Composition Control of Distillation Columns Using Multiple Temperature Measurements

Thor Mejdell

Sigurd Skogestad

Chemical Engineering
Norwegian Institute of Technology (NTH)
N-7034 Trondheim, Norway

12-16 November 1990

AIChE 1990 Annual Meeting in Chicago

Paper 239

Abstract

Temperatures and flows are often used as secondary measurements to estimate the product compositions (outputs) in distillation columns. The problem is characterized by strong collinearity (correlation) between the temperature measurements, and often between the effects of the inputs on the outputs. In a linear study three different estimator methods, the Kalman-Bucy Filter, Brosilow's inferential estimator, and Principal Component Regression (PCR) are tested for performance with mu-analysis. It is found that use of input flow measurement has a damaging effect on the estimator performance for this ill-conditioned plant (with high RGA-elements). This is the main reason why the Brosilow inferential estimator is found to perform poorly. Somewhat surprisingly, it is found that the static PCR-estimator performs well compared with the dynamic Kalman filter. The reason is that the temperatures and compostions have very similar dynamic responses. Contrary to some claims in the literature, it is found that the performance of the estimate, even when used for feedback control, generally is improved by adding temperature measurements. For high purity distillation columns and other plants with large elements in the (appropriately scaled) gain matrix, the use of input measuremets are not recommended in the present of input disturbances.

1 Introduction

This paper addresses the estimation of process outputs based on multiple secondary measurements. The application chosen here is the use of temperature and flow measurements to estimate the product compositions in a distillation column. This is an interesting application which features: i) a large number of strongly coupled measurements, and ii) a large number of disturbances and inputs with similar effects on the outputs.

The use of temperature measurements for feedback control of distillation columns is quite extensively discussed in the chemical engineering literature (eg., Nisenfeld and Seeman, 1981, p. 85-95). Temperatures are usually not used because they are of interest themselves, but as inexpensive and reliable indicators of composition. One problem is that temperature is a true indicator of the tray composition only if the mixture is binary and at constant pressure. Furthermore, even at steady state the correlation between the composition on a tray inside the column and the product composition at the end is not unique; it changes depending on the feed composition and the other product composition. These problems may be partly overcome by using several temperature measurements.

Measurement selection. Many columns have temperature sensors located at about every fifth tray in the column, that is, a typical column may have 5-10 temperature measurements. In industry all these measurements are rarely used. Rather, each composition measurement is replaced by a single temperature measurement and used for single-loop feedback control. The main problem is then to find a suitable location for this temperature. According to Nisenfeld and Seeman (1981) the most important issues are, i) that the temperature should be sensitive to changes in the composition, and ii) that the correlation between temperature and composition should be insensitive to disturbances in feed composition and in flows. Since the products are often very pure the first criteria favours placing the temperature sensor away from the products. The second criteria favours placing the sensors close to the product. However, in this paper, measurement location is not an important issue. The reason is that we use several (typical five or more) temperature measurements and then estimate the product compositions. In this case the exact location is far less important than in cases where single temperature measurements are used.

Problem definition. The objective is to obtain the best estimate $\hat{y}$ of the outputs (product compositions) using all available information, $\tilde{\theta}$. In terms of deviation variables

the linear estimator may be written

$$\hat{y}(s) = K(s)\tilde{\theta}(s) \tag{1}$$

This estimate should be obtained based on a description of the process (nominal model and expected uncertainty), the expected noise and disturbances, and a more precise definition of what we mean by "best". In the general case $\tilde{\theta}$ should include all measured dependent variables (primary measurements, y, and secondary measurements, θ), and all known independent variables (manipulated inputs, u, and measured disturbances, d). In this paper we usually have $\tilde{\theta} = \theta$, that is, the estimate is based on only secondary measurements (temperatures). The reason is that we assume no primary measurements, no measured disturbances, and we shall show for our case that the additional information contained in u is of limited value.

In this paper we consider three different approaches to the estimation problem: i) The Kalman-Bucy Filter, ii) Brosilow's Inferential Control Method, and iii) Principal Component Regression (PCR). In the last two cases we shall base the analysis on the steady-state, and use a constant gain matrix K.

Use of separate estimator. An estimator-based control scheme for the distillation column is shown in Fig.1. Note that we are implicitly assuming that the controller should be separated into two parts: one estimator which condenses all the measurements into a few estimated outputs, and a "small" (in terms of number of inputs) controller which uses these estimates for feedback control (Fig. 2). The motivation for doing this is reliability, design simplicity and robustness. In general, this solution is suboptimal compared to using one big controller which directly uses all available measurements. The reason is of course that some information is lost when the original measurements are condensed into the fewer estimated variables. In some cases it may be shown that no information is lost and this is then referred to a separation principle. In particular, this may be the case if all the states of the system are estimated since they contain all information about the system at any given point in time. However, in this paper we shall not use all states for feedback control and therefore the separation principle does not apply. And as we in our case are estimating the actual controlled outputs, we may postulate that the performance loss caused by the separation is not a major problem.

Use all available measurements ? The statement in the problem definition above that the best estimate should be based on *all* available measurements is not as obvious as one should think. Actually, a large number of authors (eg. Joseph and Brosilow, 1978, Morari

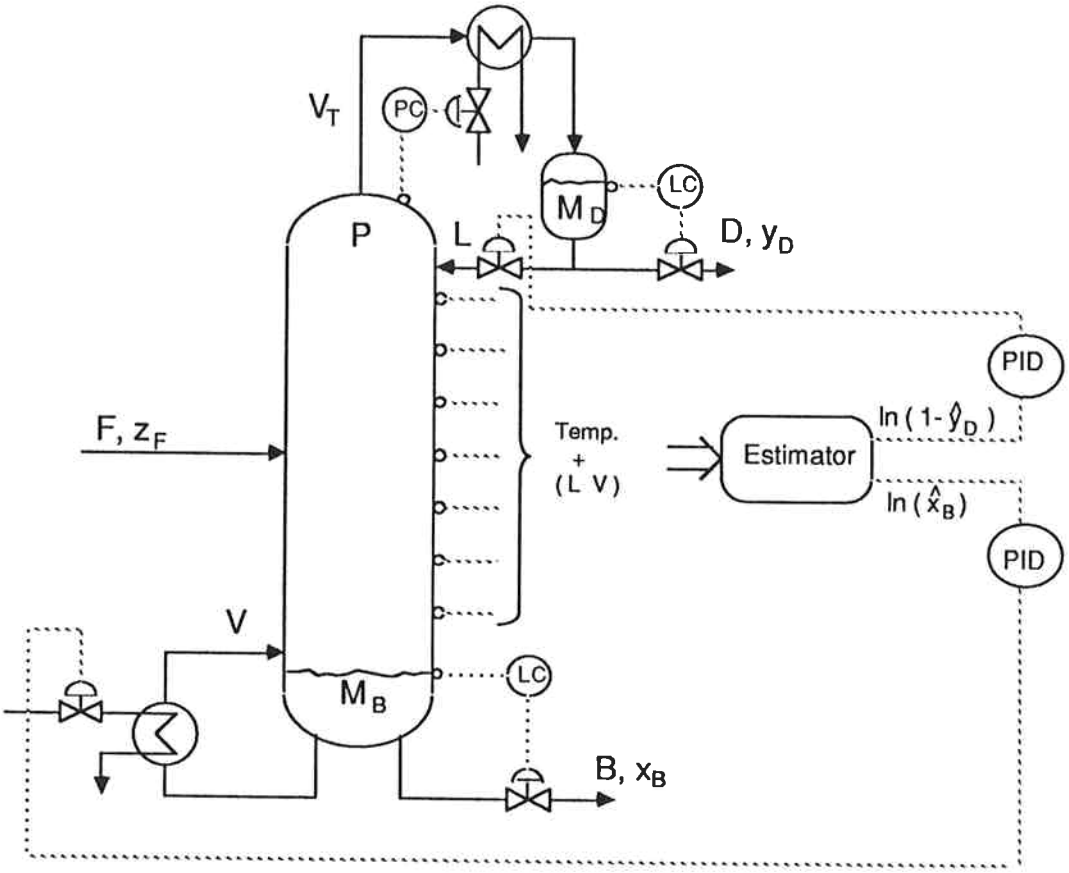

Figure 1: Control scheme based on LV configuration.

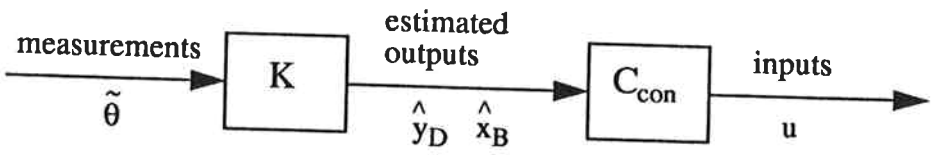

Figure 2: Controller block $C = C_{reg}C_{est}$ split in separate blocks for estimation and control.

and Stephanoplous, 1980, Patke et al., 1982, Yu and Luyben, 1986, Moore et al., 1987, Keller and Bonvin 1987) have suggested that one should only use a few of the *temperature* measurements to avoid the poorly conditioned problem of obtaining information from the strongly correlated temperatures. For example, our example column has 41 temperature measurements. That is, we need to determine 41 parameters in K for each output if all temperatures are used. However, the temperatures are of course strongly coupled and the 41 parameters must also be strongly coupled. Furthermore, our distillation column with two components, two products and constant pressure has only three degrees of freedom at steady state (eg., z_F, y_D, x_B). This implies that, at least for the linear case with small perturbations from the nominal operating point, we may determine at most three of these 41 parameters independently (irrespective of how the temperatures are coupled).

Latent variables. The above discussion points out the need for a robust way of obtaining the matrix K which avoids this overparameterization. Intuitively, this may be done by smoothening the available data, and obtaining a smaller number, k, of "latent variables", t, which are less coupled and contain most of the original information. These are subsequently used for estimation. In the linear case the latent variables may be written $t = P_t\theta$, where P_t is the projection matrix. The estimator then becomes $\hat{y} = K_t t$ where K_t is a "small" matrix with k parameters for each output (typically $k = 3$ in our examples), and the overparametrization in the regression step is avoided. The simplest "method", but certainly not the optimal one, is to delete measurements θ, and use, for example, only three temperatures as latent variables. This approach is implicit in some of the papers mentioned above. In Brosilows method estimated (inferred) disturbances are used as latent variables. In the PCR and PLS methods a few linear combinations of the secondary measurements are used as latent variables. These linear combinations are those which are found to be most sensitive based on the calibration set. In the Kalman estimator the states may be regarded as latent variables, although they are not independent as they are coupled through the model. Also, their number is often not less than the measurements.

Kalman estimator. The Kalman-Bucy filter (Kalman and Bucy, 1961) arises from the traditional "optimal" approach of modelling disturbances and noise as stochastic processes and minimizing a quadratic error function. This estimator contains a full dynamic model of the plant, and the states are updated using constant gain feedback from the measurements. Somewhat surprisingly, there are very few reports on the use of model-based Kalman filters for composition estimation in distillation columns. Apart from its

complexity the main disadvantage with the Kalman estimator is that model uncertainty is not included, and that it is difficult *a priori* to find the weights for the disturbance and noise.

Brosilow estimator. In process control, Weber and Brosilow (1972) proposed to use secondary measurements to estimate disturbances. Their justification is that measurement noise is usually less important in process control applications, and that the output variations are mainly caused by disturbances, which tend to vary slowly compared to the process dynamics. In Brosilows inferential controller, the disturbances are then assumed to be constant in the future, and the disturbance estimates are used in a sort of feed-forward scheme to counteract their expected effect on the outputs. We shall only use the estimator part of Brosilows scheme (Joseph et al., 1976) and not the feedback part. Brosilows scheme has a strong intuitive appeal and seems to have found some use in industry. However, because of measurement noise, model error and poor numerical properties caused by collinearity, we will show that the use of inferred disturbances as latent variables may not work well for ill-conditioned plants.

PCR estimator. A more direct approach is to derive a direct relationship between θ and y using a static regression estimator. The approach taken here is inspired by recent efforts by analytical chemists in their "multivariate calibration problem", e.g. Wold et al. (1987). This 'soft modelling' approach has an intuitive appeal to engineers as one seems to skip the modelling step: One does not have to obtain an explicit ('hard') model of how the independent variables affect θ and y (although typical variations should be included in the calibration set). Rather, one seeks a direct correlation between the available measurements (θ) and the variables to be estimated (y). However, fundamental knowledge may not be easily included in the estimator.

Analysis of estimators. The estimators are compared on a rigorous basis, considering both the estimation error $y-\hat{y}$ ("open loop" when the estimates are not used for feedback) and the control error $y - y_s$ ("closed-loop" when the estimates are used for feedback). Input uncertainty, disturbances and noise are explicitly included in the analysis using the structured singular value, μ, of Doyle (1982).

