Closed-Loop System Identification for Small Samples with Constraints

O. Arda Vanli

Enrique Del Castillo

Department of Industrial and Manufacturing Engineering,

The Pennsylvania State University,

University Park, PA 16802, USA

Abstract

Traditional approaches to closed-loop identification of transfer function models require a sufficiently large data set and model forms that are general enough while at the same time requiring some form of an external excitation (a "dither signal") be applied to the process. In the limit, as the dither signal dominates the control actions, identification is easier, but the operation of the process becomes closer to that of an uncontrolled (i.e. open-loop) process, which could be unacceptable. This paper proposes a closed-loop system identification procedure that aims to improve model parameter estimates by incorporating prior knowledge about the process in the form of constraints without the use of a dither signal. A Monte Carlo simulation study is presented to illustrate the small sample benefits of adding various forms of constraints. It is shown how constraints based on process knowledge which is relatively easy to know from prior experience result in best identified models among the class of constraints considered. In particular, the knowledge of the input-output delay of the process is shown to be the most important in identifying a process in closed-loop. An example based on a real process illustrates the advantages of the proposed method over the dither signal approach.

Keywords: Box-Jenkins transfer function models, constrained non-linear least squares, prior process knowledge, feedback control.

1 Introduction

Consider a process that is adjusted, in discrete time, with a feedback controller. In many industrial settings, it is of interest that the unknown model of the process be identified while

the process is being controlled. Following the control systems literature, we refer to the identification problem as finding both the true model *form* and the true parameter values of the model of the process, thus it includes estimation.

Open-loop system identification experiments (i.e. observing the process when no feedback control is exercised and when the input is persistently exciting the process), often the preferred way of identifying the model of a process, would not be feasible or cost effective in an industrial process because the process could be unstable and drift off-target dangerously or could produce expensive scrap when uncontrolled. The main difficulty with closed-loop identification is that the data collected during closed-loop (controlled) operation is less informative than open-loop data because the control input level is frequently a linear function of the output (and hence the input is not persistently exciting the process). Our purpose in this paper is to identify and estimate the model of a process in closed-loop when a feedback controller is in operation so that, according to the identified model, the controller could be re-designed in such a way that the quality characteristic under control better achieves the objective.

Another common approach for process identification is to introduce an additional random signal in the input (the controllable factor), a so-called dither signal, in a closed-loop process (Box and MacGregor, 1974 and 1976). This, unfortunately, can be too expensive if the cost of changing the controllable factor or the cost of running the process off-target is high. These are very common conditions that justify the action of a controller (Box and Luceño, 1997). In this paper, a new method for identification of transfer function models is presented that works under closed-loop operating conditions and does not require a dither signal.

The identifiability conditions developed in the control theory literature (Ljung et. al., 1974, Söderström et. al., 1975, Söderström et. al., 1976) are useful for ensuring that the input is sufficiently exciting the process so that the model parameters can be estimated from closed-loop data. The use of the identifiability conditions in practical settings, however, is limited due to two main reasons. First, they are asymptotic conditions, thus, they only guarantee convergence to the true model as the sample size tends to infinity. The main consequence of this is that, even when the identifiability conditions are satisfied, the quality of the closed-loop parameter estimates obtained with finite samples is usually poor. It has been shown that a dither signal added to the control input can result in improved parameter estimates from small closed-loop data sets (see e.g., MacGregor and Fogal, 1995). However, the limitations of the open-loop experiments apply also to the use of a dither signal: frequent changes in the input, when excessive, makes the system to practically operate in open-loop and this can be undesirable during production.

The second limitation of the identifiability conditions is that the assumed model structure must be general enough to contain the true process model as a particular case. It is worth

pointing out the severity of this second condition. *If the assumed model does not contain the true model, the true model cannot be identified.* More precisely, model mis-specification due to not including the true model (under-specification) results in biased or non-unique estimates of the true model. Therefore, in this paper we concentrate only on the case where the true model is included in the assumed model. Since the true model form is unknown, the only way to guarantee that this condition is satisfied is to use increasingly larger models. Model mis-specification due to selecting a larger model than necessary (over-specification), however, is contrary to the parsimony principle which favors models with fewer parameters because it causes inflation in the variance of the predictions.

In this study, we propose a model identification and estimation approach that utilizes prior knowledge about the process to improve the parameter estimates obtained from closed-loop data without introducing any external excitation into the system. The proposed approach, which we will call the constrained least squares (or CLS) approach, incorporates the prior knowledge about the process as constraints on the parameters. We study the bias and variance of the parameter estimates in a simulation example. In all illustrations, we consider only processes for which the identifiability conditions hold.

The rest of the paper is organized as follows. Closed-loop identification and parameter estimation of Box-Jenkins transfer function models using the proposed approach is discussed in Section 2. We adopt a Box-Jenkins model form as it is the most natural and most common in the quality control literature. Section 3 reviews possible forms of process information that can be specified as different types of constraints. Section 4 shows how adding constraints improves process identifiability. Section 5 and 6 present the main results of the paper. In Section 5 the small sample benefits of adding constraints on the properties of the parameter estimates are studied by Monte Carlo simulation and by considering different model and sample sizes. Section 6 illustrates the application of the approach on real process data and compares it against the identification method based on a dither signal.

2 Closed-Loop identification of Box-Jenkins transfer function models

In this paper we will use Box-Jenkins (BJ) transfer function models to identify and estimate controlled processes. The BJ model form is the most natural way of modelling dynamic-stochastic processes because of its independent parametrization of the transfer function and disturbance models giving a signal-plus-noise interpretation.

Denoting the level of the input (the controllable factor) by U_t and the output (the response variable) deviation from target by Y_t, a BJ transfer function model that relates the output

to the input is given as:

$$Y_t = \frac{b(\mathcal{B})}{a(\mathcal{B})}\mathcal{B}^k U_t + \frac{\theta(\mathcal{B})}{\phi(\mathcal{B})(1-\mathcal{B})^d}\epsilon_t \tag{1}$$

where $\mathcal{B}$ is the back shift operator (i.e. $\mathcal{B}Y_t = Y_{t-1}$) and the first term on the right is the process dynamics model and the second term on the right is the disturbance, or noise, model that represents the effect of all sources of variation at the output. $k \geq 0$ is the input-output delay, $d \geq 0$ is the degree of integration of the disturbance and $\{\epsilon_t\}$ is a white noise process, independently and identically distributed (i.i.d) with mean 0 and variance σ_ϵ^2. We will refer to (1) as the assumed process model, denoted $\mathcal{M}(\boldsymbol{\beta})$. The model polynomials are defined as:

$$\begin{aligned} b(\mathcal{B}) &= b_1\mathcal{B} + b_2\mathcal{B}^2 + \ldots + b_{n_b}B^{n_b} \\ a(\mathcal{B}) &= 1 - a_1\mathcal{B} - \ldots - a_{n_a}B^{n_a} \\ \phi(\mathcal{B}) &= 1 - \phi_1\mathcal{B} - \ldots - \phi_{n_\phi}B^{n_\phi} \\ \theta(\mathcal{B}) &= 1 - \theta_1\mathcal{B} - \ldots - \theta_{n_\theta}B^{n_\theta}. \end{aligned} \tag{2}$$

The degrees or orders n_a, n_b, n_ϕ and n_θ and k and d are selected by the user. The model parameters can therefore be given in a $(p \times 1)$ parameter vector as:

$$\boldsymbol{\beta}' = (a_1, ..., a_{n_a}, b_1, ..., b_{n_b}, \phi_1, ..., \phi_{n_\phi}, \theta_1, ..., \theta_{n_\theta}) \tag{3}$$

where $p = n_a + n_b + n_\phi + n_\theta$.

