturbances. In addition, one must determine how to evaluate the mean cost function  $J_{mean}$ . There are many possibilities, for example

- (1) Average cost for a finite set of disturbances
- (2) Mean cost from *Monte-Carlo* evaluation of a distribution of  $d$  and  $d_c$ .
- (3) Worst-case loss

**Step 4: Optimization.**

- (1) First solve the nominal optimization problem, that is, we find  $u_{opt}(d^*)$  by solving  $\min_u J(u, d^*)$  where  $d^*$  is nominal value of the disturbances (parameters). From this may one also obtain a table with the nominal optimal values for all other variables (including the candidate controlled variables).
- (2) In most cases (unless it involves too much effort) we then solve the optimization problem for the disturbances  $d$  in question (defined in step 3). This is needed to check whether there exists a feasible solution  $u_{opt}(d)$  for all

disturbances  $d$ , and to find the optimal cost  $J(u_{opt}, d)$  needed if we want to evaluate the loss.

**Step 5: Identify candidate controlled variables.** We normally implement the constraints that are active for all disturbances (“active constraint control”). This leaves  $N_{opt, free}$  degrees of freedom for which we want to select controlled variables. Typically, these are measured variables or simple combinations thereof. For example, based on the optimization in step 4, one may look for variables which optimal value is only weakly dependent of disturbances. The variable should also be easy to control and measure, and it should be sensitive to changes in the manipulated inputs (Skogestad and Postlethwaite, 1996). Insight and experience may also be helpful at this stage, because the possible number of combinations may be extremely large.

**Step 6: Evaluation of loss.** We compute the mean value of the loss for alternative sets of controlled variables  $c$ . This is done by evaluating the loss

$$L(u, d) = J(u, d) - J(u_{opt}(d), d) \quad (4)$$

$$u = f_c(c_s + d_c, d)$$

with fixed setpoints  $c_s$  for the defined disturbances  $d \in \mathcal{D}$  and implementation errors  $d_c \in \mathcal{D}_c$ . We usually select the setpoints as the nominal optimal values,  $c_s = c_{opt}(d^*)$ , but it is also possible to let the value of  $c_s$  be subject to an optimization.

**Step 7: Further analysis and selection.** We select for further consideration the sets of controlled variables with acceptable loss (and which thus yield self-optimizing control). These could then be analyzed to see if they are adequate with respect to other criteria that may be relevant, such like the region of feasibility and the expected dynamic control performance (input-output controllability)

#### 4. DISTILLATION CASE STUDY

We consider a binary mixture with constant relative volatility  $\alpha = 1.12$  to be separated in a distillation column with 110 theoretical stages and the feed entering at stage 39 (counted from the bottom with the reboiler as stage 1). Nominally, the feed contains 65 mole% of light component ( $z_F = 0.65$ ) and is saturated liquid ( $q_F = 1.0$ ). This is “column D” of Skogestad and Morari (1988) and represents a propylene-propane splitter where propylene (light component) is taken overhead as a final product with at least 99.5% purity, whereas unreacted propane (heavy component) is recycled to the reactor for reprocessing. We assume that the feed rate is given at 1 kmol/min and that there is no capacity limit in the column.

**Step 1: Degree of freedom analysis** With a given feed stream and a specified pressure, a conventional two-product distillation column, as shown in Figure 4,

has two degrees of freedom at steady state ( $N_u = 2$ ). (From a control point of view the column has  $N_m = 5$  degrees of freedom, but two degrees of freedom are needed to stabilize the reboiler and condenser holdups, which have no steady-state effect, and one degree of freedom is used to control the pressure at its given value). The two remaining degrees of freedom, e.g. selected to be the reflux flow  $L$  and the distillate flow  $D$ ,

$$u = \begin{bmatrix} L \\ D \end{bmatrix}$$

(note that this is not a unique choice) may be used to optimize the operation of the plant. However the question is: How should the optimal solution be implemented, that is, which two variables  $c$  should be specified and controlled during operation?