Ill-conditioned plants. The distillation column used in this paper is an example of an ill-conditioned plant. Here the plant gain is strongly dependent on the input direction, or equivalently the plant has a large condition number, $\gamma(G)$. At each frequency

$$\gamma(G) = \sigma_1(G)/\sigma_r(G) \tag{2}$$

α	N	N_F	z_F	y_D	x_B	D/F	L/F
1.5	40	21	0.50	0.99	0.01	0.500	2.706

- Feed is liquid.
- Constant molar flows.
- Ideal VLE using Raoults law.
- Constant pressure 1 atm.
- Holdup on each tray; $M_i/F = 0.5$ min

Parameters A,B and C in Antoine equation:
ln p(mm HG) = A - B/(T(K) + C)

Light component	Heavy component
15.8366, 2697.55, -48.78	15.4311, 2697.55, -48.78

Table 1: Data for distillation column example.

Here σ_1 is the largest singular value, and σ_r is the r'th (the smallest) singular value, where r is the rank of G. $\sigma_1(G)$ is a measure of the magnitude of the elements in the matrix. The smallest additive perturbation matrix which may make G loose rank has magnitude $\sigma_r(G)$. Thus, the condition number, $\gamma(G)$ gives the relative magnitude of the additive error allowed to avoid singularity (loosing rank). Consequently, matrices with a large condition number are very sensitive to numerical round-off errors (eg., Weber and Brosilow, 1972). For square matrices the relative gain array (RGA) may be used as an alternative measure. It is defined at each frequency as RGA $= G \times (G^{-1})^T$, where $\times$ denotes element-by-element multiplication. The magnitude of the RGA–elements is closely related to the optimal condition number ($\gamma(G)$ minimized with respect to input and output scaling) (Skogestad and Morari, 1987). Skogestad and Morari (1987) have shown that the RGA is also a very good indicator of how sensitive a plant's feedback control performance is to input gain uncertainty.

Example column. As an example column we use column A studied by Skogestad and Morari, 1988. The column separates a binary mixture with relative volatility 1.5, and has 40 theoretical stages, including the reboiler, plus a total condenser. Column data are given in Table 1. The liquid holdups are assumed constant, that is, the flow dynamics are neglected. This gives rise to a 41th order linear model in terms of the mole fraction of the light component on each tray. The two dominant time constants of the column are 194

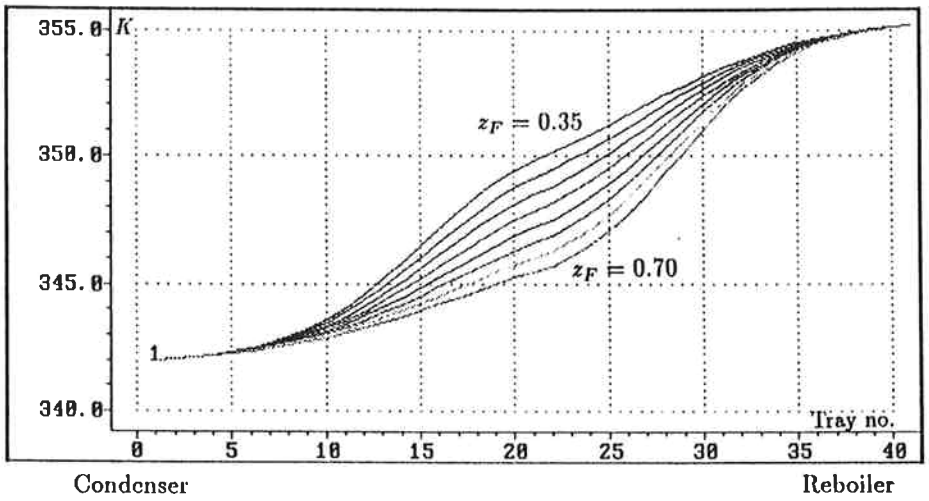

Figure 3: Temperature profiles for different feed compositions when y_D and x_B are held constant.

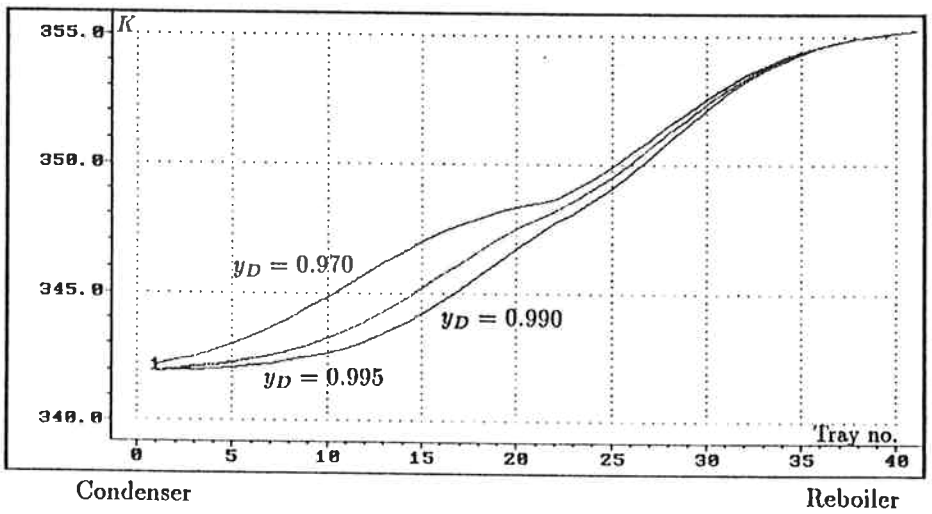

Figure 4: Temperature profiles for different top product compositions when z_F and x_B are held constant.

min and 15 min. The difference in boiling points of the two pure components is 13 °C. In Figure 3 and in Figure 4 typical temperature profiles for the column are displayed. We note that variations in temperature are small towards the ends of the columns, and that changes in feed composition have a large effect on the temperatures inside the column even though the product compositions are constant.

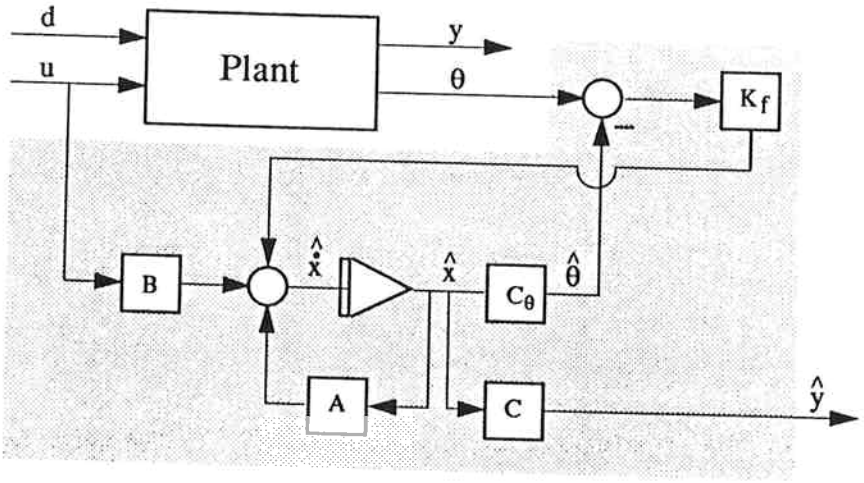

Figure 5: Block diagram of the Kalman Estimator.

2 Estimation Methods

2.1 Kalman filter.

In this scheme a dynamic state space model is used in parallel with the process, and the deviation between the outputs from the process and the model is used as feedback to the model through a filter gain K_f (Fig. 5). The linear state space model for the process is

$$\dot{x} = Ax + Bu + Ev \tag{3}$$

$$y = Cx \tag{4}$$

$$\theta = C_\theta x + w \tag{5}$$

Here x is the state vector, u the manipulated inputs, y the primary outputs to be estimated, θ the secondary measurements, v the process noise (disturbances), and w the measurement noise. v and w are assumed to be white noise processes with covariance matrices $\mathcal{V}$ and $\mathcal{W}$.

Minimizing the expected variance of $\theta - \hat{\theta}$ yields the estimated states

$$\dot{\hat{x}} = A\hat{x} + Bu + K_f(\theta - C_\theta\hat{x}) \tag{6}$$

$$= (A - K_f C_\theta)\hat{x} + Bu + K_f\theta \tag{7}$$

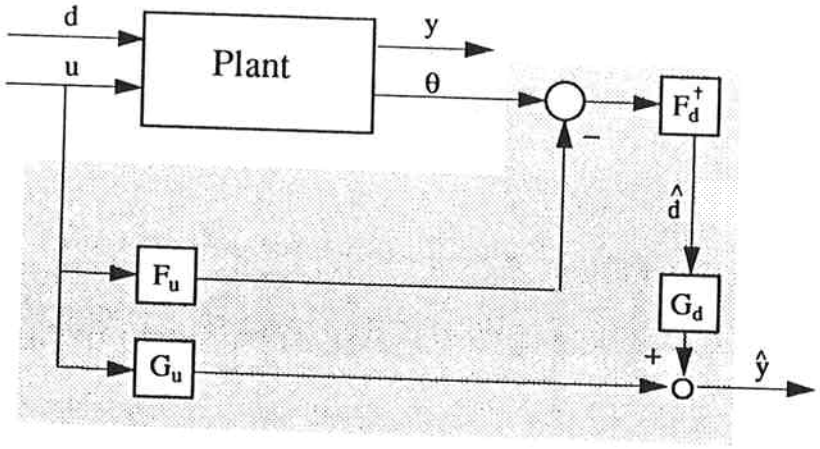

Figure 6: Block diagram of Brosilow Estimator.

where filter gain K_f is

$$K_f = \mathcal{X} C_\theta^T \mathcal{W}^{-1} \tag{8}$$

Here $\mathcal{X}$, the covariance matrix of $\hat{x}$, is found from the matrix Riccati equation

$$\dot{\mathcal{X}} = A\mathcal{X} + \mathcal{X}A^T - \mathcal{X}C_\theta^T \mathcal{W}^{-1} C_\theta \mathcal{X} + E\mathcal{V}E^T \tag{9}$$

We use constant filter gains which give $\dot{\mathcal{X}} = 0$, and Eq. (9) is reduced to an algebraic equation. The overall Kalman estimator then becomes

$$\hat{y}(s) = C(sI - A + K_f C_\theta)^{-1}(K_f \theta(s) + Bu(s)) \tag{10}$$

2.2 Brosilow estimator.

The following linear steady-state model of the column in terms of deviation variables is used

$$y = G_d d + G_u u \tag{11}$$

$$\theta = F_d d + F_u u \tag{12}$$

Here d denotes the disturbances. The matrices above are of course related to those used in the state space description in the Kalman filter. For example, $G_u = -CA^{-1}B$ and for the case $v = d$ we have $F_d = -C_\theta A^{-1}E$. Using (12) the estimated disturbances become

$$\hat{d} = F_d^\dagger(\theta - F_u u) \tag{13}$$

where the pseudoinverse $F_d^\dagger$ is the optimal inverse in the general least square sense. For the special case of more θ's than d's and independent d's (Weber and Brosilow, 1972)

$$F_d^\dagger = (F_d^T F_d)^{-1} F_d^T \tag{14}$$

For the special case of more d's than θ's and independent θ's (Joseph et al., 1976, Joseph and Brosilow, 1978)

$$F_d^\dagger = F_d^T (F_d F_d^T)^{-1} \tag{15}$$

In the general case the pseudoinverse is obtained from a SVD of F_d by deleting directions with singular values equal to zero (eg., see Strang, 1980, p. 142). Combining (11) and (13) yields the Brosliow estimator (see Fig. 6)

$$\hat{y} = K_B \theta + (G_u - K_B F_u) u \tag{16}$$

where

$$K_B = G_d F_d^\dagger \tag{17}$$

This static Brosilow estimator may be made equivalent to the Kalman filter at steady-state only if non-stationary noise is allowed for the disturbances v (Morari and Stephanopoulos, 1980).

2.3 PCR estimator.

We want to estimate p outputs (y) from q known variables (θ). The problem is then to obtain the matrix K in

$$\hat{y} = K\theta \tag{18}$$

To this end obtain n "calibration" sets of corresponding values of y and θ, and place these as *rows* in the matrices $Y^{n \times p}$ and $\Theta^{n \times q}$, respectively. [1] If the estimator was perfect we would have

$$Y = \Theta K^T \tag{19}$$

The ordinary least square solution for K is:

$$K_{LS} = Y^T \Theta [\Theta^T \Theta]^{-1} \tag{20}$$

which is the "regression estimator" used by Joseph et al. (1976). The $q \times q$ matrix $\Theta^T \Theta$ is n times the covariance matrix of the calibration measurements θ. This matrix is singular

[1] It might seem more reasonable to place y and θ as columns in the matrices, but we shall here use the standard notation in statistics.