In model (1) additional disturbance characteristics, such as a process offset or a drift can also be included:

$$Y_t = \alpha + \frac{\delta}{(1-\mathcal{B})^d} + \frac{\mathcal{B}^k b(\mathcal{B})}{a(\mathcal{B})}U_t + \frac{\theta(\mathcal{B})}{\phi(\mathcal{B})(1-\mathcal{B})^d}\epsilon_t \tag{4}$$

where δ is a drift constant useful in modelling possible trends in the process mean, such as tool wear which frequently occurs in machining operations. The value of α represents a process offset that determines the mean deviation from target when the process is uncontrolled and $\delta = 0$.

Closed-loop identification with Box-Jenkins models has been studied by various authors. Box and MacGregor (1976) applied the results by Söderström et. al. (1976) to processes under minimum mean squared error (MMSE) control. It was assumed that the true input-output delay and model orders were known. By the addition of a dither signal, they showed that the precision of the parameter estimates can be improved. Luceño (1997) investigated the case of switching between multiple PI controllers during the operation of the process and the effects this has on the quality of parameter estimates. Similarly, it was assumed that

the correct model orders and the delay were known. More recent approaches to closed-loop identification, the projection method (Forssell and Ljung, 2000) and the two-stage method (Van Den Hof and Schrama, 1993) break down the problem into two open-loop identification problems.

The idea of applying prior process knowledge in closed-loop identification to improve parameter estimates has been discussed by Box and MacGregor (1974) for the cases of prior knowledge on the dynamics transfer function and on the disturbance model. MacGregor and Fogal (1995) investigated the benefits of knowing the disturbance model by applying noise model pre-filters in identification. Ljung and Forssell (1998) have shown that the complete knowledge of the true disturbance model is sufficient for parameter identifiability. Pan and Del Castillo (2001) considered the stationarity and invertibility conditions as a form of prior knowledge and showed by example that closed-loop identification is possible by utilizing these conditions. While there have been several studies on utilizing prior process knowledge in process identification, no systematic study that addresses the relative effectiveness of different forms of prior knowledge was provided.

2.1 Identification and Parameter Estimation of the Process Model

The objective of closed-loop identification is to obtain the representation of a process from a finite set of input-output data $Z = \{Y_1, U_1, \ldots, Y_N, U_N\}$ collected during controlled operation of the process. The unknown true process model, that will be referred to as $\mathcal{S}$, can also be represented in the BJ model form (4). In this model the polynomials $\tilde{a}(\mathcal{B}), \tilde{b}(\mathcal{B}), \tilde{\phi}(\mathcal{B})$ and $\tilde{\theta}(\mathcal{B})$, the delay $\tilde{k}$ and the degree of integration $\tilde{d}$ for the true process are used. Throughout the paper we reserve the tilde notation for the true process description.

In the proposed approach the parameters of the process model are estimated from measured closed-loop output data and the knowledge of the controller by conducting two estimation steps. This approach is sometimes called the "indirect" approach to closed-loop identification (the "direct" approach, by contrast, requires only one estimation step, see e.g. Ljung, 1974).

In the first estimation step, the closed-loop equation of the true process is estimated by fitting an ARMA (autoregressive-moving average) time series model to the output data. The closed-loop equation of the true process that obeys the transfer function model (4) is obtained by eliminating the input and representing the output as a time series. To do this, we consider the ARMAX (autoregressive-moving average-exogenous) representation of this model, which is given as:

$$\tilde{A}(\mathcal{B})Y_t = \tilde{\gamma} + \mathcal{B}^{\tilde{k}}\tilde{B}(\mathcal{B})U_t + \tilde{C}(\mathcal{B})\epsilon_t \tag{5}$$

where

$$\tilde{A}(\mathcal{B}) = (1-\mathcal{B})^{\tilde{d}}\tilde{\phi}(\mathcal{B})\tilde{a}(\mathcal{B}), \qquad \tilde{B}(\mathcal{B}) = (1-\mathcal{B})^{\tilde{d}}\tilde{\phi}(\mathcal{B})\tilde{b}(\mathcal{B}), \qquad \tilde{C}(\mathcal{B}) = \tilde{\theta}(\mathcal{B})\tilde{a}(\mathcal{B}) \tag{6}$$

are the polynomials and $\tilde{\gamma} = \left[\tilde{\alpha}(1-\mathcal{B})^{\tilde{d}} + \tilde{\delta}\right] \tilde{a}(\mathcal{B})\tilde{\phi}(\mathcal{B})$ is the intercept. Suppose that a linear feedback controller of the form

$$U_t = \frac{D(\mathcal{B})}{F(\mathcal{B})} Y_t \tag{7}$$

is acting on the process. By inserting this in (5) we can obtain the ARMA closed-loop equation of the true process as:

$$\begin{aligned} \left(\tilde{A}(\mathcal{B})F(\mathcal{B}) - \mathcal{B}^{\tilde{k}}\tilde{B}(\mathcal{B})D(\mathcal{B})\right) Y_t &= F(\mathcal{B})\tilde{\gamma} + F(\mathcal{B})\tilde{C}(\mathcal{B})\epsilon_t \\ &\text{or} \\ \tilde{\Phi}(\mathcal{B})Y_t &= \tilde{\xi} + \tilde{\Theta}(\mathcal{B})\epsilon_t \end{aligned} \tag{8}$$

which can be estimated by fitting an ARMA time series model to the measured output data $\{Y_t\}_{t=1}^{N}$. Note that the AR and the MA polynomials of this model are $\tilde{\Phi}(\mathcal{B}) = 1 + \tilde{\Phi}_1\mathcal{B} + \ldots + \tilde{\Phi}_{\tilde{n}_\Phi}\mathcal{B}^{\tilde{n}_\Phi}$ and $\tilde{\Theta}(\mathcal{B}) = 1 + \tilde{\Theta}_1\mathcal{B} + \ldots + \tilde{\Theta}_{\tilde{n}_\Theta}\mathcal{B}^{\tilde{n}_\Theta}$, respectively.

