

Effect of Varying Controller Parameters in Closed-Loop Subspace Identification

Morten Bakke* Tor A. Johansen* Sigurd Skogestad**

* *Dep. of Engineering Cybernetics, NTNU, Trondheim, Norway.*

** *Dep. of Chemical Process Engineering, NTNU, Trondheim, Norway.*

Abstract: It is well known that subspace identification methods that assume open loop data without correlations between the input and noise, may give biased estimates when applied to closed loop data. The effect of the controller gain parameters on the quality of the identified model is studied when closed loop data are used. Several subspace identification methods (both open loop and closed loop methods), and different simulated data sets ranging from ideal 2 x 2 linear systems, to a fairly realistic nonlinear debutanizer process simulator, are considered. The results show that up to a point, higher controller gain during the identification experiment gives more accurate models than with lower controller gain, for both open and closed loop subspace identification methods. It is observed that the sensitivity to the controller gain is very small for the closed loop sub-space method tested for the ideal cases when its assumptions are satisfied. An explanation for this is that in this case there will be no bias, while the open loop methods may have a bias that depends on the controller parameters. Another interesting observation is that for the debutanizer example, the nonlinearities seem to lead to biased estimates also with the closed loop subspace method, and the choice of controller gain appears to be just as important as the use of a closed loop subspace identification method for the accuracy of the estimates.

Keywords: Closed-loop identification, Subspace identification, Experiment design, Debutanizer.

1. INTRODUCTION

A mathematical model is often used to develop *model-based* controllers. Likewise, in order to find good tuning parameters for conventional P/PID/PID-controllers, some information from the system to be controlled is required, like a low order model approximation. The term *identification for control* (Gevers (2004), Hjalmarsson (2004)) considers experimental conditions for identification of a model for either tuning, synthesis, updating or realization of these controllers.

When feedback control is used, a model that is accurate near the bandwidth frequency of the closed-loop system is important. The model may be mediocre for other frequencies, due to the "forgiving effect" of feedback control. This corresponds to the frequencies where the closed-loop sensitivity function is small, being low or high frequencies. A good model fit around the bandwidth frequency is needed for robust performance of the closed-loop system, see van den Hof and Schrama (1995). Use of bandwidth frequency of the closed-loop system to extract information relevant for control purposes is a well-known technique, and stems back to the work of Ziegler and Nichols (1942). For multivariable systems, the bandwidth frequencies may not be well-defined. An overview of different techniques for auto-

tuning of both single-loop and multi-loop controllers based on bandwidth frequencies is given in Hang et al. (2002).

Recently a system identification technique based on state-space methods has emerged, known as *subspace identification* (see Qin (2006) for an overview of subspace methods). A great advantage of subspace identification methods is that they are non-iterative, using well-understood linear algebra algorithms with good numerical properties. They are also known to cope well with large data sets, rendering it possible to identify large systems in a fair amount of time. An example of this is given in Juricek et al. (2001), where the Tennessee Eastman challenge process is subject for identification. It is shown that the only successful models are the state-space models produced by two different subspace algorithms.

The goal of the present work is to investigate MIMO system identification in closed loop using subspace Identification Methods. Most of the theory on subspace methods assumes that datasets are collected in open-loop, and earlier results show that ordinary subspace methods may fail when closed-loop data is applied. The problem using subspace methods with closed-loop data, is that *instrumental variables* are used to filter out noise in the data record. For noisy closed-loop data, the noise on the input data will be correlated with the noise on the output data due to the feedback mechanism. An assumption that is made when instrumental variables are used, is that noise on the input data is *uncorrelated* with noise on the output data (see Ljung (1999)). If this is not satisfied, the subspace

¹ Morten.Bakke@Honeywell.com. Present address: Honeywell Process Solutions, Asker, Norway.