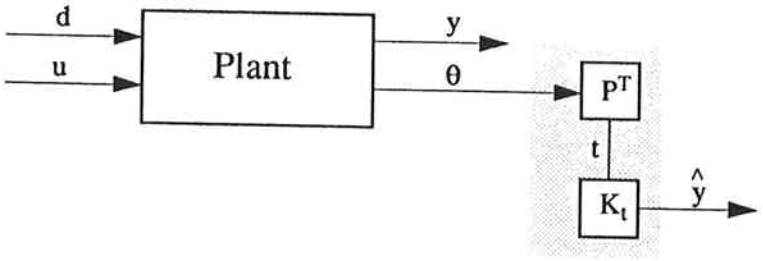

Figure 7: Block diagram of a PCR Estimator that does not use input information.

if $n < q$, that is, if we have too few calibration sets. It is also singular if strong collinearity in the temperatures exists. This will usually be the case in a column with measurements located close to each other.

To avoid these difficulties the general pseudo inverse in terms of the SVD is used, and directions corresponding to small singular values (principal components) are deleted. Using standard notation from the statistics literature, the SVD of Θ is written

$$\Theta = t_1 p_1^T + t_2 p_2^T + \cdots + t_m p_m^T \tag{21}$$

where $m \leq min(n, q)$ is the rank of Θ. Here p_1 is the eigenvector corresponding to the largest eigenvalue of $\Theta^T\Theta$, (the square of the largest singular value of Θ), and p_2 is the eigenvector corresponding to the second largest eigenvalue, and so on. The loading vectors (p's) give the directions of the principal components, while the scores (t's) give the magnitude. If all m terms in (21) are retained we obtain the generalized pseudoinverse. However, in PCR we select only those first k principal components that can be distinguished from the measurement noise. Let the matrices $P^{q\times k}$ and $T^{n\times k}$ include only these k most important directions. Define the new latent variables as $t = P^T\theta$. Note that $P^T = P^{-1}$ since P is orthonormal. The least square solution to $y = K_t t$ becomes $K_t = Y^T T[T^T T]^{-1}$, and the overall estimator gain matrix becomes (see Fig 7)

$$K_{PCR} = Y^T T[T^T T]^{-1} P^T \tag{22}$$

Gain		$\mathcal{V}$				
		L	V	F	z_F	
K1	diag{	200	200	0.01	0.01	}
K2	diag{	0.10	0.10	0.01	0.01	}
K3	diag{	0.01	0.01	0.01	0.01	}
K4	diag{	0.0	0.0	0.01	0.01	}

Table 2: Process disturbance covariance matrix of Kalman filter gains. In all cases $\mathcal{W} = I.I\Delta\mathcal{I}$

In the general case θ may be replaced by $\tilde{\theta}$ which includes also the inputs and measured disturbances.

3 Estimators for the example column.

In this section we describe how the different estimators were obtained for the example column with 41 stages.

3.1 Kalman filter gains.

The covariance matrix of the measurement noise $\mathcal{W}$ was set to $0.04I$, where I is the identity matrix. This corresponds to $0.2\ ^\circ C$ noise on each temperature. The process noise is here defined as $v^T = [L, V, F, z_F]$ (reflux, boilup, feedrate and feed composition). Its covariance matrix, $\mathcal{V}$, was assumed diagonal and was varied in order to tune the filter. Four different values of the variance on L and V were selected (Table 2) and the corresponding filter gain matrices are denoted K1 to K4. The assumption of white noise process disturbances is somewhat unrealistic in a distillation column, but the estimator is not expected to be very sensitive to this assumption.

3.2 Brosilow estimator.

With $d^T = [z_F, F]$ and $u^T = [L, V]$, the matrices F_d,F_u,G_d and G_u in equations (11) and (12) were found by linearizing the model at the nominal operating point. The estimator was obtained using Eq. (16).

A modified estimator $K_{B_{mod}}$ was formed by *not* using information about the manipulated inputs u, and instead using L, V and z_f as the disturbances d' to be inferred. The

z_f	y_d	x_b	z_f	y_d	x_b
0.4000	0.9810	0.0190	0.4000	0.9810	0.0010
0.4000	0.9990	0.0190	0.4000	0.9990	0.0010
0.6000	0.9810	0.0190	0.6000	0.9810	0.0010
0.6000	0.9990	0.0190	0.6000	0.9990	0.0010
0.4500	0.9855	0.0145	0.4500	0.9855	0.0055
0.4500	0.9945	0.0145	0.4500	0.9945	0.0055
0.5500	0.9855	0.0145	0.5500	0.9855	0.0055
0.5500	0.9945	0.0145	0.5500	0.9945	0.0055

Table 3: Data to simulate stationary temperature profile. $q_F = 1.0$, P = 1.0 atm.

estimator then becomes $\hat{y} = K_{Bmod}\theta$ where

$$K_{Bmod} = G'(F'^T F')^{-1} F'^T \tag{23}$$

and F' and G' are the process matrices formed by these three variables. In the linear case with no errors in the matrices G' and F', this estimator is identical to the PCR-estimator. This is also clear if we compare (22) and (23) and imagine using changes in L, V and z_F to generate the calibration sets. In both cases we obtain the least square estimate, and if we disregard numerical problems it does not matter which latent variables we use.

3.3 PCR estimator.

In this paper the calibration sets are obtained from a *linear* steady state column model. A factorial design method was used to select 16 different runs around the operating point (Table 3). When stated random noise of magnitude 0.1 °C was added on all temperatures in the calibration sets, but the default is no noise. The specified variables were chosen as the outputs y_D and x_B and the feed composition z_F. Note that the column conditions are independent of the load (increasing all flows proportionally), and it is not necessary to simulate different feed rates. The temperature data were reduced to the desired number of principal components and K_{PCR} was computed from (22).

Strictly speaking, with a linear model we need only three runs (in addition to the nominal steady state) to generate the data, but we used 16 runs to better study the effect of measurement noise and to get better statistical information.

It is important to note that with this approach we may freely vary the *outputs*, y_D and x_B, and are thus able to span all directions in the output space. This is different from the Brosilow approach, which is based on an open-loop model in terms of the *inputs*

Case.	Location (tray no.)
41	at every tray
5a	1,12,21,30,41
5b	10,15,22,29,33
3a	2,22,41
3b	6,22,36
3c	10,22,33
3d	10,17,33
2a	1,41
2b	6,36
2c	9,33
2d	10,30

Table 4: Location of temperature measurements.

(L, V, F, z_F), and where the output space will not be properly spanned for ill-conditioned plants with strongly coupled outputs y.

3.4 Number of measurements and their locations

The estimation methods above were applied to different locations and numbers of temperature measurements. The various cases are summarized in Table 4. Here tray no. 41 denotes the reboiler, no. 21 the feedtray, and no. 1 the condenser.

4 Analysis of the Estimators.

The objective is to evaluate the different estimation methods described above. In this section we define our criteria for the evaluation.

4.1 Evaluation criteria

- *Open-loop evaluation (OL).* One obvious criteria for evaluating the different estimators is their ability to follow the true composition value. The error e_1 in fig. 8 is the difference between the real (y) and the estimated output $(\hat{y})$. The column is assumed to operate under feedback, since this is more close to a real situation than a pure open loop test. The term "open loop" is still used since the controller uses the actual y, that is, there is no feedback from the estimate $\hat{y}$. We use single-loop PID controllers since this is the most common choice in practice. The tunings in

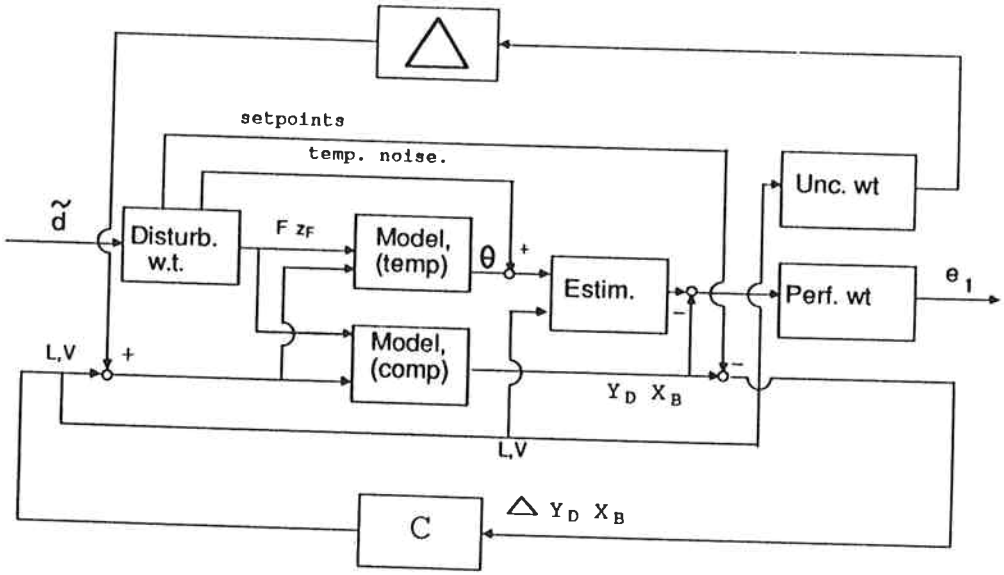

Figure 8: Block diagram for "open-loop" μ-test. We use $\Delta = 0$ (nominal performance) unless otherwise stated.

PID-Parameters

Loop	K_c	τ_i	τ_d
y_D	0.589	9.53	0.620
x_B	0.555	4.42	0.332

Table 5: PID-parameters for the distillation column example.

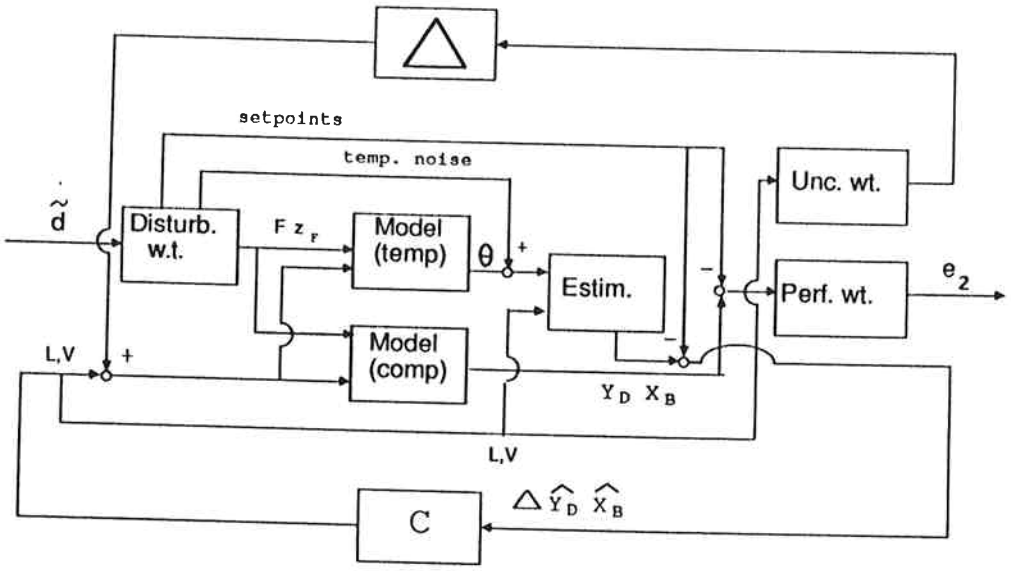

Figure 9: Block diagram for "closed-loop" test.

Table 5 yield optimal robust performance (minimize μ) when the estimate is exact. To make our results only weakly dependent on the controller used, we shall usually consider the *nominal performance* in this test, i.e., without any uncertainty. This makes the comparison independent of the robust stability requirement of the system which depends strongly on the controller.

- *Closed-loop evaluation (CL).* The main objective of the estimator is to replace the primary measurement of y, that is, use the estimate $\hat{y}$ for feedback control. The error, e_2, of interest to be minimized, is then the control error, i.e. the difference between y and y_{setp} (Fig. 9). We use the same controller as for the open-loop comparison, that is, a PID controller tuned optimally for perfect estimates. Using the same PID controller for all estimators will bias the comparison somewhat, as the optimal controller in each case will depend on the estimator used.

4.2 μ-analysis.