In the second estimation step, the parameters of the assumed model (4) are estimated by minimizing the difference between the (estimated) closed-loop equation (8) of the true process and the closed-loop equation of the assumed model. The closed-loop equation of the assumed model is $(AF - \mathcal{B}^k BD)Y_t = F\gamma + FC\epsilon_t$ (here, as in later sections we omit back shift operator for clarity) where the polynomials $A(\mathcal{B}), B(\mathcal{B})$ and $C(\mathcal{B})$ are obtained by inserting in the expressions (6) the polynomials $a(\mathcal{B}), b(\mathcal{B}), \phi(\mathcal{B})$ and $\theta(\mathcal{B})$ and the degree of integration d of the assumed model. Therefore, this equation can be written in terms of the BJ model polynomials as follows:

$$\begin{aligned} (1-\mathcal{B})^d\phi\left(Fa - \mathcal{B}^k bD\right) Y_t &= F\phi a\left[(1-\mathcal{B})^d\alpha + \delta\right] + F\theta a\epsilon_t \\ &\text{or} \\ \Phi(\mathcal{B})Y_t &= \xi + \Theta(\mathcal{B})\epsilon_t \end{aligned} \tag{9}$$

where $\Phi(\mathcal{B}) = 1 + \Phi_1\mathcal{B} + \ldots + \Phi_{n_\Phi}\mathcal{B}^{n_\Phi}$ and $\Theta(\mathcal{B}) = 1 + \Theta_1\mathcal{B} + \ldots + \Theta_{n_\Theta}\mathcal{B}^{n_\Theta}$. Note that for process identifiability the model must have $n_\Phi \geq \tilde{n}_\Phi$ and $n_\Theta \geq \tilde{n}_\Theta$. The second estimation step, therefore, consists of computing the estimates $\hat{\boldsymbol{\beta}}$ by minimizing the sum of $m = n_\Phi + n_\Theta$ squared errors obtained by comparing the coefficients of like powers of $\mathcal{B}$ in (9) and in (8):

$$Q(\boldsymbol{\beta}) = \frac{1}{2}\sum_{i=1}^{m} e_i^2(\boldsymbol{\beta}) = \frac{1}{2}\sum_{i=1}^{n_\Phi}(\Phi_i - \tilde{\Phi}_i)^2 + \frac{1}{2}\sum_{i=1}^{n_\Theta}(\Theta_i - \tilde{\Theta}_i)^2. \tag{10}$$

Deg of integ, d	Controller	AR part, $\Phi(\mathcal{B})$	Intercept, ξ	MA part, $\Theta(\mathcal{B})$
0	P	$\phi(a - \mathcal{B}^k b K_p)$	$\phi a \alpha$	θa
1	PI	$\phi\left[(1-\mathcal{B})a - \mathcal{B}^k b(c_1 + c_2\mathcal{B})\right]$	$\phi a \delta$	θa
0	PI	$\phi\left[(1-\mathcal{B})a - \mathcal{B}^k b(c_1 + c_2\mathcal{B})\right]$	0	$(1-\mathcal{B})\theta a$

Table 1: Closed-loop equations with PI controllers

In this study we limit our attention to processes controlled by proportional-integral (PI) controllers. In controller (7), P control is obtained with $F(\mathcal{B}) = 1$ and $D(\mathcal{B}) = c_1$, and PI control is obtained with $F(\mathcal{B}) = 1 - \mathcal{B}$ and $D(\mathcal{B}) = c_1 + c_2\mathcal{B}$, where c_1 and c_2 are the controller constants. It will be assumed that under the actions of the feedback controller, the process output is stabilized. Therefore, for the disturbance in (1), we consider models with $d = 0$ for a P controlled process and $d \in \{0, 1\}$ for a PI controlled process (higher values of d would yield a non-stationary behavior for these controllers).

Considering the possible cases we can have with PI controllers, the closed-loop model (9) can be further simplified as shown in Table 1. Notice that with $d = 0$ and a PI controller, there is a unit root in the MA part of the closed-loop equation. This fact can be exploited to determine the correct value of $\tilde{d}$ by looking at the MA part $\tilde{\Theta}(\mathcal{B})$ of the estimated closed-loop equation of the process. Statistical tests that have been developed to detect over-differencing in ARIMA model building can be used to determine if $\tilde{\Theta}(\mathcal{B})$ contains a unit root (Saikkonen and Luukkonen, 1993).

After the value of $\tilde{d}$ is determined and by knowing the controller, the appropriate model closed-loop equation in Table 1 is equated to the true process closed-loop equation (8) to compute the sum of squared errors (10) and the minimization of $Q(\boldsymbol{\beta})$ yields the parameter estimates $\hat{\boldsymbol{\beta}}$. The least squares parameter estimation procedure is presented next.

2.2 Least squares estimation of the process model

The parameters in $\boldsymbol{\beta}$, as defined in (3), are estimated by minimizing the sum of squared errors function (10). After solving for $\boldsymbol{\beta}$, the remaining parameters δ and α can be estimated from the intercept terms of the equations (9) and (8), as they do not appear in any of the other coefficients.

Unconstrained non-linear least squares

We refer to estimation without any prior knowledge on the process parameters as the unconstrained case. It can be noticed that the error terms $e_i(\boldsymbol{\beta})$ in a BJ model are nonlinear in $\boldsymbol{\beta}$

due to the presence of the product terms $\phi a, \phi b$, and θa in (9). The solution $\hat{\boldsymbol{\beta}}$ that minimizes (10) can be obtained by solving for the $(m \times 1)$ vector of errors $\boldsymbol{e}(\boldsymbol{\beta})$ using non-linear least squares techniques. A numerical approach to this is Newton's method, used to iteratively solve for the linearized form of the error vector (Myers, 1990, pp. 426-428). Let $\boldsymbol{\beta}^{(j)}$ denote the solution that was obtained in the jth iteration. In Newton's method we represent $\boldsymbol{e}(\boldsymbol{\beta})$ using a first order Taylor series around this solution, namely,

$$\boldsymbol{e}(\boldsymbol{\beta}) \doteq \boldsymbol{e}(\boldsymbol{\beta}^{(j)}) + \nabla \boldsymbol{e}(\boldsymbol{\beta}^{(j)})'(\boldsymbol{\beta} - \boldsymbol{\beta}^{(j)}) \tag{11}$$