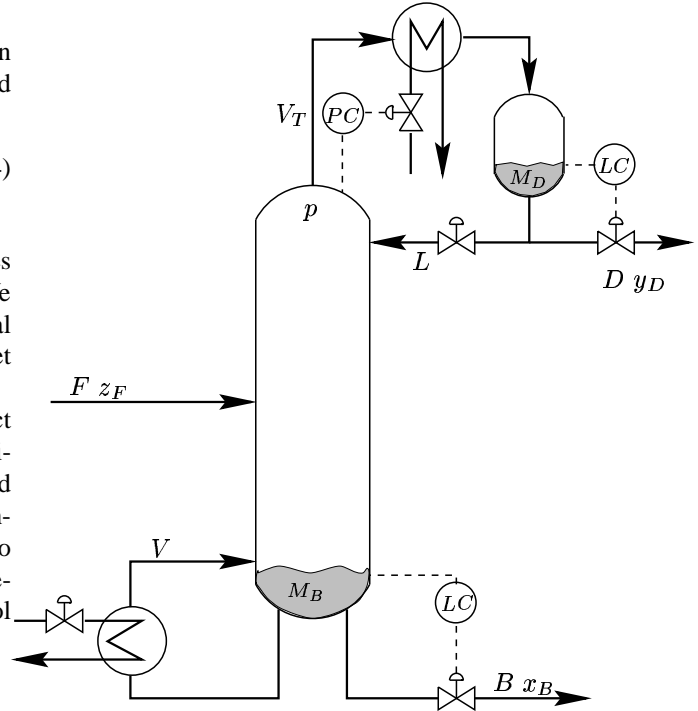


Fig. 4. Typical distillation column controlled with the LV-configuration

**Step 2: Cost function and constraints** Ideally, the optimal operation of the column should follow from considering the overall plant economics. However, to be able to analyze the column separately, we introduces prices for all streams entering and exiting the column and consider the following profit function  $P$  which should be maximized (i.e.  $J = -P$ )

$$P = p_D D + p_B B - p_F F - p_V V \quad (5)$$

i.e.  $J = -P$ . We use the following prices [\$/kmol]

$$p_D = 20, \quad p_B = 10 - 20x_B, \quad p_F = 10, \quad p_V = 0.1$$

The price  $p_V = 0.1$  [\$/kmol] on boilup includes the costs for heating and cooling which both increase proportionally with the boilup  $V$ . The price for the

feed is  $P_F = 10$  [\$/kmol], but its value has no significance on the optimal operation when the feed rate is given. The price for the distillate product is 20 [\$/kmol], and its purity specification is

$$x_D \geq 0.995$$

There is no purity specification on the bottoms product, but we note that its price is reduced in proportion to the amount of light component (because the unnecessary reprocessing of light component reduces the overall capacity of the plant).

With a nominal feedrate  $F = 1$  [kmol/min], the profit value of the column is of the order 4 [\$/min], and we would like to find a controlled variable which results in a loss  $L$  less than 0.04 for each disturbance ( $L = 0.04$  \$/min gives a yearly loss of about \$20000).

**Step 3: Disturbances** We consider five disturbances:

- $d_1$ : increase feedrate  $F$  from 1 to 1.3 kmol/min.
- $d_2$ : decrease feed composition  $z_F$  from 0.65 to 0.5
- $d_3$ : increase feed composition  $z_F$  from 0.65 to 0.75
- $d_4$ : decrease feed liquid fraction  $q_F$  from 1.0 (pure liquid) to 0.5 (50% vaporized)
- $d_5$ : increase the purity of distillate product  $x_D$  from 0.995 (its desired value) to 0.996

The latter is an implementation error (safety margin) for  $x_D$ . In addition, we will consider implementation errors for the other selected controlled variable (see below).

**Step 4: Optimization** In Table 1 we give the optimal operating point for the five disturbances; larger feedrate ( $F = 1.3$ ), less and more light component in the feed ( $z_F = 0.5$  and  $z_F = 0.65$ ), a partly vaporized feed ( $q_F = 0.5$ ), and a purer distillate product ( $y_D = 0.98$ ). In addition, we have considered the effect of a higher price for the distillate product ( $p_D = 30$ ) and a five times higher energy price ( $p_V = 0.5$ ).

As expected, the optimal value of all the variables listed in the table ( $x_D, x_B, D/F, L/F, V/F, P/F$ ) are insensitive to the feedrate, since the columns has no capacity constraints, and the efficiency is assumed independent of the column load.