Our tool is the Structural Singular Value (μ) analysis (Doyle,1982). In this framework we rearrange our system to fit the general form shown in fig. 10. Here M denotes the generalized nominal plant including the plant and the weights, $\tilde{d}$ denotes external input

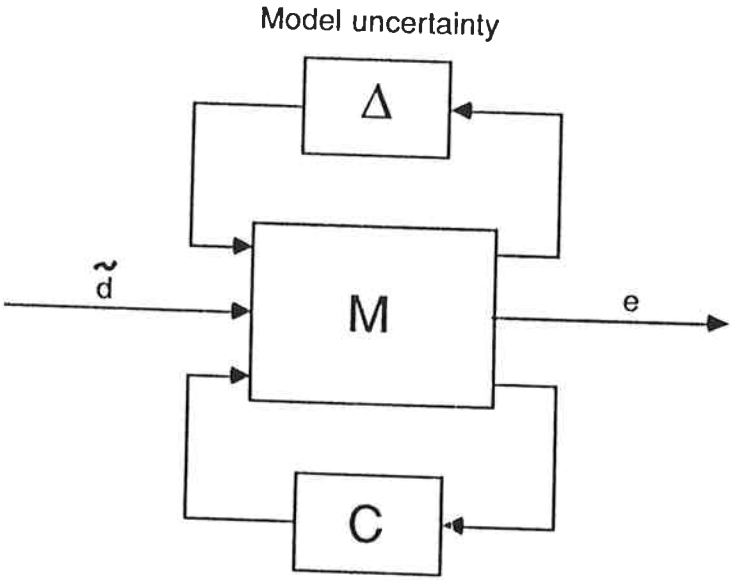

Figure 10: General structure for studying any linear control problem.

disturbances and setpoint changes, and e is the "error" we want to keep small. We have one Δ-block loop, which represent the model uncertainty, and one controller loop. In the μ analysis we evaluate the maximum amplification from $\tilde{d}$ to e at each frequency. Weights are used to scale the signals, $\tilde{d}$ and e, and the uncertainty Δ to be less than 1. These weights are discussed below. μ expresses the worst-case error at a given frequency, and the performance requirement for the error is satisfied if μ is less than one at all frequencies.

4.2.1 Uncertainty weights.

The most important source of uncertainty is assumed to be on the inputs (L and V). We shall use the same uncertainty weight as Skogestad and Morari (1988), which is given by

$$w_I(s) = 0.2\,\frac{5s+1}{0.5s+1} \tag{24}$$

The weight is shown graphically in fig 11a. In the low frequency range it allows for a 20% uncertainty in flow *changes* (L and V are deviation variables), due to the inaccuracy of valve settings. The uncertainty increases at higher frequencies, reaching a value of 100% at about $\omega = 1$ min^{-1}. The increase at high frequencies will allow for a time delay of about 1 min between L and V and the outputs y_D and x_B.

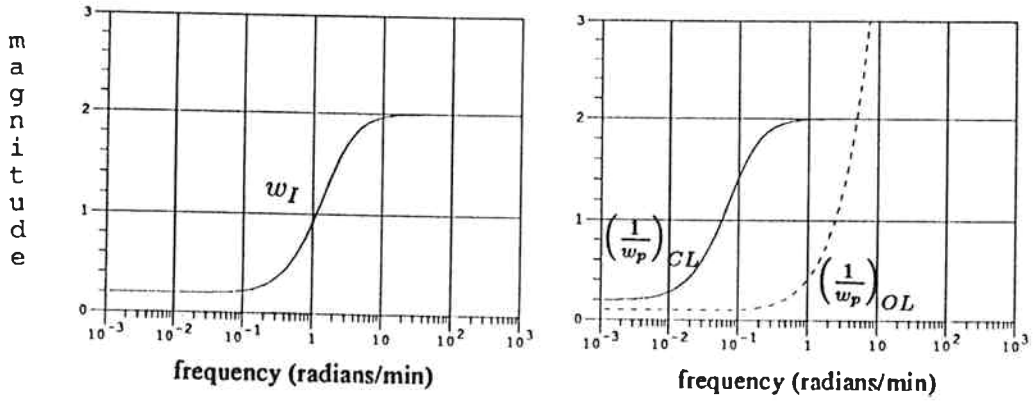

Figure 11: Weights for μ-analysis. a) Uncertainty weight. b) Inverse performance weights. Solid line: Closed loop, dotted line: Open loop.

4.2.2 Performance weights.

In the "Open Loop" test we use the following performance weight

$$w_p(s) = \frac{10}{4s+1} \tag{25}$$

which is shown in fig 11b. This weight requires less than 10% estimation error for $(1-\hat{y}_D)$ and $\hat{x}_B$ at steady-state ($\omega \leq 0.1$ min^{-1}). At higher frequencies the weight increases to one at $\omega = 2.5$ min^{-1}. This allows for an error greater than 100% at frequencies above 2.5 min^{-1}. In the "Closed Loop" test we chose a performance weight

$$w_p(s) = 5\frac{10s+1}{100s+1} \tag{26}$$

This implies that the deviation of $y - y_s$ should be within 20% at steady state, i.e, we tolerate a deviation of the product composition of about 0.2 mole%. Our feedback system should be effective up to about $\omega = 0.05$ min^{-1} and the amplification at high frequencies should never exceed 2. Except for the allowed steady state offset this weight is the same as the one used by Skogestad and Morari (1988).

4.2.3 Weights for external inputs.

The external inputs to the systems (the $\tilde{d}$'s in the block diagrams for mu-analysis) consist of setpoints, as well as ordinary disturbances and noise. They are normalized by specifying their maximum values at any frequency using weights. The maximum setpoint changes

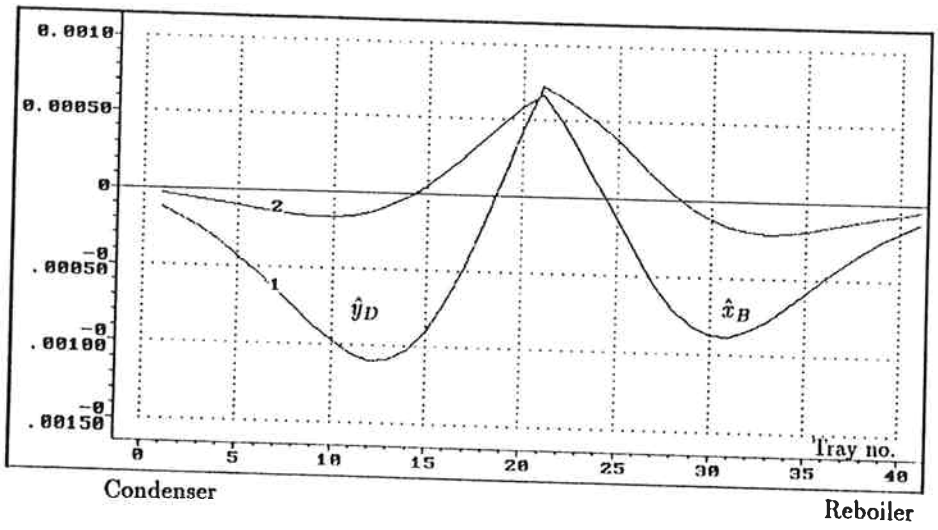

Figure 12: Elements in matrix K for estimation of 1) y_D, 2) x_B.

are set to 100% of x_B and $(1-y_D)$. Since the operating point is 0.01 and 0.99 this implies that the x_{Bset} may vary from 0 to 0.02, and y_{Dset} from 0.98 to 1.0. The disturbances in the feedrate F and the feed composition z_F are set to 20%, i.e. z_F may vary from 0.4 to 0.6 in mole fraction. Noise was generated by adding a constant vector of random values with normal distribution and a standard deviation of 0.2 $°C$ to all 41 temperatures. No noise is used in the mu-analysis unless otherwise stated.

5 Results.

5.1 Insights into the collinearity using PCR.

The elements in the matrix K_{PCR} for the case with 41 temperatures, are plotted in fig. 12. In fig. 13 the three largest loading vectors p are displayed. These show how the different measurements are summed up to make the latent variables (principal components). The first component is mainly due to changes in the external flows, D and B, and reflects moving the entire temperature profile up and down the column. The second component is due to changes in internal streams (with D and B constant), and reflects stretching or compressing the profile (changing the separation in the column). The third component is due to changes in the feed composition. From the figures we see that the temperatures near the product streams are weighted little compared to the ones towards the middle of the column. The reason is that the temperature variation is small at the ends, and the

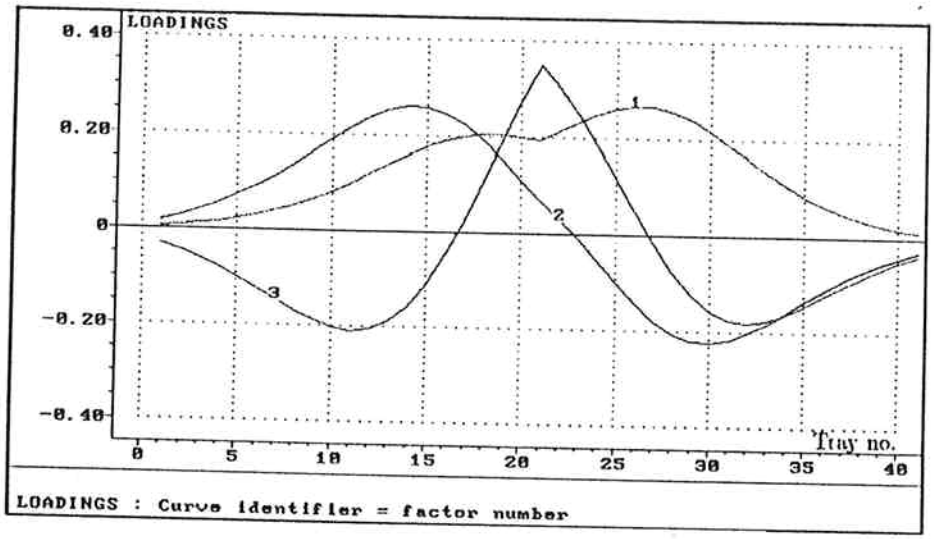

Figure 13: Loading plot of the first 3 principal components. Curve identifier: component number.

measurements are therefore much more sensitive to noise. Pressure variations were not included in this study, but the temperatures near the ends of the column would be useful to compensate for such variations. The forth vector is displayed in fig. 14. We see that this vector contains only numerical noise, and there are, as expected, only three different directions in the temperature space when pressure is kept constant. This is also confirmed by Figures 15 and 16. They show how the different principal components account for the total variance in the calibration set both in y-space and in θ-space.

5.2 Number of measurements and their location.

The μ-plots in Figure 17 for the PCR estimator shows the effect of using varying numbers of measurements. It demonstrates that adding temperature measurements improves the estimates and the control performance. The main difference is between two and three measurements. With less than three measurements all principal components in the temperature space can not be recovered, unless they are placed towards the ends where the dimension of temperatures shrinks to one (see Figures 3 and 4). A comparison of various locations of the two temperatures are shown in the μ-plots in Figure 18. Without noise the best location is of course at the ends (trays 1 and 41, Case 2a in Figure 18a) and a perfect estimate is obtained. With noise it is better to use measurements closer to the middle of the column where the temperature changes are much larger (trays 6 and 36, Case 2b in Figure 18b). The same conclusion applies to cases with three (Fig. 19)

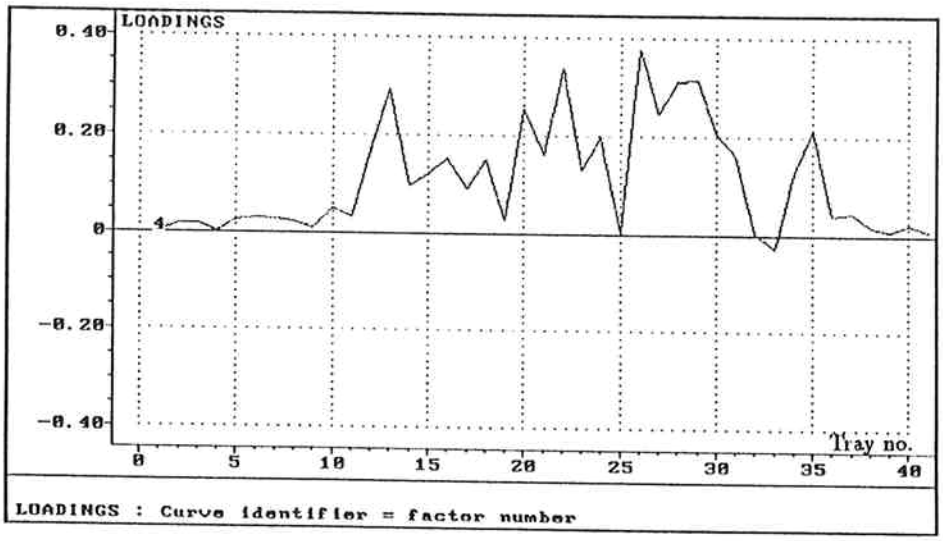

Figure 14: Loading plot for the fourth principal component.