where $\nabla \boldsymbol{e}(\boldsymbol{\beta}) = \left[\begin{array}{ccc} \nabla e_1(\boldsymbol{\beta}) & \ldots & \nabla e_m(\boldsymbol{\beta}) \end{array} \right]$ is the Jacobian matrix of $\boldsymbol{e}(\boldsymbol{\beta})$. The solution $\boldsymbol{\beta}^{(j+1)}$ at the next iteration is obtained by setting $\boldsymbol{e}(\boldsymbol{\beta})$ in (11) equal to $\mathbf{0}$ and solving the resulting linear system for $\boldsymbol{\beta}$. Note that by defining $\boldsymbol{y} = \boldsymbol{e}(\boldsymbol{\beta}^{(j)})$,$\boldsymbol{X} = -\nabla \boldsymbol{e}(\boldsymbol{\beta}^{(j)})'$ and $\tilde{\boldsymbol{\beta}}^{(j)} = \boldsymbol{\beta} - \boldsymbol{\beta}^{(j)}$, this linear system can be represented as a multiple linear regression model:

$$\boldsymbol{y} = \boldsymbol{X}\tilde{\boldsymbol{\beta}}^{(j)} + \boldsymbol{\varepsilon}^{(j)} \tag{12}$$

where $\boldsymbol{\varepsilon}^{(j)}$ is the model error. Thus, $\boldsymbol{\beta}^{(j+1)}$ that minimizes the squared norm of $\boldsymbol{\varepsilon}^{(j)}$ is obtained from the ordinary least squares solution $\tilde{\boldsymbol{\beta}}^{(j)} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\beta} - \boldsymbol{\beta}^{(j)}$, or solving for $\boldsymbol{\beta}$ and making it equal to $\boldsymbol{\beta}^{(j+1)}$

$$\boldsymbol{\beta}^{(j+1)} = \boldsymbol{\beta}^{(j)} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}. \tag{13}$$

Starting with an arbitrary vector of estimates, $\boldsymbol{\beta}^{(0)}$, the iterative procedure is continued using (13) until convergence is achieved. The solution is repeated multiple times, each time with a different starting point randomly selected from the parameter space, in an attempt to obtain the global minimum.

Constrained non-linear least squares (CLS) approach

Prior information on the process parameters is incorporated in the parameter estimation procedure as constraints. Prior information that can be specified as identities for functions of the parameters can be represented as equality constraints on $\boldsymbol{\beta}$:

$$\nu_i(\boldsymbol{\beta}) = 0 \quad i = 1, 2, \ldots, r. \tag{14}$$

Prior information that can be specified as ranges on functions of the parameters can be represented as inequality constraints on $\boldsymbol{\beta}$:

$$g_i(\boldsymbol{\beta}) \leq 0 \quad i = 1, 2, \ldots, s. \tag{15}$$

Point estimates of the functions of the parameters that are obtained from historical data can be incorporated as equality constraints. However, precise estimation of the parameters is not always available in practice. Estimation error can be taken into account through the use of linear inequality constraints that represent confidence intervals on the true values of the parameters.

In the case of equality constraints only, the parameter estimates can be obtained by minimizing a "constrained" sum of squares function :

$$Q_c(\boldsymbol{\beta}) = Q(\boldsymbol{\beta}) + \frac{1}{2}\sum_{i=1}^{r} \nu_i^2(\boldsymbol{\beta}) \tag{16}$$

by using Newton's method. We remark that this solution works regardless of whether the equality constraints (14) are linear or nonlinear functions of the parameters.

In the case of inequality constraints only, the parameter estimates can be obtained by minimizing the squared norm of the model error of the linear system (12) subject to the constraints (15). This is done by defining an objective function $Q_L^{(j)} = ||\boldsymbol{\varepsilon}^{(j)}||^2 = ||\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}^{(j)}||^2$ and solving the minimization:

$$\begin{aligned} \min_{\tilde{\boldsymbol{\beta}}^{(j)}} \quad & Q_L^{(j)} \\ \text{s.t.} \quad & g_i(\tilde{\boldsymbol{\beta}}^{(j)}) \leq 0 \;\; i = 1, 2, \ldots, s \end{aligned} \tag{17}$$

for each iteration j until convergence is achieved. Since the objective function is quadratic in $\tilde{\boldsymbol{\beta}}^{(j)}$ this can be solved using quadratic programming techniques when the inequality constraints are linear in $\tilde{\boldsymbol{\beta}}^{(j)}$ (Bazaraa et. al., 1993). Similarly to Newton's method, this procedure is also repeated from several starting points. In the case we have a mix of equality and inequality constraints we minimize (16) subject to (15). Henceforth, we will refer to this approach as the constrained least squares (CLS) approach.

Bayesian methods are also commonly employed for incorporating prior information in parameter estimation by modeling the prior information in the form of probability density functions. A Bayesian analysis of the ARMA model (9) requires sampling from non-standard posterior distributions (see e.g. Marriott et. al., 1996). In addition, in order to incorporate prior information about functions of parameters the covariance of these parameters must also be properly modeled. In contrast, a constrained least squares approach can incorporate a relatively wide class of prior knowledge while requiring a considerably less computational effort.

2.3 Identifiability conditions

A process $\mathcal{S}$ is said to be *parameter identifiable* if $\hat{\boldsymbol{\beta}}$ estimated by minimizing $Q(\boldsymbol{\beta})$ is unique and $\mathcal{M}(\hat{\boldsymbol{\beta}})$ corresponds to $\mathcal{S}$ (Ljung et. al. 1974). It has been shown that parameter

identifiability in closed-loop can be achieved in several ways: i) applying an independent signal in the controller (e.g. a time varying set-point or a dither signal) that is persistently exciting of sufficient order (Söderström et. al., 1976); ii) switching between different linear controllers (Ljung et. al., 1974); and iii) having a complex enough controller in the feedback loop (Söderström et. al., 1975).

Söderström et. al. (1975) derived the following necessary and sufficient parameter identifiability condition for a true process $\mathcal{S}$ that obeys an ARMAX model form (5) and that is controlled by the controller (7):

$$\max\{n_F - \tilde{n}_B, n_D + \tilde{k} - \tilde{n}_A\} - \tilde{n}_P \geq 0 \tag{18}$$

where $\tilde{n}_P$ is the degree of the polynomial $\tilde{P}$, which is defined as the common factor between $\tilde{A}F - \mathcal{B}^{\tilde{k}}\tilde{B}D$ and $\tilde{C}$. It is evident that this is a condition on both the process and the controller. If (18) is not satisfied, it is not possible to uniquely identify the true process parameters regardless of the model structure assumed. Condition (18) can be rewritten as:

$$\max\{n_F + \tilde{n}_A, n_D + \tilde{k} + \tilde{n}_B\} - \tilde{n}_P \geq \tilde{n}_A + \tilde{n}_B$$

which more clearly states that in the closed-loop model (8) the number of linearly independent equations from the AR part $\tilde{A}F - \mathcal{B}^{\tilde{k}}\tilde{B}D$, after cancelling the common factors, must be at least equal to the number of unknowns. Note that the MA part $F\tilde{C}$ does not enter the identifiability condition unless it has a common factor with the AR part.