² Tor.Arne.Johansen@itk.ntnu.no

³ Sigurd.Skogestad@chemeng.ntnu.no

algorithms will in general return biased system parameters. Recently, several closed-loop subspace methods try to account for this Verhagen (1993); Ljung and McKelvey (1996); Qin and Ljung (2003); Jansson (2003); Qin (2006); Katayama et al. (2005); Wang and Qin (2006); Gilson and Mercere (2006); Jansson (2005), and consistency analysis of different closed-loop subspace methods is given in Lin et al. (2004) and Chiuso and Picci (2005).

While the asymptotic variance can be computed explicitly for some closed-loop subspace methods under certain assumptions, Chiuso (2006), we conduct an empirical simulation study to investigate the performance of different subspace methods when closed-loop data sets are applied. The systems that are subject to investigation here are all multivariable, and operate under feedback from decentralized PI controllers. The main contribution of this work is investigations on the performance of subspace identification methods when the controller parameters are varied, also under conditions with process nonlinearities and finite data sets. Insight into this contributes to explain why closed-loop experiments in some cases may fail, and provides knowledge on how the controllers should be tuned to give good identification conditions. Additional results and simulations are available in Bakke (2009).

2. BACKGROUND

2.1 Closed-Loop Identification

Open-loop identification is a common method for system identification. Consider Figure 1: The experiment design involves manipulation of the variable u_m in order to identify the dynamics from u_m to y . There are many systems where prior knowledge tells us that it is *practically impossible* to break the feedback loop in order to identify the open-loop dynamics. This is in particular true for processes with unstable behavior, and systems with inherit feedback effects (e.g. economic and biological systems). As stated by van den Hof (1998), it will in many cases be too high a risk, or cost, to break feedback loops for identification purposes, since the outputs may drift away from their nominal operating points during the identification experiment due to unknown disturbances.

When a system/process is identified in *closed-loop*, the open-loop dynamics of the given system is estimated while the controllers already present are calculating the input u to the system. When a process G_0 is subject to feedback control by a controller C , the degree of freedom for the user to excite the system is usually moved from u to the controller reference signal r . Consider the block diagram in Figure 2: Here the reference signal is denoted r_2 , and an additional external input signal r_1 is directly added to the system input u . Throughout this work, it is assumed that r_1 is an unknown signal in the form of process noise, and the reference r_2 is instrumental to externally excite the system dynamics. The common approach to closed-loop identification is to generate an input sequence to the system by varying the reference signal, and collect u and y data from the system to be used for identification. There are three common methods used for closed-loop identification, according to Ljung (1999). These are *the direct method*, *the indirect method*, and *the joint input-output method*. The direct method operates directly on

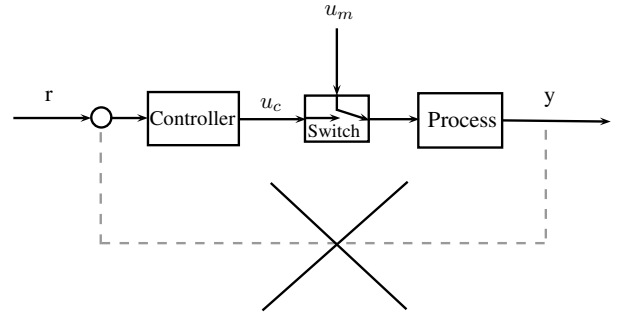


Fig. 1. Feedback loop broken before identification experiment

u and y data from the process, ignoring the feedback. The indirect method assumes a known, linear controller. The process is identified from r to y , and the closed-loop transfer function is estimated. From knowledge of the controller, the open-loop transfer matrix is calculated. The joint input-output method assumes that the closed-loop system is a system with one input (r), and two outputs (u and y). The system is identified from r to y (\hat{G}_{ry}), and from r to u (\hat{G}_{ru}). The open-loop estimate is calculated as $\hat{G} = \hat{G}_{ry} \cdot (\hat{G}_{ru})^{-1}$.