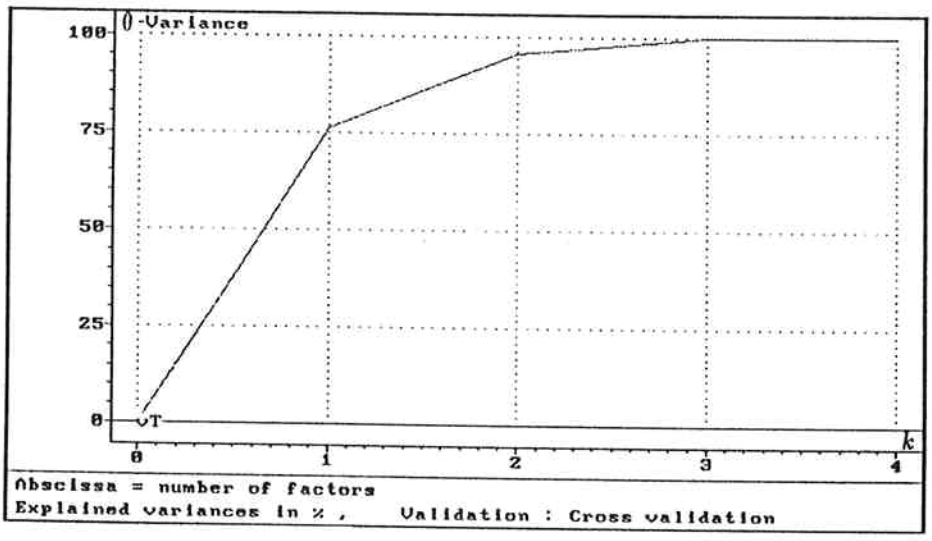

Figure 15: Explained Θ-variance (%). Abscissa: Number of principal components.

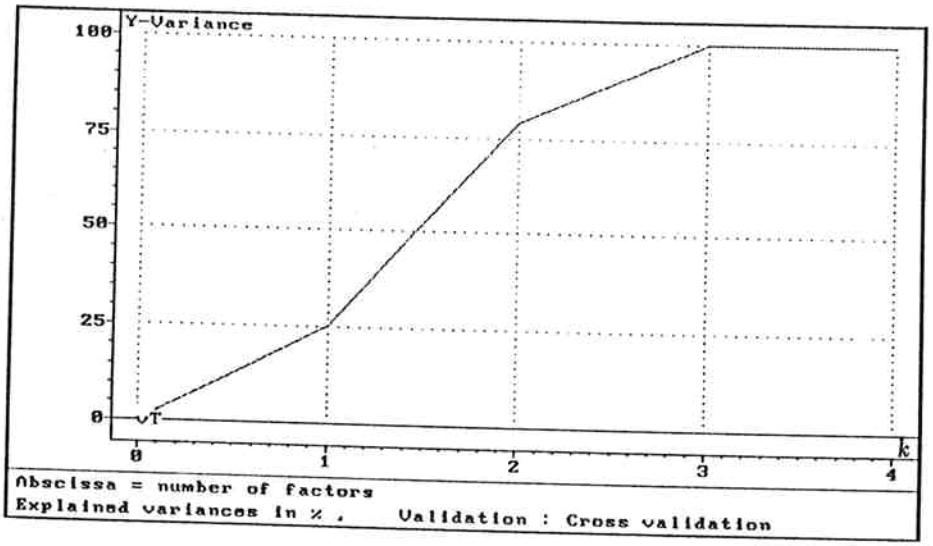

Figure 16: Explained Y-variance (%). Abscissa: Number of principal components.

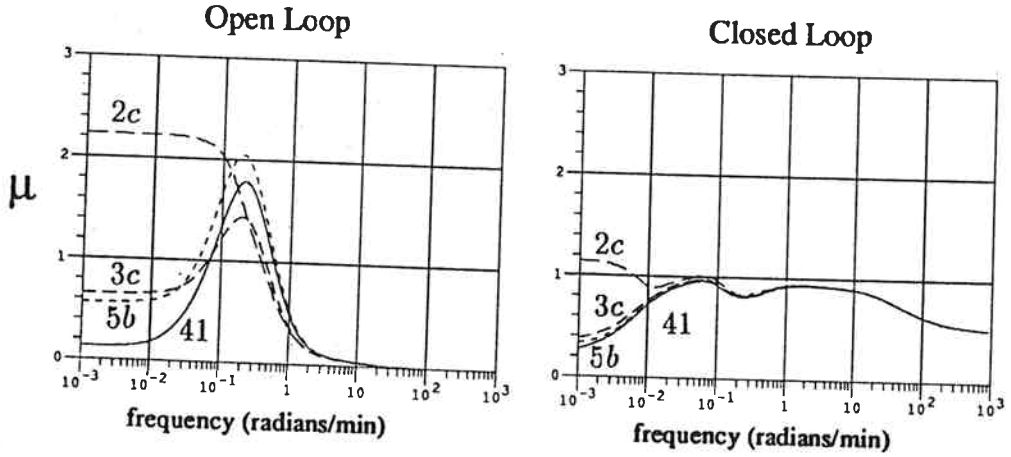

Figure 17: Effect on μ of number of temperatures for PCR-estimator. The temperatures in the calibration set are corrupted with 0.1 °C noise.

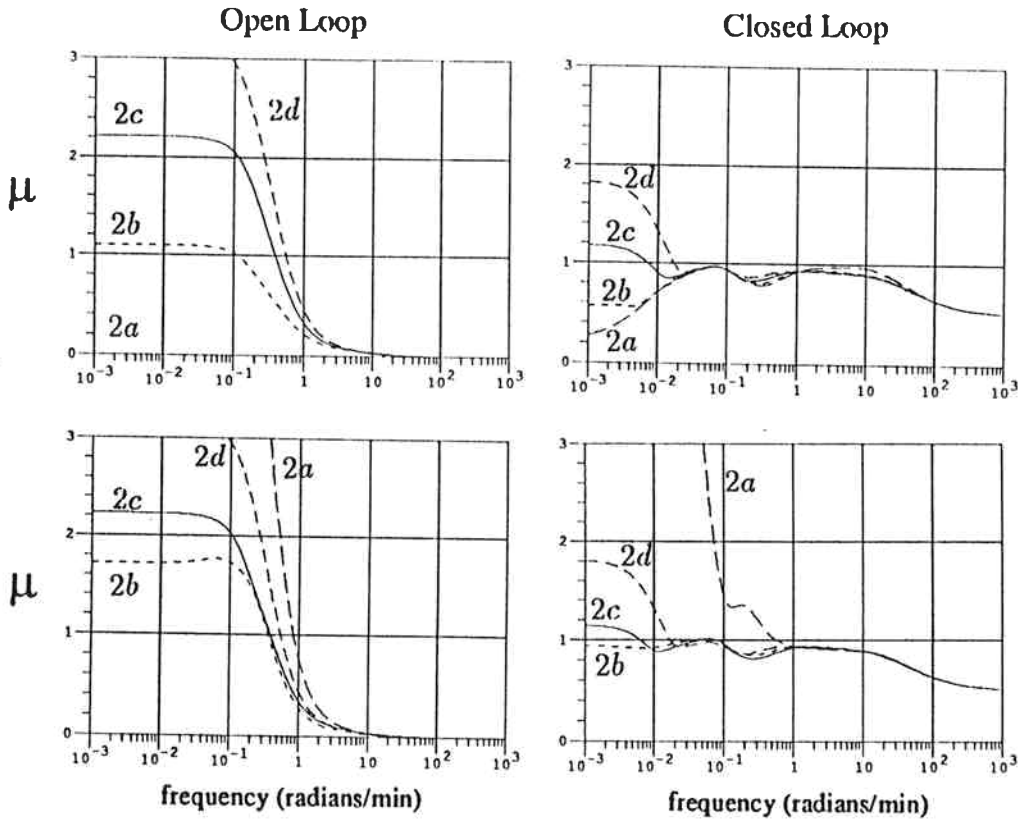

Figure 18: Effect on μ of noise and location for two measurements. a) PCR without noise in calibration set, b) PCR with 0.1 °C noise in calibration set.

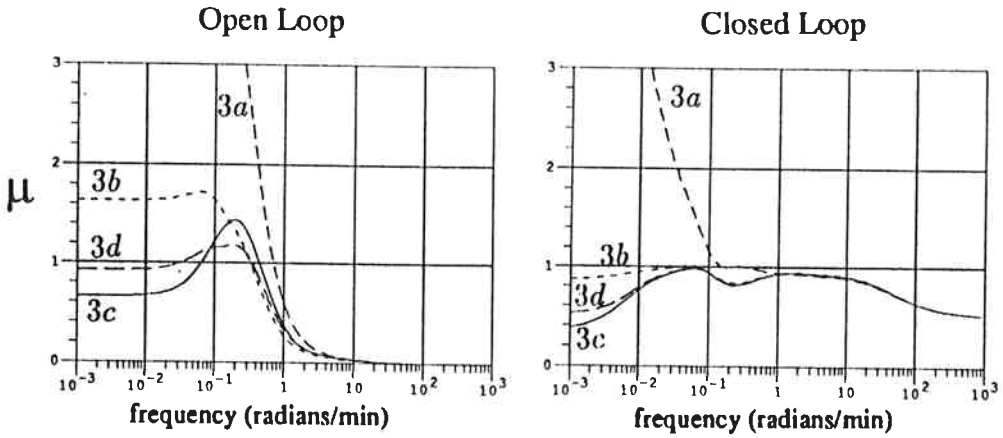

Figure 19: Effect on μ of location for three measurements. PCR with 0.1 °C noise in calibration set.

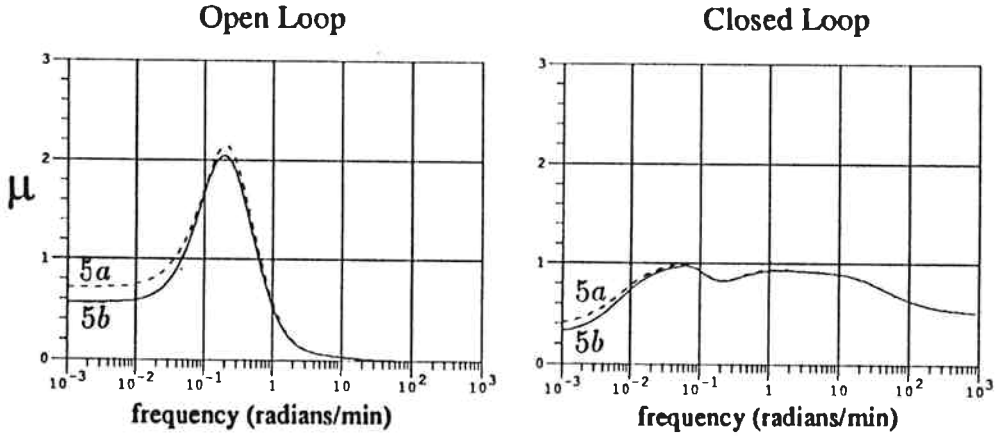

Figure 20: Effect on μ of location for five measurements. PCR with 0.1 °C noise in calibration set.

and five measurements (Fig. 20), but the location of measurements of course becomes less important as additional measurements are used, provided they are reasonably evenly spaced.

Also the Kalman filter is improved by adding measurements. This is illustrated by the μ-plots in Figure 21.

5.3 Comparison of Kalman filter and static PCR estimator.

In fig. 22 we compare the μ-plots of the Kalman and PCR estimators, using 41 temperatures. The first thing to note is how well the simple static estimator $\hat{y} = K_{PCR}\theta$

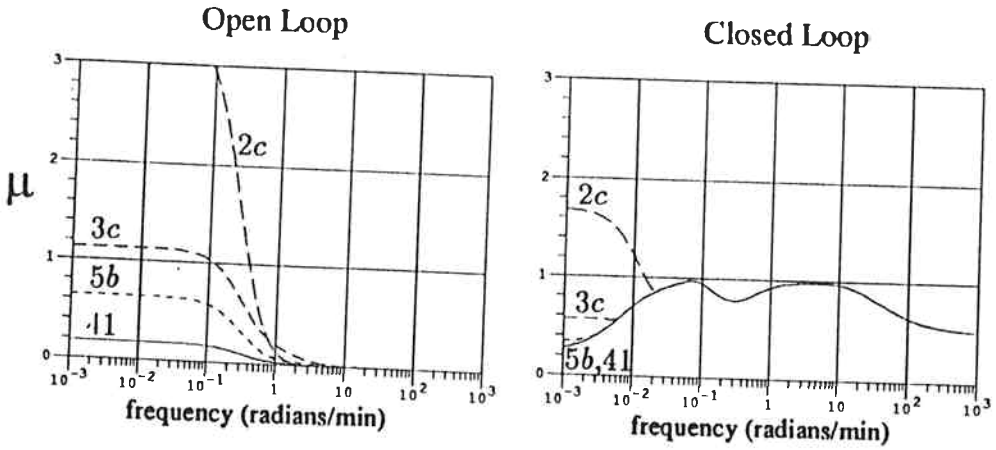

Figure 21: Effect on μ of number of measurements for Kalman Filter (K1). No Noise.