It has been proved that if the process satisfies the identifiability conditions *and* the model structure is general enough to contain the true process representation, i.e. $\mathcal{S} \in \mathcal{M}(\boldsymbol{\beta})$, then the true open-loop representation $\mathcal{S}$ can be obtained in the limit as the sample size N tends to infinity (Ljung et. al. 1974, Ljung 1978).

The parameter identifiability condition for a process that obeys a BJ model form can be derived as follows. From the polynomials (6) written for the true process we have that $\tilde{n}_A = \tilde{d} + \tilde{n}_\phi + \tilde{n}_a$, $\tilde{n}_B = \tilde{d} + \tilde{n}_\phi + \tilde{n}_b$, and $\tilde{n}_C = \tilde{n}_\theta + \tilde{n}_a$. Furthermore, for the controller (7), we have $n_D = n_F = 0$ for a P controller and $n_D = n_F = 1$ for a PI controller. Therefore, for the different controller and disturbance cases, we can write the condition (18) for a BJ model form as:

$$\begin{array}{lll} \tilde{d} = 0 \quad \text{and P controller:} & \max\{\tilde{n}_a, \tilde{k} + \tilde{n}_b\} - \tilde{n}_p & \geq \tilde{n}_s \\ \tilde{d} = 1 \quad \text{and PI controller:} & \max\{\tilde{n}_a, I_{PI} + \tilde{k} + \tilde{n}_b - 1\} - \tilde{n}_p & \geq \tilde{n}_s \\ \tilde{d} = 0 \quad \text{and PI controller:} & \max\{1 + \tilde{n}_a, I_{PI} + \tilde{k} + \tilde{n}_b\} - \tilde{n}_p & \geq \tilde{n}_s \end{array} \tag{19}$$

where $\tilde{n}_s = \tilde{n}_\phi + \tilde{n}_a + \tilde{n}_b$ and $I_{PI} = 1$ with PI control and $I_{PI} = 0$ with I control. $\tilde{P}$ is defined in the same manner as in (18), that is, as the polynomial common to $\tilde{A}F - \mathcal{B}^{\tilde{k}}\tilde{B}D$ and $\tilde{C}$.

3 Some constraint types and their specification

In this section we review some possible forms of process information that are commonly available in practice and show how they can be represented as equality and inequality constraints as discussed in section 2.2.

One type of prior knowledge that can be expressed as an equality constraint of the form (14) is the asymptotic gain of the transfer function model (1). The asymptotic gain, denoted by $\tilde{g}$, is equal to Y_∞ when $U_t = 1$ for all t. For this model, the asymptotic gain can be written by using the final value theorem of $\mathcal{Z}$ transforms and the algebraic equivalence between the back shift operator and the $\mathcal{Z}$ transform (i.e. $\mathcal{B} \equiv z^{-1}$) as follows:

$$\tilde{g} = Y_\infty = \lim_{t\to\infty} Y_t = \lim_{\mathcal{B}\to 1} \frac{\mathcal{B}^{\tilde{k}}\tilde{b}(\mathcal{B})}{\tilde{a}(\mathcal{B})}. \tag{20}$$

In practice, this limit operation implies that one must wait long enough for the process to settle in order to observe the steady-state response Y_∞. This is common practice in analysis methods that study steady-state effects, such as design of experiment (DOE) techniques. The asymptotic gain of a dynamic system can therefore be estimated by fitting a regression model to the steady-state input and output data of the process; the estimated regressor coefficient corresponds to the gain of the process.

As a numerical example, consider the transfer function $Y_t = \frac{5\mathcal{B}}{1+0.5\mathcal{B}}U_t$. By letting $\mathcal{B} \to 1$, the asymptotic gain is calculated as:

$$\tilde{g} = \frac{5(1)}{1+0.5(1)} = 3.33.$$

Another characteristic of a process that is often known from prior experience with the process is the input-output delay. The process delay $\tilde{k}$ is defined as the number of whole time periods during which no change in the output is observed in response to a change in the input. The delay can be measured by off-line step-response experiments (see, e.g., Björklund and Ljung, 2003). The knowledge of the delay can be easily incorporated in identification by simply writing the assumed model with the correct delay.

Knowledge of the disturbance model can also be available from observing the process in open-loop provided that this is feasible and not too expensive over the period of time where open-loop observations are needed (usually 75 to 100 observations are needed for fitting time series models, see Box et. al., 1994). In a BJ model, the response measured in open-loop, that is, when the input is a constant or zero, is completely due to the disturbance process. For instance, in the case when $U_t = 0$, from model (1) we have the response as $Y_t = \frac{\theta(\mathcal{B})}{\phi(\mathcal{B})(1-\mathcal{B})^d}\epsilon_t$ which is identical to the disturbance process. The knowledge on the true values of the coefficients of $\phi(\mathcal{B})$ and $\theta(\mathcal{B})$ can be incorporated in the identification procedure by writing an equality constraint for each coefficient in these polynomials.

The usual stationarity and invertibility conditions of the disturbance model, i.e. that the roots of $\phi(\mathcal{B})$ and $\theta(\mathcal{B})$ must lie outside the unit circle, and the stability condition of the dynamics transfer function, i.e. that the roots of $a(\mathcal{B})$ must lie outside the unit circle, can be formulated as inequality constraints. These are standard assumptions of ARMA and transfer function models that usually hold in practice. Assuming an ARMA(2,2) disturbance model in (1), the stationarity conditions are $-1 \leq \phi_2 \leq 1$, $\phi_1 + \phi_2 \leq 1$ and $-\phi_1 + \phi_2 \leq 1$; and the invertibility conditions are $-1 \leq \theta_2 \leq 1$, $\theta_1 + \theta_2 \leq 1$ and $-\theta_1 + \theta_2 \leq 1$. Assuming a second order transfer function denominator polynomial, the stability conditions are $-1 \leq a_2 \leq 1$, $a_1 + a_2 \leq 1$ and $-a_1 + a_2 \leq 1$.

4 Effects of adding constraints on identifiability

As defined in Section 2.3, a process $\mathcal{S}$ is parameter identifiable if $\hat{\boldsymbol{\beta}}$ estimated by minimizing $Q(\boldsymbol{\beta})$ is unique and $\mathcal{M}(\hat{\boldsymbol{\beta}})$ corresponds to $\mathcal{S}$ (Ljung et. al. 1974). In this section we investigate the effects of introducing equality constraints on the identifiability of the process by looking at the relation between the convexity properties of the sum of squares function $Q(\boldsymbol{\beta})$ and the parameter identifiability.