Since the indirect method is limited to linear, known controllers, and the joint input-output method relies upon calculating the inverse of a possibly close-to-singular transfer function, the direct method is used throughout this work. One of the main reasons why open-loop experiments

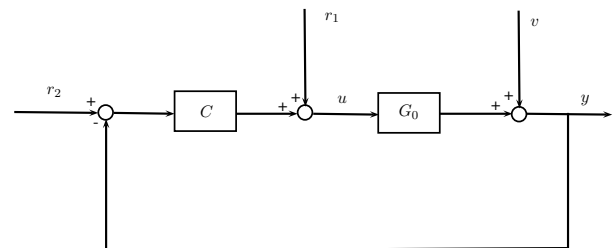


Fig. 2. A system operating in closed-loop with process and measurement noise

are preferred is the lack of correlation between process- and measurement noise, and the input u to the system. Under closed-loop operation, the input to the process is the controller command u_c , which is a function of the output(s) of the system. Hence, the disturbances that affect the output variables y will also affect the controller command u_c through the feedback loop.

2.2 Subspace Identification Methods

The family of subspace algorithms seek to estimate the system matrices A , B , C and D of a discrete-time, linear state space model from input-output data of the system. A general subspace identification algorithm is given in Ljung (1999). All subspace identification methods consist of three main steps:

- (1) Estimating the predictable subspace for multiple future steps

- (2) Extract the state variables from the estimated subspace
- (3) Fitting the estimated states to a state space model

In this paper, three different subspace identification algorithms are applied to closed-loop data; the MOESP algorithm (discussed in Verhaegen (1994)), the MATLAB implementation of the N4SID algorithm (see van Overschee and de Moor (1994)), and the DSR_e algorithm (Di Ruscio (2008), see also Di Ruscio (2003) and Di Ruscio (1996)). Out of the three methods used in this work, only the DSR_e method is particularly designed for closed-loop data. This method identifies the Kalman filter of the system, using filtered outputs in the identification routine.

3. RESULTS

3.1 Preliminary discussion

In order to generate identification and validation data from the systems to be simulated, PRBS-signals are generated and applied to the references of the systems. Different PRBS-sequences with the same frequency spectrum are generated for identification and validation respectively.

It is useful to define some performance metrics for the models identified. Two indices are defined in order to compare performances of identified models:

$$\text{MRSE} = \frac{1}{l} \sum_{i=1}^l \sqrt{\frac{\sum_{j=1}^N (y_i(j) - \hat{y}_i(j))^2}{\sum_{j=1}^N y_i(j)^2}} \quad (1)$$

$$\text{MVAF} = \frac{1}{l} \sum_{i=1}^l \left(1 - \frac{\text{variance}(y_i - \hat{y}_i)}{\text{variance}(y_i)}\right) \quad (2)$$

Here, y_i means system outputs, and \hat{y}_i means model outputs (one-step ahead predictions). The MRSE index given in (1) is used to measure the Mean Relative Squared Error between the real process outputs and the outputs predicted by the model. As seen by (1), an MRSE index of 0 indicates a perfect model.

MVAF in (2) stands for Mean Variance Accounted For, and is a measure for evaluating the dynamic properties of the produced models. If the ratio $\frac{\text{variance}(y_i - \hat{y}_i)}{\text{variance}(y_i)}$ is small, the MVAF index is close to 1.

White noise is added to both the process inputs and outputs. For the idealized case studies, the identification experiments are first performed without any noise, and repeated with both process and measurement noise. When there is noise present in the loop, Monte Carlo simulations are used with 10 simulations, reporting only the mean value of the performance indices from these experiments. For the nonlinear debutanizer case, only the scenario with both process and measurement noise is used.

Three different controller gain schedules are applied for all systems; low gain, medium gain and high gain. The same subspace algorithm parameters and PRBS signals are used for all gain schedules.