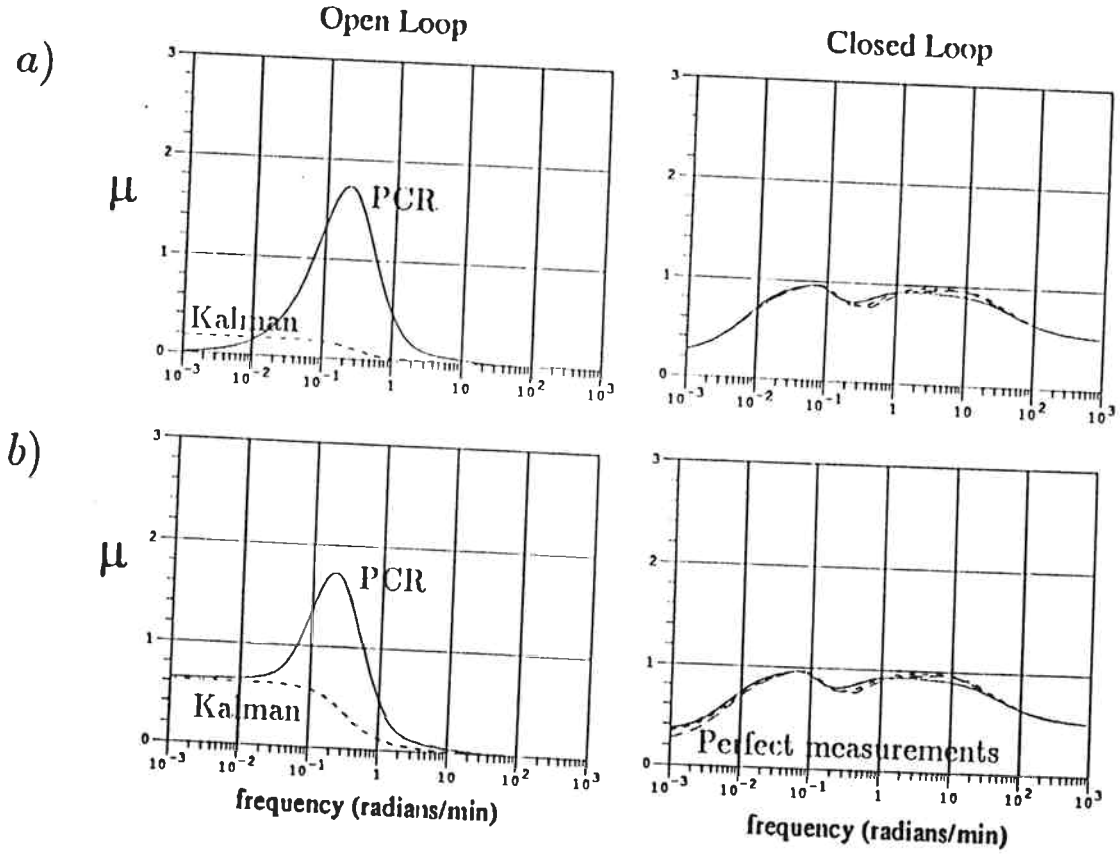

Figure 22: Comparison of Kalman (K1) and PCR estimator. a) without noise in μ-analysis, b) with noise.

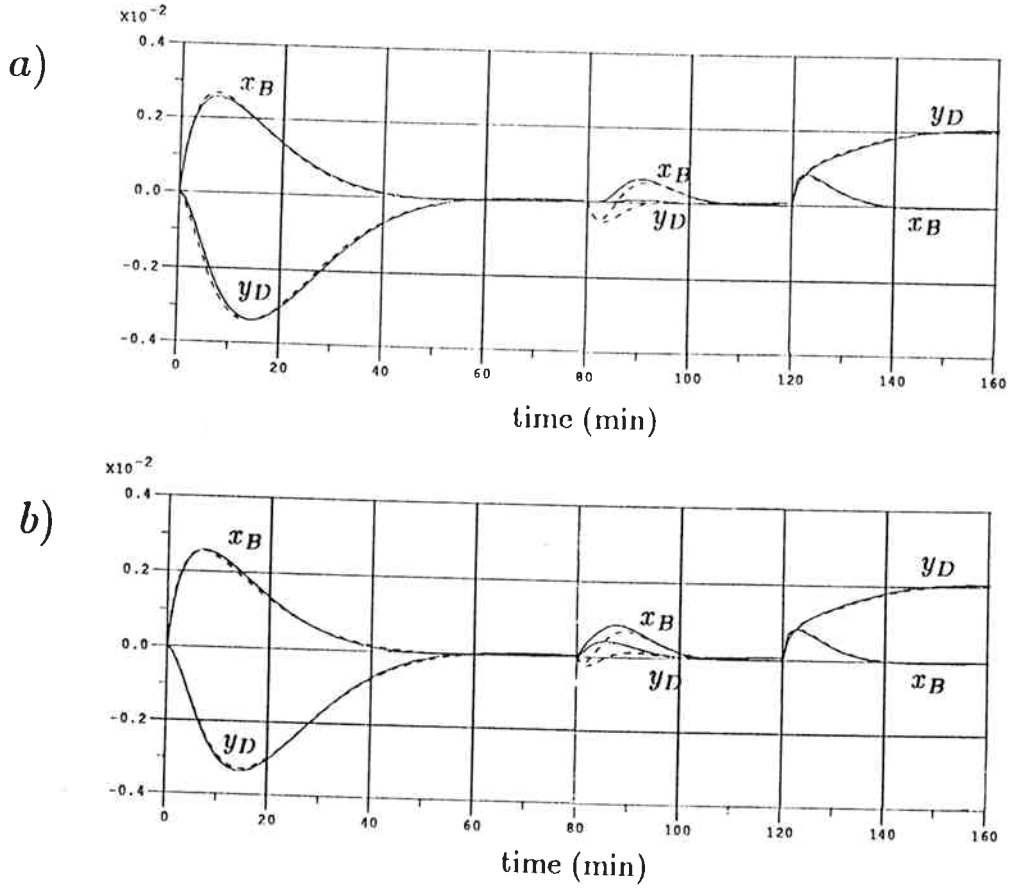

Figure 23: Comparison of output $y(t)$ (solid) and PCR-estimate $\hat{y}(t)$ (dotted line). Responses under feedback control are shown for a 20 % increase in feedrate at t=0, a 20 % increase in feed composition at t=80 min, and a setpoint change in y_D at t=120 min. a) y used for feedback control, b) $\hat{y}$ used for feed backcontrol.

performs. The main reason is that the dynamic responses of the temperatures θ and the compositions y are very similar. This will be the case for most distillation columns, at least for sections of the column, but may of course not be the case for other applications.

In the Open Loop analysis the Kalman filter is significantly better at higher frequencies. This is due to the dynamics included in this estimator. On the other hand, the "Closed Loop" test shows that the estimators will perform about equally well when used for feedback, and also as well as using perfect measurements. Actually, for some frequencies, the PCR estimator is even better than using perfect measurements. The reason is that the temperatures in the middle of the column generally change slightly faster than at the ends, and the steady state estimator will therefore have a small inherent "feedforward" effect. The simulation responses in Figure 23 confirm that the PCR-estimate is almost equal to the true value. One exception is for feed composition disturbances, where it shows a small inverse response.

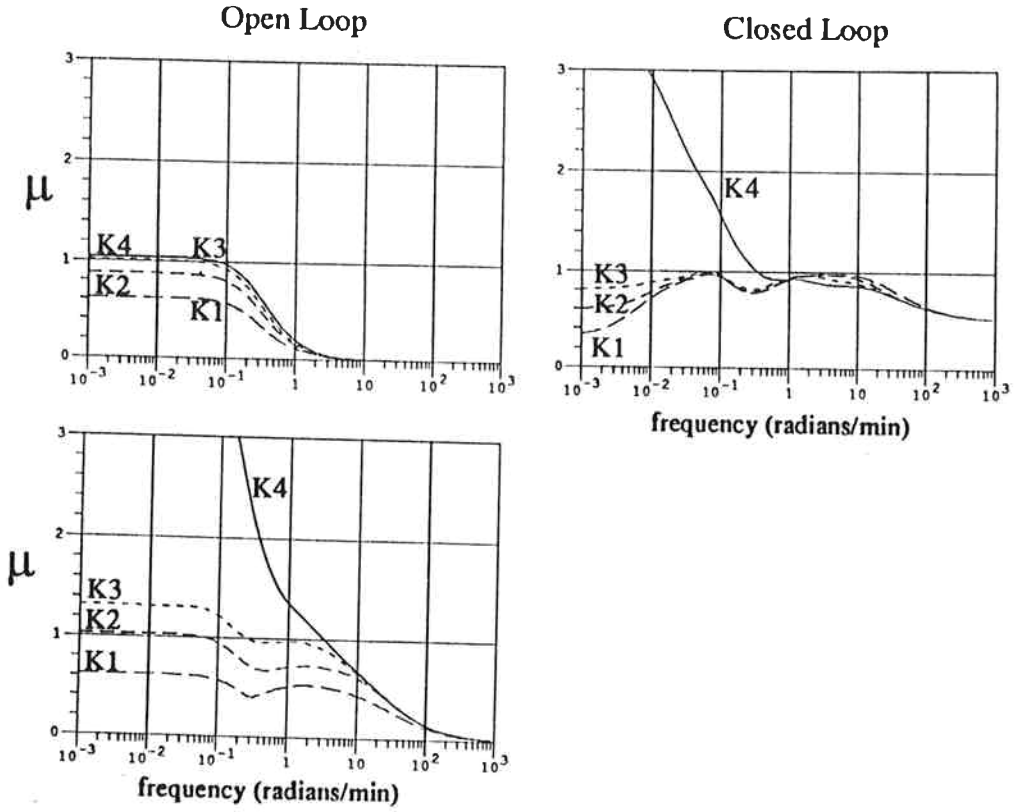

Figure 24: Different Kalman filter gains (Table 4.2). Upper left: Nominal estimation error. Lower left: Robust estimation error. Upper right: Robust control error. Solid line: K4. Short dotted line: K3. Medium dotted line: K2. Long dotted line: K1.

5.4 Different Kalman filters and use of inputs in estimator.

Figure 24 shows μ-plots for the Kalman filters obtained using the four different levels of process noise on L and V in Table 2. The best Kalman filter, K1, is the one that was compared with PCR above. The remarkable thing with this best estimator is the very large assumed variance on the inputs u (L and V). In effect, this variance is so large that the transfer function from u to $\hat{y}$ in Eq.(10) is approximately zero, that is, the estimator does not use the information about the input signals.

The worst Kalman filter, K4, assumes disturbances (noise) of magnitude 0.1 for F and z_F, but assumes no disturbances on the inputs. This estimator performs reasonably well in the μ-test when there is no uncertainty (upper left part in Figure 24). (But note that disturbances on the inputs are not included in the μ-analysis.) However, it is extremely poor when input uncertainty is added (lower left).

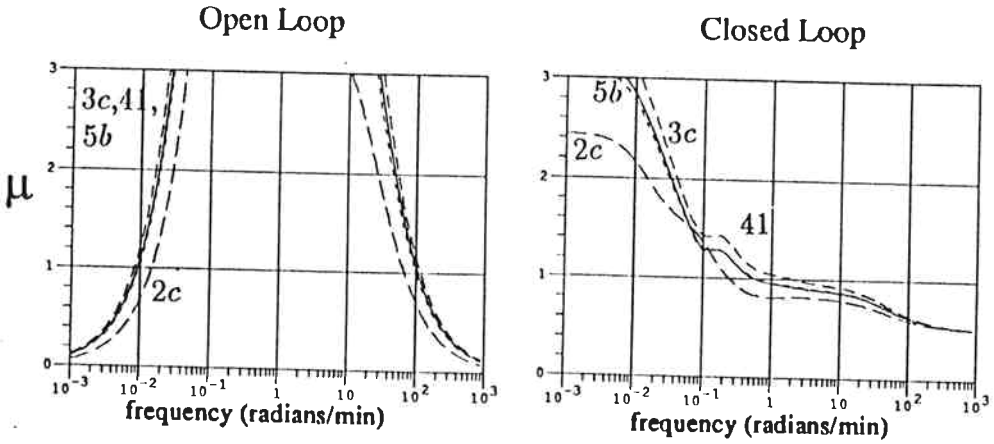

Figure 25: Effect of number of measurements for Brosilow Inferential Estimator. No noise.

The PCR estimator in this paper uses only temperatures, but we did also evaluate the effect of adding inputs. However, the improvement in estimator performance was very small even at steady state. Furthermore, the dynamic behaviour of the static estimator is much worse when inputs are used.

5.5 Brosilow estimator.

The Brosilow inferential estimator for the system with different numbers of measurements is shown in Figure 25. It clearly demonstrates that the estimator as originally proposed performs poorly, and its performance does not improve with increasing number of measurements. The "Open-loop" test shows that the estimator nominally works well at very low frequencies ($\omega < 0.001$ min^{-1}). The poor dynamic performance (intermediate frequencies) is due to the fact that the estimator uses the input signals u (L and V) as shown in Eq. (16). The dynamic behaviour of u and the compositions y are very different and using a constant matrix $G_u - K_B F_u$ does not work well. This problem could have been corrected using a low-pass filter on the inputs with a large time constant, e.g., 194 minutes (that is, add dynamics to G_u and F_u). However, even this estimator would not perform well in practice, as the "Closed-loop" test shows that the robust performance is poor even at low frequencies. This is due to the input uncertainty, that is, the actual values of L and V are different from what the estimator thinks they are.