Consider the true process

$$(1 + 0.5\mathcal{B})Y_t = 5\mathcal{B}U_t + \epsilon_t \tag{21}$$

which is controlled by a P controller $U_t = K_p Y_t$ where $K_p = 0.5$. The true process closed-loop equation (8) is obtained by inserting the controller equation in (21); it is equal to:

$$(1 + 0.25\mathcal{B})Y_t = \epsilon_t. \tag{22}$$

To identify the process we assume a model in the same form as (21), therefore the parameter vector is $\boldsymbol{\beta} = (a_1, b_1)$ and the delay is $k = 0$. The assumed model closed-loop equation is obtained by inserting the controller equation in the assumed model; it is equal to:

$$(1 + \Phi_1\mathcal{B})Y_t = \epsilon_t \tag{23}$$

where $\Phi_1 = a_1 - K_p b_1$.

It can be seen that (21) is an ARMAX model form because it corresponds to (1) with same denominator polynomials in the transfer function and the disturbance; that is $a(\mathcal{B}) = \phi(\mathcal{B})(1-\mathcal{B})^d = (1+0.5\mathcal{B})$. For a process that obeys an ARMAX model we have that, in the closed-loop equation (9), the $\phi(\mathcal{B})(1-\mathcal{B})^d$ and $a(\mathcal{B})$ terms of the AR and MA polynomials cancel and we have

$$\Phi(\mathcal{B}) = F(\mathcal{B})a(\mathcal{B}) - B^k b(\mathcal{B})D(\mathcal{B}) \text{ and } \Theta(\mathcal{B}) = F(\mathcal{B})\theta(\mathcal{B})$$

which have all coefficients linear in the parameters. Therefore, the sum of squared errors function (10) is a quadratic (convex) function of $\boldsymbol{\beta}$.

For this process $\tilde{n}_A = 1, \tilde{n}_B = 1, \tilde{n}_C = 0, n_F = 0, n_D = 0$ and $\tilde{k} = 0$, and by applying (18) we see that $\max\{1, 1\} = 1 \not\geq 2$, thus, the process is not parameter identifiable. An alternative way to check this is as follows. Since parameter identifiability is equivalent to the existence of a unique minimizing solution to $Q(\boldsymbol{\beta})$, a sufficient condition for parameter identifiability of an ARMAX process is that $Q(\boldsymbol{\beta})$ is strictly convex. This can be checked by an eigenvalue test for positive definiteness of the Hessian matrix $H = \left(\frac{\partial^2 Q}{\partial\beta_i \partial\beta_j}\right)\ i, j = 1, ..., p$. For positive definiteness the eigenvalues λ_i of H must be $\lambda_i > 0$ for $i = 1, ..., p$.

The sum of squared errors function that is computed by comparing (22) and (23) is $Q(\boldsymbol{\beta}) = \frac{1}{2}e_1^2 = \frac{1}{2}(a_1 - K_p b_1 - 0.25)^2$ which has the Hessian matrix:

$$H = \begin{bmatrix} 1 & -K_p \\ -K_p & K_p^2 \end{bmatrix}.$$

The eigenvalues of H are $\lambda_1 = 0, \lambda_2 = 1 + K_p$. Thus, $Q(\boldsymbol{\beta})$ is not strictly convex, which implies that there are infinitely many minimizing solutions, and the process is not parameter identifiable.

To study the effect of adding a constraint, suppose that we know the gain $\tilde{g} = \frac{5}{1+0.5} = 3.33$ of the true process. This can be written as a constraint on a_1 and b_1 as $\frac{b_1}{1+a_1} = 3.33$ or in the form of an equality constraint (14) as $\nu = 3.33 + 3.33a_1 - b_1$. The Hessian matrix of the constrained sum of squared errors function (16) is:

$$H_c = \begin{bmatrix} 1 + 3.33^2 & -K_p - 3.33 \\ -K_p - 3.33 & K_p^2 + 1 \end{bmatrix} \tag{24}$$

which has the characteristic equation $\lambda^2 - \lambda(K_p^2 + 13.089) + (1 - 3.33K_p)^2 = 0$. The process is parameter identifiable as long as the eigenvalues are $\lambda_1, \lambda_2 > 0$, which is satisfied when the constant term $(1 - 3.33K_p)^2$ is non-zero, or when $K_p \neq \frac{1}{3.33} = 0.3$.

$Q(\boldsymbol{\beta})$ and $Q_c(\boldsymbol{\beta})$ are plotted as a function of a_1 and b_1 in Figure 1 (logarithmic scale is shown). As it can be seen, $Q(\boldsymbol{\beta})$ has a locus of minimizing solutions along a straight line which give equivalent objective function values. The line includes the true solution $a_1 = 0.5, b_1 = 5$, however, the solution is not unique. This illustrates an instance of a non-parameter identifiable process. As the result of adding the constraint, the sum of squared errors function becomes strictly convex which results in a unique minimizing solution that corresponds to the true process (plot on the right, Figure 1). Therefore, the process after adding the constraint is parameter identifiable.

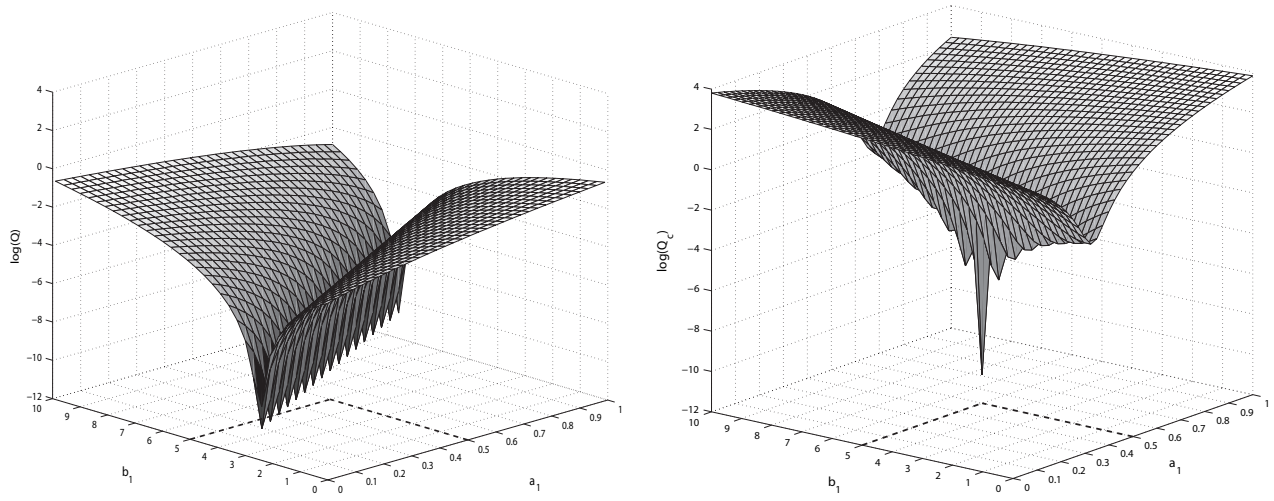

Figure 1: Effect of adding a constraint on the identifiability of the process model given in equation (21). Graphs show the logarithm of the sum of squared error surfaces: unconstrained case (left); constrained with gain information (right). Without any constraint, the surface has infinitely many alternative minima. Knowing the gain, the surface has a unique minimum which corresponds to the true process.