We find that, except in the last case with a much higher energy price, the optimal bottom composition stays fairly constant around  $x_B = 0.04$ . This indicates that a good strategy for implementation may be to control  $x_B$  at a constant value. On the other hand, the value of  $D/F$  varies considerably, so we expect this to be a poor choice for the controlled variable.

**Step 5: Candidate controlled variables** It is clear that one of the controlled variables should be the distillate composition,  $x_D$ . This follows since the optimal solution is always obtained when the product

purity specification for the most expensive product is “active”, i.e. in our case when  $x_D = 0.995$ . We are then left with one “unconstrained” degree of freedom which we want to specify by keeping a controlled variable at a constant value.

We found above that the optimal value of the bottom composition is relatively insensitive to disturbances and other changes, and this indicates that  $x_B$  is a good candidate controlled variable. However, there at least two practical problems associated with this choice. First, on-line composition measurements are often unreliable and expensive. Second, dynamic performance may be poor because it is generally difficult to control both product compositions (“two-point control”) due to strong interactions (e.g. SM and Shinsky). Thus, if possible, we would like to control some other variable.

In the following the following alternative controlled variables are considered (in addition to  $x_D$  which should always be kept constant at its lower limit of 0.995):

$$c_1 = x_B; c_2 = L/F; c_3 = V/F; c_4 = L/D; c_5 = D/F$$

We consider implementation errors of about 20% in all variables, including  $x_D$ .

**Step 6: Evaluation of loss** In Table 2 we show for  $F = 1$  [kmol/min] the loss  $L = P_{opt} - P$  [\$/min] when each of the five candidate controlled variables are kept constant at their nominally optimal values. Recall that we would like the loss to be less than 0.04 [\$/min] for each disturbance.

We have the following comments to the results for the loss given in Table 2:

- (1) As expected, we find that the loss is small when we keep  $x_B$  constant.
- (2) Somewhat surprisingly, for disturbances in feed composition it is even better to keep  $L/F$  or  $V/F$  constant
- (3) Not surprisingly, keeping  $D/F$  (or  $D$ ) constant is not an acceptable policy, e.g., operation is infeasible when  $z_F$  is reduced from 0.65 to 0.5.
- (4) All alternatives are insensitive to disturbances in feed enthalpy ( $q_F$ ).
- (5)  $L/D$  is not a good controlled variable, primarily because its optimal value is rather sensitive to feed composition changes.
- (6) For a implementation error (overpurification) in  $x_D$  where  $x_D$  is 0.996 rather than 0.995 all the alternatives give a loss of about 0.09. For a larger overpurification where  $x_D = 0.998$  (not shown in Table) the loss ranges from 0.43 ( $x_D$  constant) to 0.79 ( $L/D$  constant). The loss is rather large, so we conclude that we should try to control  $x_D$  close to its specification.
- (7) For reflux  $L$  and boilup  $V$  one should include “feedforward” action from  $F$ , and keep  $L/F$  and

$V/F$  constant (e.g., the loss is 0.514 if we keep  $L$  constant and  $F$  increases by 30%).

- (8) Using  $L/F$  or  $V/F$  as controlled variables is rather sensitive to implementation errors.
- (9) Other controlled variables have also been considered (not shown in Table). For example, a constant composition (temperature) on stage 19 (towards the bottom),  $x_{19} = 0.20$ , gives a loss of 0.064 when  $z_F$  is reduced to 0.5, but otherwise the losses are similar to those with  $x_B$  constant.
- (10) We have not computed the effect of changes in prices in the Table, because these do not effect column behavior, so all alternatives behave the same (with the same loss). Thus, if there are price changes, then one must recompute new optimal values for the variables.

**Step 6: Selection of controlled variables** The following three candidate sets of controlled variables yield the lowest losses

$$c_1 = \begin{bmatrix} x_B \\ x_D \end{bmatrix}; \quad c_2 = \begin{bmatrix} L/F \\ x_D \end{bmatrix}; \quad c_3 = \begin{bmatrix} V/F \\ x_D \end{bmatrix};$$

The “two-point” control structure  $c_1$  where both compositions are controlled, is known to result in a difficult control problem due to strong interaction, and we may not be able to keep the compositions at their specifications. The loss will then be larger than indicated, and it is probably better to keep  $L/F$  or  $V/F$  constant.