We therefore conclude that using the measured input signals u (which are inaccurate) does not improve the estimate. A better approach then seems to be to regard the inputs L

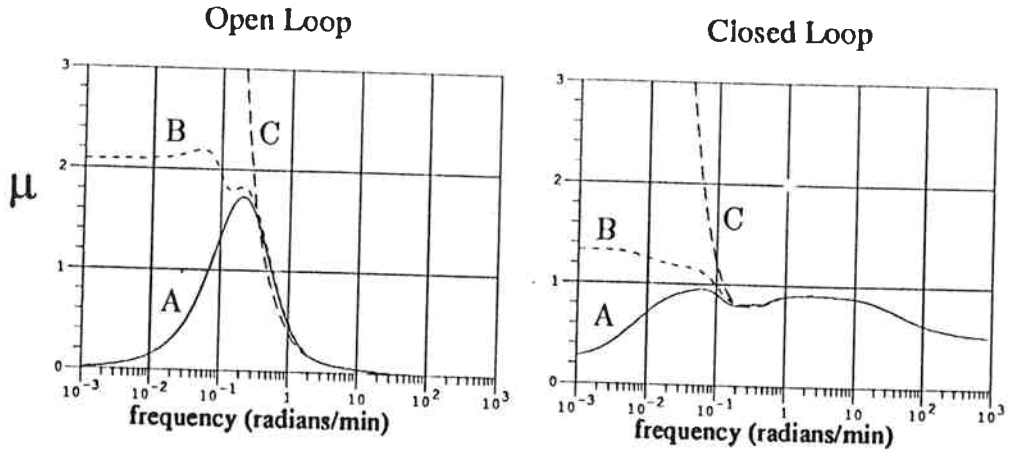

Figure 26: Modified Brosilow estimator based on temperatures only. 41 temperatures. A: Perfect model. B: Model rounded to 3 digits. C: Model when 1% random noise is added to process matrices.

and V as unknown disturbance together with z_f. This gives rise to the modified estimator $\hat{y} = K_{Bmod}d'$ where $d' = [L, V, z_F]^T$. This estimator performs much better as seen from curve A in Figure 26. The estimated values of the latent variables L, V and z_F may not be correct, but this error is not important as long as the estimate $\hat{y}$ is accurate. However, using L, V and z_F as latent variables has very poor numerical properties. For example, curve C in Figure 26 shows the drastic deterioration in performance caused by adding 1% random error to the elements of the matrices G' and F'.

6 Discussion

6.1 Kalman filter.

Model uncertainty is not included explicitly when obtaining the Kalman filter and it may require physically unrealistic values of the noise weights, V and W, in order to obtain the best estimator when uncertainty is included. This is illustrated by the large value needed for noise (disturbances) on the inputs in order to obtain the best Kalman filter, K1. Otherwise, the Kalman Filter performed well in the μ-tests and was undoubtedly the best estimator in the open loop μ-test. The main reason is its inherent dynamics, which can track the changes in the process tightly. Furthermore, because of the weights, it is flexible, and it may be tuned to perform well for ill-conditioned plants a well. As mentioned above this is done by adding (artificial) large noise (disturbances) on the inputs

to the process.

6.2 Brosilow estimator.

As discussed above the Brosilow Inferential estimator as originally proposed suffers from four main weaknesses:

W1. *For ill-conditioned plants (with large RGA-values) input uncertainty causes poor estimates when the estimator uses information about the manipulated inputs u.*

W2. *Even for plants which are not ill-conditioned, the dynamic behaviour of a static estimator which directly uses inputs is often poor. The reason is the dynamic "lag" which usually exists between the inputs u and the outputs y.*

W3. *It does not handle collinearity among the variables in an appropriate way.* If the number of disturbances are less than the number of measurements, like in our example column, the problem arises when there is collinearity among the disturbances. This makes the results sensitive to small numerical errors as shown above. On the other hand, if the number of disturbances is larger than the number of measurements, like in the work of Joseph and Brosilow (1978), the collinearity between temperatures creates problems. Instead of using only selected measurements as proposed by Joseph and Brosilow (1978), one should rather delete small *directions* in F_d using the singular value decomposition.

W4. *For ill-conditioned plants (with large condition numbers) the use of inputs and disturbances as latent variables is ill-conceived .*

Weaknesses W1 and W2 may be corrected using the "modified" Brosilow estimator, and also W3 may be corrected using an appropriate pseudoinverse of F_d. However, the use of secondary measurements to infer the disturbances and then estimate the primary measurements is the key idea in the Brosilow estimator, and W4 can not be corrected. To illustrate W4, consider an ill-conditioned plant, where we, in order to get a good estimate must require that:

1) The estimate of the disturbances and the inputs is very accurate (this implies that models F_u and F_d, which are used to infer the disturbances, also must be very accurate).

2) The model from disturbances and inputs to disturbances (G_d and G_u) is very accurate (it must capture the low-gain direction as well).

If the condition number of any one of these four matrices is large then the estimate may be sensitive to small numerical errors. The same applies to the modified Brosilow estimator if the matrices G' or F' are ill-conditioned. For our example column the condition numbers of G' are F' are 165 and 321. This explains the poor results in Fig.26. We want to stress that this sensitivity to errors in the matrix elements is different from the sensitivity to input uncertainty in W1, which is discussed in more detail below.

The estimation scheme of Brosilow is based on explaining the observations by estimating the inputs using a causal input-output model. This approach may be satisfactory in many cases, but not for ill-conditioned plants [2]. However, for such systems there may still be a rather simple direct relationship between various *dependent* variables, for example, between temperatures and composition in a distillation column, and a simple regression estimator may work well.

6.3 PCR estimator.

The PCR-estimator does not have the same weaknesses as the Brosilow estimator. First, the estimator used here does not use the input values, and does not suffer from uncertainty with respect to their exact value and poor dynamic performance. Secondly, and more important, its numerical properties are much better. The matrix to invert in PCR, the score matrix T in Eq. (22), is generally much better conditioned than F' used by the modified Brosilow estimator. For example, for our column the condition number of T is 4.7, whereas the condition number of F' is 321. To get a well-conditioned T one must ensure that excitations of the weak directions are included in the calibration set. To ensure such excitations, one should use data from the column with feedback (that is, with specified outputs), for example, by specifying the product compositions together with the feed composition in an factorial design like in Table 3. One should *not* use open loop data, like step responses etc., which will excite only the strong directions (The gain matrices in Brosilow's scheme will typically result from such excitations).

In an earlier study (Joseph et al., 1976) it was found that the Brosilow estimator performed better than a regression estimator. However, they used the simple least-square estimator in (20) which suffers from the same poor numerical properties as the Brosilow estimator.

[2] The extreme of an ill-conditioned plant is a chaotic system where it is impossible to back-calculate the inputs which have caused the observed outputs

Conceptually, it is simple to generalize the static PCR estimator to obtain a dynamic estimator. This may be done using the PCR method to derive an ARMA model relating time series data for θ and y.

Partial Least Square (PLS) estimator.

The PLS estimator is an alternative regression estimator, which also takes into account the directions in Y when finding the approximate pseudo inverse of Θ (Höskuldson 1988). In the PLS method this is done by considering the eigenvalues of $\Theta^T Y Y^T \Theta$ rather than of $\Theta^T \Theta$ used in PCR. This takes into account the directions in Θ which have the largest covariance with Y, and thus ensures that these directions are not deleted. For the linear distillation example studied in this paper the PCR and PLS methods gave almost identical results. However, when nonlinear data were used we found PLS to be somewhat better.

6.4 Use of inputs in estimator

.

The μ-results showed that when the inputs are used explicitly by the estimator, the Brosilow estimator and the Kalman filter (case K4) are very sensitive to input errors. By input error we mean the differences between the actual plant inputs, u, and the desired input, u_c, computed by the controller and which are used for estimation (see Figure 27). We consider two source of input error: 1) input disturbances, and 2) model uncertainty at the inputs. We then have

$$u = (I + \Delta_I)(u_c + u_d) \tag{27}$$

where u_d is the disturbance on the inputs, and Δ_I is the relative input uncertainty (see below). To understand the effect of the input error, consider the somewhat unlikely case when there are no secondary measurements, that is, $F_d(s) = F_u(s) = 0$. In this case both the Brosilow estimator and Kalman filter become

$$\hat{y} = G_u(s) u_c \tag{28}$$

Assume also that there are no other disturbances (this assumption may easily be relaxed). Then the actual plant output is

$$y = G_u u \tag{29}$$

and the estimation error becomes

$$e_1 = y - \hat{y} = G_u(u - u_c) \tag{30}$$

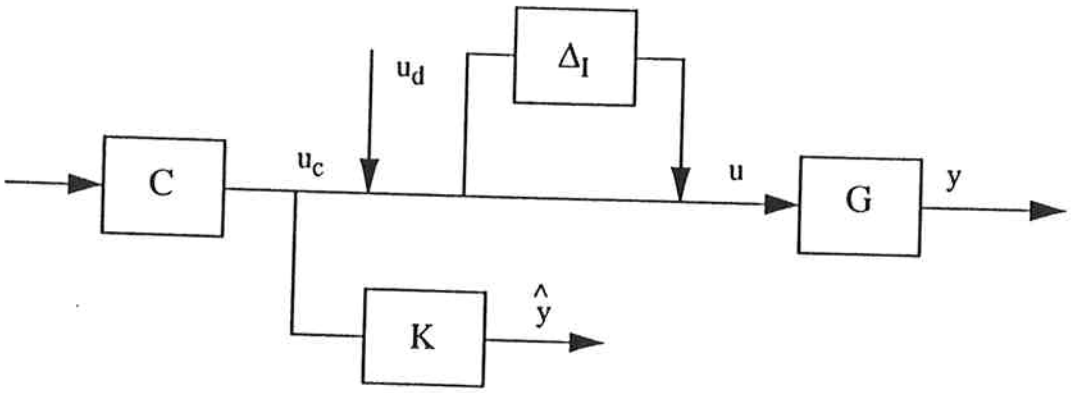

Figure 27: Actual input, u, may differ from value, u_c, used by the estimator because of 1) input disturbance u_d and 2) input uncertainty Δ_I.

We shall now consider separately the two sources of input error.

1. Input disturbances. In this case $u - u_c = u_d$ and the estimation error becomes

$$y - \hat{y} = G_u u_d \tag{31}$$

We note that the estimation error may be large even for small disturbances, u_d, if the elements in the matrix $G_u(j\omega)$ are large, that is, if $\bar{\sigma}(G_u(j\omega)$ is much larger than 1. This assumes that G_u has been scaled such that at any frequency expected input disturbances have magnitude 1, and the allowed estimation error has magnitude 1. In cases where $\bar{\sigma}(G_u)$ is much larger than 1, it is probably not advisable (or at least not very helpful) to use the input signals for estimation. This is typically the case for distillation columns with high-purity products. For example, in our case we have at steady-state (Skogestad and Morari, 1988)

$$G_u(0) = \begin{pmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{pmatrix} \tag{32}$$

Here the gain matrix is scaled such that the allowed estimation error is 1 in mole% (corresponds to about 100% error of the nominal impurity) and the allowed disturbances on the inputs, L and V, are equal to the feed rate (corresponds to about 30% of the nominal inputs). The largest singular value, $\bar{\sigma}(G_u(0))$, is 197.2. We conclude that the

use of input signals will not be very helpful for estimation in this case. This was also confirmed by the results in this paper.

We might consider basing the estimation on, for example, D and V (DV- -configuration) rather than L and V. In this case we have $u = [DV]^T$ and the gain matrix becomes (Skogestad et al., 1988)

$$G_u^{DV}(0) = \begin{pmatrix} -87.8 & 1.4 \\ -108.2 & -1.4 \end{pmatrix} \tag{33}$$

This assumes disturbances on D of magnitude F. This seems large, and a value of $0.1F$ may seem more reasonable. Rescaling the gain matrix gives

$$G_u^{DV}(0) = \begin{pmatrix} -8.78 & 1.4 \\ -10.82 & -1.4 \end{pmatrix} \tag{34}$$

However, also in this case the elements in the matrix (and therefore also $\bar{\sigma}(G_u)$) are rather large, and the estimate will be sensitive to input errors.