5 Simulation Example: Small sample benefits of adding constraints

In this section we apply the CLS approach on simulated process data to study the small sample properties of the parameter estimates obtained by using constraints. We consider the process (this is the "true" process we simulate):

$$Y_t = \frac{0.1}{1-\mathcal{B}} + \frac{10\mathcal{B}}{1-0.8\mathcal{B}}\mathcal{B}U_t + \frac{1-0.3\mathcal{B}}{1-\mathcal{B}}\epsilon_t \tag{25}$$

which is controlled with a PI controller $U_t = \frac{c_1+c_2\mathcal{B}}{1-\mathcal{B}}Y_t$ where $c_1 = 0.005$, $c_2 = -0.01$. The white noise process is $\epsilon_t \overset{iid}{\sim} N(0,1)$.

This process obeys the BJ model form (4) with delay $\tilde{k} = 1$, degree of integration $\tilde{d} = 1$, the drift constant $\tilde{\delta} = 0.1$, the model polynomial orders $(\tilde{n}_a, \tilde{n}_b, \tilde{n}_\theta, \tilde{n}_\phi, \tilde{n}_P) = (1,1,1,0,0)$ and the PI controller polynomial orders $(n_D, n_F) = (1,1)$. Therefore, in the parameter identifiability condition (19) we have that $\max\{1,2\} \geq 2$, hence the parameter identifiability condition is satisfied by the process. The closed-loop equation (8) of the true process that corresponds to the given PI controller is:

$$(1 - 1.8\mathcal{B} + 0.75\mathcal{B}^2 + 0.1\mathcal{B}^3)Y_t = 0.02 + (1 - 1.1\mathcal{B} + 0.24\mathcal{B}^2)\epsilon_t \tag{26}$$

which is an ARMA(3,2) model.

The process is identified by assuming the Box-Jenkins model form (4). We first identify the degree of integration in the disturbance model. Since the closed-loop equation (26) does not have a unit root in the MA polynomial and a PI controller was used, we determine (correctly) the degree of integration as $d = 1$ (see Table 1).

The form of the true process model is unknown, however, in selecting the assumed model we suppose that some prior knowledge on the true model is available in order to guarantee that the true process is contained in the assumed model. In particular we consider:

- a small model: $Y_t = \frac{\delta}{1-\mathcal{B}} + \frac{b_1\mathcal{B}+b_2\mathcal{B}^2}{1-a_1\mathcal{B}}\mathcal{B}^k U_t + \frac{1-\theta_1\mathcal{B}}{(1-\mathcal{B})(1-\phi_1\mathcal{B})}\epsilon_t$ and $k = 0$, and

- a large model: $Y_t = \frac{\delta}{1-\mathcal{B}} + \frac{b_1\mathcal{B}+b_2\mathcal{B}^2+b_3\mathcal{B}^3+b_4\mathcal{B}^4}{1-a_1\mathcal{B}-a_2\mathcal{B}^2}\mathcal{B}^k U_t + \frac{1-\theta_1\mathcal{B}-\theta_2\mathcal{B}^2}{(1-\mathcal{B})(1-\phi_1\mathcal{B}-\phi_2\mathcal{B}^2)}\epsilon_t$ and $k = 0$

which reflect, respectively, the cases of a more certain and a less certain prior knowledge of the true model. We can see that both models have larger orders in each polynomial and a smaller delay than the true process (i.e. the true process is contained in both models), thus, the identifiability condition is satisfied by both models.

We considered the constraint types that were discussed in section 3. The symbols given in Table 2 will be used to refer to the constraints used in a particular simulation/estimation test. We incorporate these constraints, for example in the small model, as follows. For the delay, we set $k = 1$; for the gain we set $\frac{b_1+b_2}{1-a_1} = \tilde{g}$ where $\tilde{g} = \frac{10}{1-0.8} = 50$ is the true gain; and for the noise model we set $\theta_1 = 0.3$ and $\phi_1 = 0$. The constraint scenarios considered in the simulations are the single-constraint cases that are shown in Table 2 and the two-constraint cases (g, n), (k, g) and (k, n).

Constraint	Symbol
Unconstrained	u
Stationarity, invertibility, stability	st
Known asymptotic gain	g
Known input-output delay	k
Known noise model	n

Table 2: Constraints used in simulations

5.1 Estimation of the process model from the simulated realizations

The controlled operation of the process (25) was simulated under the given PI controller. In order to evaluate the effect of using different sample sizes, we generated $N = 50, 200$ and 400 observations in the simulation. The simulation was repeated 1000 times to obtain a distribution of the parameter estimates.

For each simulated realization, the closed-loop equation (26) was estimated by fitting an ARMA(3,2) model to the simulated output $\{Y_t\}_{t=1}^{N}$ which yielded the estimates of $\tilde{\Phi}(\mathcal{B}), \tilde{\Theta}(\mathcal{B})$ and $\tilde{\xi}$ (MATLAB's *armax* function was used for ARMA model fitting, see Ljung, 2002). These estimates were then used in the minimization of the sum of squares function (10) to obtain $\hat{\boldsymbol{\beta}}$. (Note that the closed-loop equation (9) that corresponds to the assumed models and a PI controller is an ARMA(5,3) for the small model and an ARMA(8,5) for the large model, thus $m = 8$ for the small model and $m = 13$ for the large model.)

In the scenarios with only equality constraints, the iterative Newton's method was applied to solve the nonlinear least squares problem; in the scenarios that include inequality constraints MATLAB's *quadprog* function was used to solve the iterative quadratic programming problem. An estimate $\hat{\boldsymbol{\beta}}$ is obtained by solving the problem several times with a different starting point, and by retaining the solution giving the smallest objective function value as the global minimum. In each run the solution is repeated from $100p$ different starting points.

5.2 Assessment of the benefits of using constraints

In this section we discuss the benefit of adding constraints by considering *i*) the bias and variance properties of the parameter estimates and *ii*) the control performance of the parameter estimates.

Bias and Variance of the Parameter Estimates

We define the squared bias of the random vector estimator $\hat{\boldsymbol{\beta}}$ as

$$(Bias\ \hat{\boldsymbol{\beta}})^2 = \sum_{i=1}^{p} (E\hat{\beta}_i - \beta_i)^2$$

that is, as the sum of the squared biases of its individual components. Similarly, we define the variance as the sum of the variances of its individual components:

$$Var\hat{\boldsymbol{\beta}} = \sum_{i=1}^{p} Var\hat{\beta}_i.$$

Note that this does not consider the covariance of the estimates. The magnitudes of the covariances were observed to be relatively small, however, and including them in the sum did not change the values significantly. Thus, to have a simpler and easier to interpret variance definition we did not consider the covariances.