3.2 Idealized Case Studies

In this section, two linear, multivariable systems with different dynamic properties are simulated. In both cases,

the systems are operating under decentralized PI-control. The two systems are defined as follows, on transfer matrix form:

$$\mathcal{S}_1 = \frac{1}{2s+1} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad (3)$$

$$\mathcal{S}_2 = \begin{bmatrix} \frac{2s-1}{(3s+1)(s+1)} & \frac{1}{2s+1} \\ \frac{1}{50s+1} & \frac{1}{(70s+1)(30s+1)} \end{bmatrix} \quad (4)$$

Systems \mathcal{S}_1 and \mathcal{S}_2 corresponds to G_0 in Figure 2. The controller which is used for both systems is

$$\begin{bmatrix} C_1(s) & 0 \\ 0 & C_2(s) \end{bmatrix} = \begin{bmatrix} K_{c1} \frac{\tau_i s + 1}{\tau_{i1} s} & 0 \\ 0 & K_{c2} \frac{\tau_i s + 1}{\tau_{i2} s} \end{bmatrix} \quad (5)$$

System \mathcal{S}_1 is a 2×2 , weakly coupled system with a common denominator. The RGA of the system is calculated as (see Skogestad and Postlethwaite (2005))

$$\Lambda(G_0) = \begin{bmatrix} \lambda_{11} & 1 - \lambda_{11} \\ 1 - \lambda_{11} & \lambda_{11} \end{bmatrix}; \quad \lambda_{11} = \frac{1}{1 - \frac{g_{12}g_{21}}{g_{11}g_{22}}} = 2 \quad (6)$$

System \mathcal{S}_2 is also a 2×2 , weakly coupled system, but this system is also relatively stiff. The RGA of this system is $\frac{2}{3}$. The ratio between the highest and lowest eigenvalues is for this system about 70. Table 1 shows the results from the simulated identification experiments for systems \mathcal{S}_1 and \mathcal{S}_2 respectively. Figure 3 shows the indices as a function of controller gains, when the gains are varied in parallel, in small steps. For each step calculated, 10 simulations are used, and the mean values of the indices are used as the actual values. The closed-loop method DSR_e is less affected by increasing controller gain than the open-loop methods, which is reasonable since it (unlike the other methods) is guaranteed to give unbiased estimates.

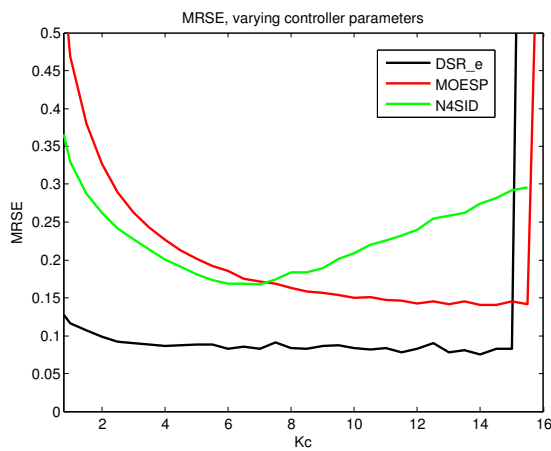
3.3 Nonlinear Debutanizer Case

Figure 4 shows a Process Flow Diagram for the debutanizer. The debutanizer is implemented in UniSim Design by Honeywell, see UniSim (2005).

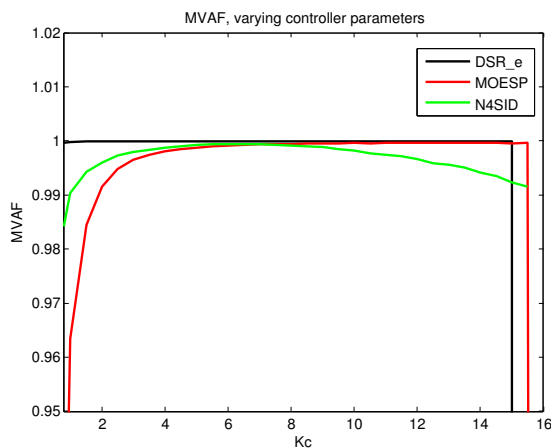
The reboiler heats the input flow so that hot vapour is sent back into the column, in order to control the temperature at a specific stage in the column. The effect of this reboiler is controlled by the temperature controller TIC-100, which in turn is controlled by a master bottoms composition controller XIC-101. The reflux flow is controlled by the flow controller FIC-100, which is a slave controller to the distillate composition controller XIC-100. In addition, there are two pressure controllers that controls the column pressure in a bypass configuration. As seen by the process flow diagram, this pressure controller will induce a flow into the reflux drum, which is under level control. This means that pressure control action will act as an "unknown disturbance" on the reflux flow through the reflux drum level controller LIC-102, seen from the flow controller FIC-100. This will in turn have effect on the compositions, since the flow controller adjust the flow to achieve its setpoint, and change of reflux flow means change of compositions. This will also affect the column temperature, as there is a physical interaction between the reflux flow into the column and the temperature.