2. Input uncertainty. In this case we have

$$u = (I + \Delta_I)u_c \tag{35}$$

where the uncertainty matrix $\Delta_I = diag\{\Delta_j\}$ is a diagonal matrix consisting of the relative input errors on each input channel j. The estimation error becomes

$$y - \hat{y} = G_u \Delta_I u_c \tag{36}$$

From Eq. (28) we have $u_c = G_u^{-1}\tilde{y}$ and we get

$$e_1 = y - \hat{y} = G_u \Delta_I G_u^{-1} \hat{y} \tag{37}$$

Skogestad and Morari (1987) found that the i'th diagonal element of the term $G_u \Delta_I G_u^{-1}$ is given by $\Sigma_j \lambda_{ij}(G_u)\Delta_J$ where λ_{ij} denotes the ij'th RGA-elements. Consequently, in the presence of input uncertainty, the estimation error e_1 is likely to be very large for plants with large RGA-elements. Note that this result is independent of the controller used. The model Eq. (32) used throughout this paper has diagonal RGA-values of 35.1, and we obtain at steady-state

$$G_u \Delta_I G_u^{-1} = \begin{pmatrix} 35.1\Delta_1 - 34.1\Delta_2 & -27.7\Delta_1 + 27.2\Delta_2 \\ 43.2\Delta_1 - 43.2\Delta_2 & -34.1\Delta_1 + 35.1\Delta_2 \end{pmatrix} \tag{38}$$

The elements in this matrix may be large even for very small relative input gain, Δ_j. This is consistent with the μ-analysis of the estimation error ("open-loop") where we

observed very large μ-values for the Brosilow and Kalman(K4) estimators for the case with uncertainty (eg., lower left in Fig.24).

For plants with small RGA-elements the diagonal elements in the error term $G_u \Delta_I G_u^{-1}$ are small, but the off-diagonal elements may still be large. However, for the DV-configurations mentioned above the offdiagonal elements represent no problem. The model in (34) has diagonal RGA-values of 0.45 and we obtain

$$G_u^{DV} \Delta_I G_u^{DV-1} = \begin{pmatrix} 0.45\Delta_1 + 0.55\Delta_2 & 0.45\Delta_1 + 0.45\Delta_2 \\ 0.55\Delta_1 + 0.55\Delta_2 & 0.55\Delta_1 + 0.45\Delta_2 \end{pmatrix} \tag{39}$$

All elements in this matrix are small. This implies that an estimator which uses information about D and V will *not* be sensitive to input uncertainty. However, as noted above it may still be sensitive to input disturbances. Furthermore, we found for the PCR estimator, that adding input information did not improve the estimate significantly even the case of no input error (caused by disturbances or uncertainty). The reason is that the temperature measurements contain most of the relevant information.

When secondary measurements, θ, are used by the estimator, then some of the input error may be detected and corrected for. Nevertheless, the results above demonstrate that the estimator should *not* use information about the input signals for plants where either 1) G_u (when appropriately scaled) contains large elements, or 2) G_u has large RGA-elements. Both these cases are often encountered for ill-conditioned plants. Note that the RGA is independent of scaling.

The conclusion for our distillation column is to base the estimate on temperature measurements only. Input information does not improve the estimate because of 1) sensitivity to input error, 2) poor dynamic response when used in a static estimator, and 3) the fact that the temperatures contain so much information that the estimate is not improved significantly (even if we disregard the first two items).

6.5 μ-analysis of estimators

The structured singular value, μ is a powerful tool for comparing multivariable linear systems with unknown disturbances and uncertainty, without having to perform a large number of simulations. Since μ is a worst case measure, this tool discovers explicitly the weak spots in a system. For example, it would have been much more difficult to discover the estimators' sensitivity to input uncertainty from simulations. However, the test requires additional modelling effort to capture the uncertainty in an adequate way.

While using the μ analysis we encountered problems with how to include measurement noise. Modelling it as independent disturbances would give a worst-case combination which would be extremely unlikely to occur when there are many temperatures. Therefore, in the μ-analysis we added the noise as $n = kn_0$, where k is a frequency-dependent constant to be varied in the μ-analysis, but where n_0 is a constant random vector. This approach works well when comparing estimators with the same location and number of measurements. However, in other cases the specific value of the random numbers in the noise vector n_0 may be important and may bias the μ-values. When comparing various PCR-estimators we therefore did not include noise in the μ-analysis. However, here we included noise on the calibration sets.

6.6 Nonlinearity

All models used in this paper are linear. This simplifies the problem and is necessary for using the μ-analysis. But distillation columns are known to be very nonlinear, so the effect of nonlinearity should be taken into consideration. Nevertheless, in general a system that does not perform well in the linear case, will not perform well in the nonlinear case, and a linear study is therefore a good first step in a performance evaluation.

The Kalman Filter may be extended to the nonlinear case using the so-called Extended Kalman Filter, where the process matrices and the gains are updated on-line. For distillation columns this may give a heavy computer load. For the PCR/PLS estimator the use of additional principal components may be used to eliminate some of the nonlinearity.

For distillation columns an alternative way to counteract nonlinearity is to use logarithmic transformations of the compositions (Joseph and Broslilow, 1978, Skogestad and Morari, 1988). This approach may be used for all estimators. In another paper by the authors, the questions of nonlinearity and multicomponent mixtures will be discussed in detail.

6.7 Obtaining and implementing the estimators

Both the Kalman filter and the Brosilow estimator require a linear open-loop model. On the other hand, the PCR approach only deals with the data. This is an advantage, especially when experimental data are used, but also when we do have a good model, as in this paper, since we save a significant effort in obtaining the linear model matrices.

To obtain the Kalman filter one must specify weighting matrices for noise and disturbances. These may be difficult to determine *a priori*, especially since the best value of these weights may not be physically meaningful. The Brosilow estimator has the advantage of having essentially no tuning parameters, but this makes it inflexible, and it does not work for ill- conditioned plants. Although not discussed in this paper, the PCR/PLS estimator depends strongly on the scaling of the variables. These scalings are then effective tuning parameters, which are used primarily to reflect the measurement noise. In this paper no variable scaling was applied.

As for implementation, the static Brosilow and PCR estimators are of course much simpler than the dynamic Kalman filter. For all estimators it is necessary to have some scheme for dealing with measurement failures, that is, to detect and correct outliers.

Pressure variations were not included in this study. Pressure compensation is easily included in the PCR estimator if different pressure levels and pressure drops are included in the calibration data set.

7 Conclusions

1. With the Kalman and PCR estimators, the estimate is improved by adding temperature measurements. With more than three temperatures the improvement for our example column is mainly to reduce the effect of measurement noise. The Brosilow estimator does not handle collinearity well and the estimate is not improved by adding temperatures. In general, one should not use few measurements (that is, delete measurements), but rather use only a few *combined* measurements (in the dominant directions of the measurement space).

2. From a theoretical point of view it is obvious that one may always improve the estimate by "appropriate" use of additional information (measurements). However, in some cases the usefulness of the additional information may be minimal (see use of inputs below). In other cases the improvement of the estimate must be traded off against the cost of obtaining the measurements and the increased chance of failures. Therefore, in practice one may not always want to use additional measurements.

3. For plants with large elements in the appropriately scaled gain matrix, G_u, the presence of *input disturbances* implies that the use of input signals does *not* improve the estimate. This will be the case for most high-purity distillation columns.

4. For ill-conditioned plants with large RGA-elements for G_u, the presence of *input uncertainty* implies that the use of input signals does *not* improve the estimate. This was illustrated for our example column by the Kalman filter where the best tuning corresponds to not using information about the input flows.

5. The Brosilow estimator uses the inputs directly and the estimate may be very sensitive small errors in input measurements. In the modified Brosilow estimator, introduced in this paper, the inputs are regarded as disturbances and this sensitivity is avoided. However, the estimate remains sensitive to small model errors if the condition number is large.

6. In the case of perfect models the modified Brosilow estimator and the (linear) PCR estimator are equivalent. This is quite obvious since both minimize the 2-norm of the estimation error. However, for ill-conditioned plants, PCR is better behaved numerically and is less sensitive to model errors.

7. When the dynamic response of the process outputs and the secondary measurements are similar, a static estimator may be sufficient. This is the case for our distillation example when inputs are *not* used.

8. For our distillation example, the PCR and Kalman estimators were almost identical in the closed-loop μ-test. The Kalman filter is more difficult to implement, requires more computer time, and needs initialization of the states. Thus the much simpler static PCR estimator is preferable.

9. The PCR and PLS estimators gave very similar results for our linear distillation example.

10. The exact location of the temperature measurements is important when few measurements are used, but is less critical for our example when we have about four or more measurements. 10. It is important to check the performance of an estimator both in "open loop" (estimation error) and in "closed loop" (control error). Some errors in the "open loop" estimation may have only minor influence in closed loop. One disadvantage with the closed loop test is that it depends heavily on the controller chosen.

11. Although the μ-analysis has some difficulties of representing noise, it was found to be

most suitable for studying the performance of the estimators for this ill-conditioned plant.

In conclusion, we believe that our study presents a number of results which may prove useful in practical control of distillation columns. Temperature measurements are reliable and without delay, and the need for on-line GC measurements, which are very unreliable, may be eliminated. However, a less frequent update based on, for example, using off-line GC may be needed. We also believe that our comparisons of various estimators, and the analysis of sensitivity to input error, are of interest from a general point of view.

NOMENCLATURE.

d - disturbances
D - distillate flow rate
$\tilde{d}$ - external inputs in μ-analysis
F - feed flow rate
F_i ,F' - Gain matrixes from inputs to secondary measurements (temperatures)
G_i, G' - Gain matrixes from inputs to primary outputs.
K - estimator matrix
K_f - Kalman filter gain
L - reflux flow rate
N - number of theoretical trays
N_F - location of feed tray
p - loading vector (direction of principal component) or no. of y-variables
PCR - Principal Component Regression.
q - no. of θ-varibles
q_F - fraction liquid inn feed
t - principal component (score), latent variable
T - matrix of scores
u - manipulated inputs ($= (L, V)^T$)
v - process noise (disturbance)
$\mathcal{V}$ - process noise covariance matrix.
V - boilup rate from reboiler
w - measurement noise
w_i - input uncertainty weight
w_p - performance weight

$\mathcal{W}$ - measurement noise covariance matrix.

x_B - mole fraction of light component in bottom product

y - output vector $= (y_D, x_B)^T$

y_D - mole fraction of light component in distillate

z_F - mole fraction of light component in feed

Greek symbols

α - relative volatility

Δ - uncertainty block

$\gamma(A)$ - condition number of matrix A

μ - Structural Singular Value

ω - frequency (min^{-1})

$\sigma_i(A)$ - The i'th largest singular value of matrix A

θ - secondary measurements (temperature vector)

$\tilde{\theta}$ - vector of all available information

Θ - data matrix of θ

REFERENCES

Doyle, J. C., 1982, "Analysis of Feedback Systems with Structured Uncertainties", *IEEE Proc.*, **129**, D, 242–250.

Höskuldsson, A., 1988, "PLS Regression Methods", *Jour. of Chemometrics*, **2**, 211–228.

Joseph, B., C. B. Brosilow, J. C. Howell and W. R. D. Kerr, 1976, "Multi-temps give better control." *Hydrocarbon Processing*, **3**, 127–131.

Joseph, B., C. B. Brosilow, 1978, "Inferential Control of Processes. Part I. Steady State Analysis and Design.",*AIChE Journal*, **24**, 485–492.

Kalman, R. E., and R. S. Bucy, 1961, "New results in Linear Filtering and Prediction Theory", ASME, *J. Basic Eng.*, **83**, 95–108.

Keller, J.P, and D. Bonvin, 1987, "Selection of inputs for the purpose of model reduction and Controller Design", Presented at the 10^{th} *World Congress on Automatic Control (IFAC)*, 1987 in Munich.

Moore, C., Hackney, J. and Canter, D., 1987, "Selecting Sensor Location and Type for Multivariable Processes" *Shell Proc. Contr. Workshop*, Butterworth Publishers, Boston.

Morari, M. and Stephanopoulos, G., 1980, "Optimal Selection of Secondary Measurements within the Framework of State Estimation in the Presence of Persistent Unknown Disturbances", *AIChE Journal*, **26**, 247–259.

Nisenfeld, A. E. and R. C. Seeman, "Distillation Columns", ISA Monograph Series 2, 1981.

Patke, N. G., P. B. Deshpande and A. C. Chou, 1982, "Evaluation of Inferential and Parallel Cascade Schemes for Distillation Control", *Ind. Eng. Chem. Process Des. Dev.* **21**, 266–272.

Skogestad, S. and M. Morari, 1987 "Implication of Large RGA-Elements on Control Performance", *Ind. & Eng. Chem. Res.*, **26**, 11, 2121–2330.

Skogestad, S. and M. Morari, 1988, "LV-Control of a High-Purity Distillation Column", *Chem. Eng. Sci.*, **43**, 1, 33–48.

Skogestad, S., M. Morari and J.C. Doyle, 1988, "Robust Control of Ill-Conditioned Plants: High-Purity Distillation", *IEEE Automatic Control*, **33**, 12, 1092–1105.

Strang, G., 1980, "Linear Algebra and its Applications", Harcourt Brace Jovanovich, New York.

Weber, R. and Brosilow, C., 1972, "The Use of Secondary Measurements to Improve Control", *AIChE Journal*, **18**, 614–623

Wold, S., Espensen, K. and Geladi, P., 1987, "Principal Component Analysis", *Chemometrics and Intel Lab Syst*, **2**, 37–52.

Yu, C. C. and W. L. Luyben, 1987 "Control of Multicomponent Distillation Columns using Rigorous Composition Estimators." *I. Chem. E. Symp. Ser.*,**104**, 29–69.