It is probably easier to keep  $L/F$  than  $V/F$  constant (less implementation error). On the other hand, we have already noted that it is important to control  $x_D$  close to its specification, and this is probably more easily done using reflux  $L$ . In conclusion, we have:

*Proposed control system.*

- $V$  is used<sup>1</sup> to keep  $x_D = 0.995$ .
- $L/F = 15.07$  is kept constant.

*Alternative control system.*

- $L$  is used to keep  $x_D = 0.995$ .
- $V/F = 15.70$  is kept constant.

If it turns out to be difficult to keep  $L/F$  (or  $V/F$ ) constant, then we may instead manipulate  $L$  (or  $V$ ) to keep a temperature towards the bottom of the column constant.

In summary, the distillation case study shows the importance of selecting the right controlled variables when implementing the optimal solution, and how the column may limit the maximum throughput of the plant. The analysis was mostly based on economic considerations (loss), but the bottom composition  $x_B$  was excluded as a controlled variable based on other considerations, namely the cost of measurement and controllability.

We note that the implementation error was not important in this case study, but we stress that it should be included in the analysis. For example, the implementation error is the main reason why we rarely select temperatures near the column ends as controlled variables (because the measurement error is too large compared to its sensitivity), but instead control a temperature away from the column end.

	$x_D$	$x_B$	$D/F$	$L/F$	$V/F$	$L/D$	$P/F$
Nominal	0.995	0.040	0.639	15.065	15.704	23.57	4.528
$F = 1.3$	0.995	0.040	0.639	15.065	15.704	23.57	4.528
$z_F = 0.5$	0.995	0.032	0.486	15.202	15.525	31.28	2.978
$z_F = 0.75$	0.995	0.050	0.741	14.543	15.284	19.62	5.620
$q_F = 0.5$	0.995	0.040	0.639	15.133	15.272	23.68	4.571
$x_D = 0.996$	0.996	0.042	0.637	15.594	16.232	24.47	4.443
$p_D = 30$	0.995	0.035	0.641	15.714	16.355	24.51	
$p_V = 0.5$	0.995	0.138	0.597	11.026	11.623	18.47	

Nominal values:  $F = 1$ ,  $z_F = 0.65$ ,  $q_F = 1.0$ ,  $p_D = 20$ ,  $p_V = 0.1$

Table 1. Optimal operating point

	$x_B =$	$D/F =$	$L/F =$	$V/F =$	$L/D =$
	0.04	0.639	15.065	15.704	23.57
Nominal	0	0	0	0	0
$F = 1.3$	0	0	0	0	0
$z_F = 0.5$	0.023	inf.	0.000	0.001	1.096
$z_F = 0.75$	0.019	2.530	0.006	0.004	0.129
$q_F = 0.5$	0.000	0.000	0.001	0.003	0.000
$x_D = 0.996$	0.086	0.089	0.091	0.091	0.093
20%impl.error	0.012	inf.	0.119	0.127	0.130

inf. denotes infeasible operation

Table 2. Loss for distillation case study.

## 5. REFERENCES

- Findeisen, W, F.N. Bailey, M. Brdys, K. Malinowski, P. Tatjewski and A. Wozniak (1980). *Control and coordination in Hierarchical Systems*. John Wiley & sons.
- Foss, C.S. (1973). Critique of chemical process control theory. *AIChE Journal* **19**(2), 209–214.
- Morari, M. (1982). Integrated plant control: A solution at hand or a research topic for the next decade?. In: *CPC-II*. pp. 467–495.
- Morari, M., G. Stephanopoulos and Y. Arkun (1980). Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the classification of the control task. Analysis of the optimizing control structures.. *AIChE Journal* pp. 220–232.
- Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control*. John Wiley & Sons.
- Skogestad, S. and M. Morari (1988). Understanding the dynamic behavior of distillation columns. *Ind. Eng. Chem. Res.* **27**(10), 1848–1862.

<sup>1</sup> There are other possible choices for controlling  $x_D$ , e.g. we could use the distillate flow  $D$ . However,  $V$  has a more direct effect.