Table 1. Performance indices from identification of system \mathcal{S}_1 : PN = Process Noise, MN = Measurement Noise, LG = Low Gain, MG = Medium Gain, HG = High Gain. "Failed" means that MRSE > 100% and MVAF < 85%.

System	PN	MN	Method	MRSE/MVAF (LG)	MRSE/MVAF (MG)	MRSE/MVAF (HG)
\mathcal{S}_1	No	No	DSR_e	0%/100%	0%/100%	0%/100%
			MOESP	0%/100%	0%/100%	0%/100%
			N4SID	0%/100%	0%/100%	0%/100%
\mathcal{S}_1	Yes	Yes	DSR_e	13%/99.99%	10.5%/100%	10%/100%
			MOESP	51%/94%	33.5%/98%	24.5%/99%
			N4SID	36.5%/98%	27%/99%	22%/99.4%
\mathcal{S}_2	No	No	DSR_e	0%/100%	0%/100%	0%/100%
			MOESP	0%/100%	0%/100%	0%/100%
			N4SID	0%/100%	0%/100%	0%/100%
\mathcal{S}_2	Yes	Yes	DSR_e	19%/99.1%	12.1%/99.9%	27%/98.1%
			MOESP	failed	failed	failed
			N4SID	14.2%/99.8%	12.1%/99.9%	55.3%/90.4%



(a) MRSE



(b) MVAF

Fig. 3. Performance of identified models for system \mathcal{S}_1 for increasing controller gains, $K_c = K_{c1} = K_{c2}$

The debutanizer model itself is nonlinear, and the valves are modeled to have linear characteristics. These are fast compared to the rest of the system. In addition, the process is interactive, since the reflux-flow will directly affect the bottoms composition, and the temperature inside the column will also affect the distillate composition.

The references, inputs and outputs of this system are defined as follows:

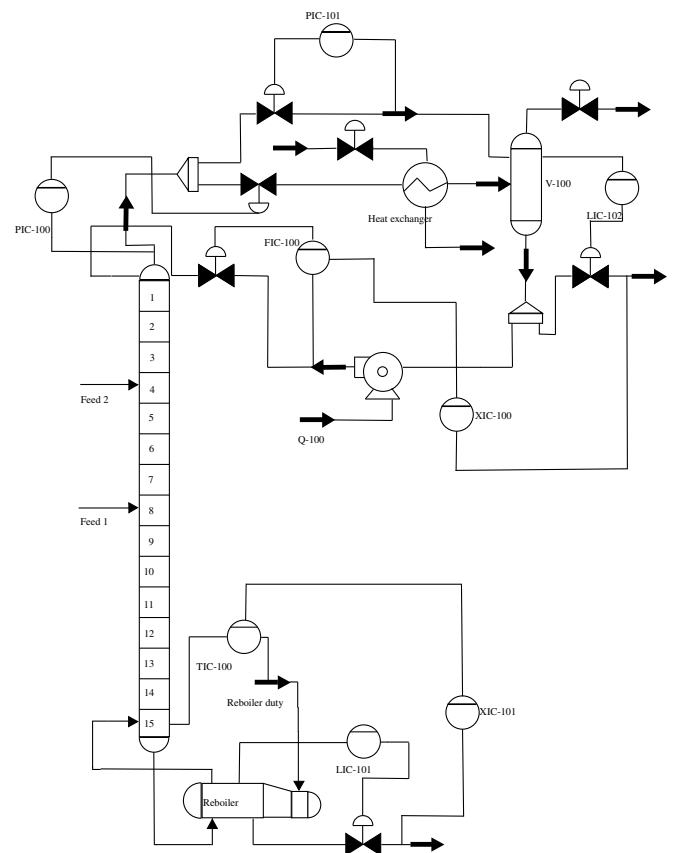


Fig. 4. Process Flow Diagram for the debutanizer

- r_1 - setpoint for the controller XIC-100 [Mole Fraction]
- r_2 - setpoint for the controller XIC-101 [Mole Fraction]
- u_1 - Setpoint to the flow controller FIC-100, which controls the reflux flow into the column $\left[\frac{m^3}{h}\right]$
- u_2 - Setpoint for TIC-100, which controls stage 8 temperature $[\text{°C}]$
- y_1 - Composition of heavy key component (i-pentane) in the distillate flow [Mole Fraction]
- y_2 - Composition of light key component (n-butane) in the bottoms flow [Mole Fraction]

Three different controller gain schedules are used also in this case.

- **Configuration 1** - $K_{c1} = 1.29, K_{c2} = 0.319$
- **Configuration 2** - $K_{c1} = 1.935, K_{c2} = 0.4785$
- **Configuration 3** - $K_{c1} = 2.58, K_{c2} = 0.638$

In all cases, the integral times are $\tau_{i1} = 38.3[\text{min}]$ and $\tau_{i2} = 13.5[\text{min}]$. Table 2 shows the performance indices for the different controller tuning configurations. Figure 5 shows the performance of the DSR_e method for these configurations.

Table 2. Performance indices, debutanizer

Configuration	Method	MRSE	MVAF
1 (Low gain)	DSR_e	70.11%	76.28%
	MOESP	89.46%	49.83%
	N4SID	73.13%	69.93%
2 (Medium gain)	DSR_e	56.32%	89.88%
	MOESP	87.9%	43.74%
	N4SID	60.01%	86.22%
3 (High gain)	DSR_e	49.76%	93.68%
	MOESP	91.7%	46.55%
	N4SID	56.43%	88.67%

4. DISCUSSION

As shown for both system S_1, S_2 , the identification performance was degraded for too high a value for the gains. This is shown to be true also for the debutanizer process, see Bakke (2009). A possible explanation to this is that when the controller gains are increased, frequencies in the input signals will be more concentrated around the achievable bandwidth of the controller, which produces system outputs with more information of the frequency response around this bandwidth frequency. If the controller gains are pushing this limit, the controller will be so aggressive that it dominates over the open-loop dynamics, and more of the controller is seen on the identified models.

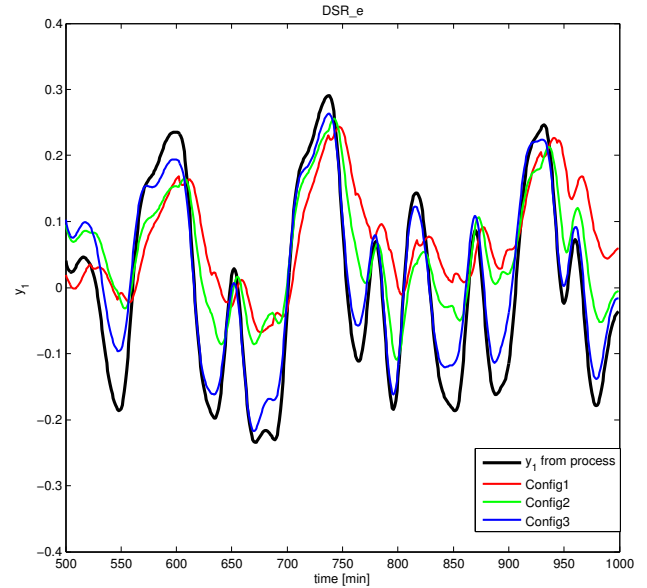
Ljung (1999) mentions on p. 434 that the part of the system input that originates from the feedback has no information value when it comes to identifying the open-loop dynamics. On the other hand, it is the part of u (controller output, system input) that stems from the reference signal that will reveal information from the open-loop system, and give a lower signal to noise ratio. In this context, it is explained why the models identified with high gains in the feedback loops performed better than those identified with the low gains in the loops. Consider the controller equation that has been used throughout this work:

$$u(s) = \begin{bmatrix} C_1(s) & 0 \\ 0 & C_2(s) \end{bmatrix} (r - y) \quad (7)$$

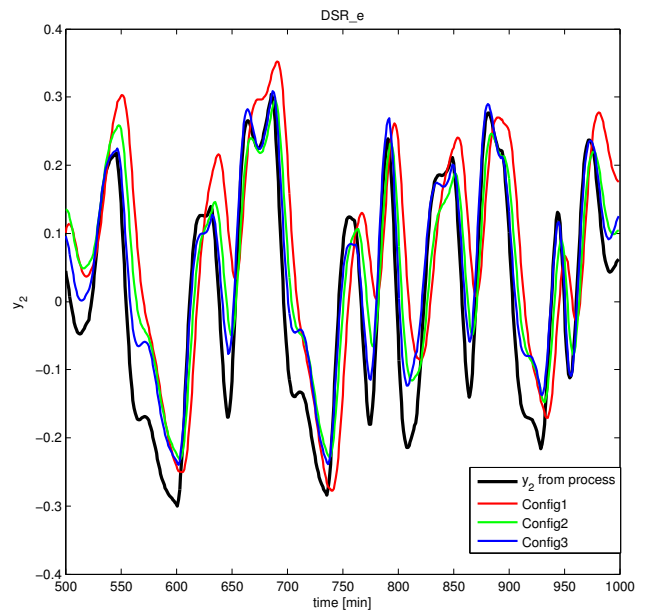
The same reference sequences are used when comparing all the methods, for all controller parameter configurations. Notice that u consists of two different signals, namely the part that stems from the feedback (y), and the part that stems from the reference signal r . When the reference signal r is manipulated, it is the controller gain that determines how much the external reference signal is amplified. Ljung (1999) defines the covariance of the open-loop transfer function estimate as

$$\text{Cov } \hat{G}_0 = \frac{n}{N} \frac{\Phi_v(\omega)}{|S_0|^2 \Phi_r(\omega)} \quad (8)$$

Here, $\Phi_v(\omega)$ is the power spectrum of the noise, and $\Phi_r(\omega)$ is the power spectrum of the reference signal. $|S_0|^2$ is the



(a) Validation sequence for y_1



(b) Validation sequence for y_2

Fig. 5. Validation sequences for the debutanizer process using the DSR_e method

squared absolute value of the sensitivity function. As seen by equation (8) the amplification of the reference signal will directly lower the "noise to signal ratio". As stated in Ljung (1999), it is this ratio that determines how well the open-loop transfer function may be estimated when there is noise present in the loop.

In the debutanizer example, all methods may give biased estimates due to the nonlinear process and linear model. It is interesting to observe that in this case it seems like an open loop subspace method like N4SID with a high gain

tend to give more accurate estimates than a closed loop subspace method like DSR_e with a low gain.

5. CONCLUSION

The results from this work show empirically that tuning of existing controllers has a significant impact on the performance of the identified models when using subspace methods. It is shown that when controller gains are increased, controller outputs with higher signal to noise ratios are generated, and the information content of these signals is higher because of the amplification of the input reference signals. This is believed to be the main reason why data produced by the relative high gain systems yield identified models with better performance than for systems with lower gain, since noise that corrupts the system will be dominated by the power of these controller outputs. It is also shown by simulations that there exists a limit to how much the gains can be increased before the performance of the identified models are degraded.

ACKNOWLEDGEMENTS

The authors would like to thank dr. David Di Ruscio for providing Matlab code for the DSR_e method, and for insightful comments on the work.

REFERENCES

- Bakke, M. (2009). *Subspace Identification using Closed-Loop Data*. Master's thesis, Norwegian University of Science and Technology (NTNU).
- Chiuso, A. (2006). Asymptotic variance of closed-loop subspace identification methods. *IEEE Trans. Automatic Control*, 51, 1299–1314.
- Chiuso, A. and Picci, G. (2005). Consistency analysis of some closed-loop subspace identification methods. *Automatica*, 41, 377–391.
- Di Ruscio, D. (1996). Combined deterministic and stochastic system identification and realization: DSR - a subspace approach based on observations. *Modelling, identification and control*, 17, 193–230.
- Di Ruscio, D. (2003). Subspace system identification of the Kalman filter. *Modelling, identification and control*, 24, 125–157.
- Di Ruscio, D. (2008). Subspace system identification of the kalman filter: Open and closed-loop systems. *International Multi-Conference on Engineering and Technological Innovation, Orlando, Florida*, 1, 253–255.
- Gevers, M. (2004). Identification for control: Achievements and open problems. Center for Systems Engineering and Applied Mechanics, Universit Catholique de Louvain.
- Gilson, M. and Mercere, G. (2006). Subspace-based optimal IV method for closed-loop system identification. In *14th IFAC Symposium of System Identification, Newcastle, Australia*.
- Hang, C.C., Wang, Q.G., and Astrom, K.J. (2002). Relay feedback auto-tuning of process controllers - a tutorial overview. *Journal of Process Control*, 12, 143–162.
- Hjalmarsson, H. (2004). From experiment design to closed-loop control. *Automatica*, 41, 393–438.
- Jansson, M. (2003). Subspace identification and arx modelling. In *Proceedings 13th IFAC Symp. System Identification, Rotterdam*.
- Jansson, M. (2005). A new subspace identification method for open and closed-loop data. Royal Institute of Technology, Stockholm, Sweden.
- Juricek, B.C., Seborg, D.E., and Larimore, W.E. (2001). Identification of the Tennessee Eastman challenge process with subspace methods. *Control Engineering Practice*, 9, 1337–1351.
- Katayama, T., Kawauchi, H., and Picci, G. (2005). Subspace identification of closed loop systems by the orthogonal decomposition method. *Automatica*, 41, 863–872.
- Lin, W., Qin, S.J., and Ljung, L. (2004). On Consistency of Closed-loop Subspace Identification with Innovation Estimation. In *43rd IEEE Conference on Decision and Control, Bahamas*.
- Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall.
- Ljung, L. and McKelvey, T. (1996). Subspace identification from closed loop data. *Signal Processing*, 52, 209–215.
- Qin, S.J. (2006). An overview of subspace identification. *Computers and Chemical Engineering*, 30, 1502–1513.
- Qin, S.J. and Ljung, L. (2003). Closed-loop subspace identification with innovation estimation. In *Proceedings 13th IFAC Symp. System Identification, Rotterdam*.
- Skogestad, S. and Postlethwaite, I. (2005). *Multivariable Feedback Control - Analysis and design*. John Wiley & Sons, Ltd.
- UniSim (2005). *UniSim Reference Guide*. Honeywell, R350 edition.
- van den Hof, P. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, 22, 173–186.
- van den Hof, P. and Schrama, R. (1995). Identification and control - Closed-loop issues. *Automatica*, 31, 1751–1770.
- van Overschee, P. and de Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30, 75–93.
- Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models in innovations form from input-output data. *Automatica*, 30, 61–74.
- Verhagen, M. (1993). Application of a subspace identification technique to identify LTI systems operating in closed loop. *Automatica*, 29, 1027–1040.
- Wang, J. and Qin, S.J. (2006). Closed-loop subspace identification using the parity space. *Automatica*, 42, 315–320.
- Ziegler, J. and Nichols, N. (1942). Optimum settings for automatic controllers. *ASME Transactions*, 64, 759–768.