The sample squared bias and variance of the simulated realizations of $\hat{\boldsymbol{\beta}}$ obtained with different constraints are plotted in Figure 2. A log scale was used to more clearly see the small values. It can be seen that if we use a smaller (more parsimonious) assumed model

or a larger sample size then the estimators under all constraints have smaller variance and bias.

The improvement attained in the variance and the bias of the unconstrained estimates by adding constraints can be summarized as follows. The knowledge of the delay is crucial regardless of the model size; it gives the greatest improvement in both the variance and the bias. The knowledge of the gain when the delay is also known (i.e. the constraints k, g) is better for reducing the variance and the bias when a small model is used; it gives moderate improvement with a large model. By contrast, the knowledge of the noise model is better for reducing the variance and the bias with a large model; with a small model it improves the bias, however, it inflates the variance. Utilizing the stationarity, invertibility and the stability constraints did not provide significant reduction in the bias or the variance for either of the models. Among all constraint scenarios, the knowledge of the delay and the gain (i.e. the constraints k, g) with a small model, and the knowledge of the delay and the noise model (i.e. the constraints k, n) with a large model give the best identified models that have the smallest variance and bias values.

The bias and variance results also suggest that one must exercise caution when adding constraints so that one does not overconstrain the system. When we have a more certain prior knowledge about the true model and use a small model, the unconstrained case already gives relatively good estimates and adding constraints can only inflate the variance as it was evidenced with the noise model constraint; a large model on the other hand benefits more from adding constraints and has less risk of being overconstrained.

Control Performance of the Parameter Estimates

The performance of an estimated process model from a control point of view can be measured by re-designing the controller according to this model in such a way that the process output achieves some optimality criterion. Here, we consider minimum mean squared error (MMSE) controllers as the optimal controller, although any other controller objective can be used, as desired. As the performance metric of the designed controller we consider the variance of the controlled process.

The MMSE control law for a process that obeys the BJ model (1) can be derived by considering the ARMAX form $A(\mathcal{B})Y_t = \mathcal{B}^{k+1}B'(\mathcal{B})U_t + C(\mathcal{B})\epsilon_t$ of this model. Here, $B'(\mathcal{B})$ is defined as $B'(\mathcal{B}) = (1-\mathcal{B})^d\phi(\mathcal{B})\frac{b(\mathcal{B})}{\mathcal{B}}$ and the $A(\mathcal{B})$ and $C(\mathcal{B})$ polynomials are defined in the same way as in (6). The MMSE control law is given as (Åström, 1970):

$$U_t^* = -\frac{H(\mathcal{B})}{B'(\mathcal{B})L(\mathcal{B})}Y_t. \tag{27}$$

The $H(\mathcal{B})$ and $L(\mathcal{B})$ polynomials are defined as $H(\mathcal{B}) = h_0 + h_1\mathcal{B} + ... + h_{n-1}\mathcal{B}^{n-1}$ and

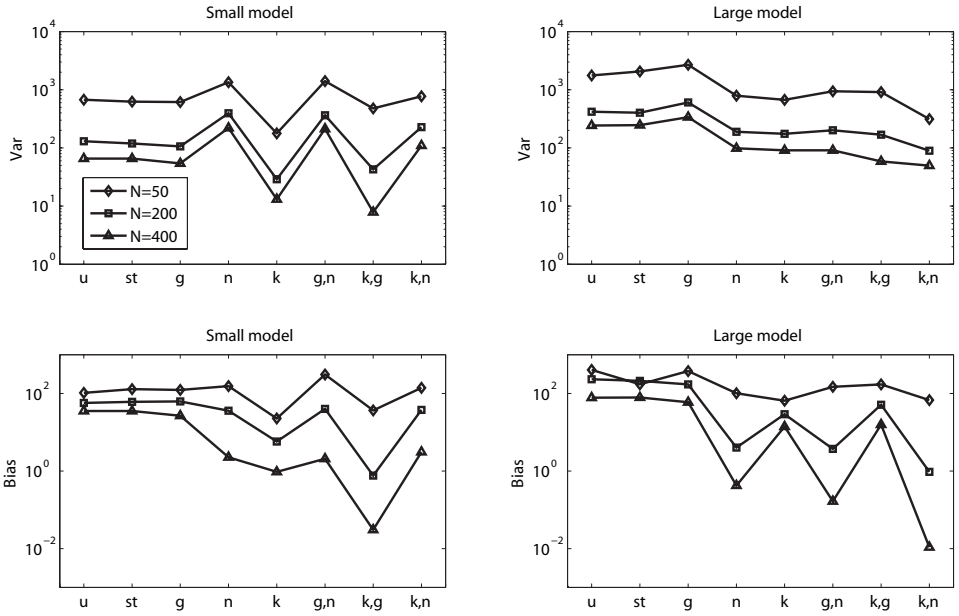

Figure 2: The sample squared bias and variance of the identified models under different constraints and sample sizes. A log scale is used. The bias and variance were estimated from the 1000 simulated realizations. **Left panels**: identification by using a small model, **right panels**: identification by using a large model.

$L(\mathcal{B}) = 1 + l_1\mathcal{B} + ... + l_k\mathcal{B}^k$ where $n = \max\{n_A, n_{B'}, n_C\}$ and their coefficients are obtained by solving the Diophantine identity

$$C(\mathcal{B}) = A(\mathcal{B})L(\mathcal{B}) + \mathcal{B}^{k+1}H(\mathcal{B}). \tag{28}$$

For each $\hat{\boldsymbol{\beta}}$ and the $\widehat{A(\mathcal{B})}, \widehat{B'(\mathcal{B})}$ and $\widehat{C(\mathcal{B})}$ polynomials estimated in the simulations, the estimated optimal controller $\widehat{U_t^*}$ can be computed, first, by solving (28) for $\widehat{H(\mathcal{B})}$ and $\widehat{L(\mathcal{B})}$ and, then, by inserting the estimated polynomials in expression (27). It can be shown that the true optimal controller, or the optimal controller computed from the true process (25), is $U_t^* = -\frac{0.7(1-0.8\mathcal{B})}{10(1-\mathcal{B})(1+0.7\mathcal{B})}Y_t$.

The variance of the process controlled by the true optimal controller can easily be calculated from the closed-loop equation of the process obtained with this controller; this is equal to $Y_t = (1+0.7\mathcal{B})\epsilon_t$ which is an MA(1) process with variance 1.49. The variance of the process controlled by the estimated optimal controller is calculated numerically. For each estimated controller $\widehat{U_t^*}$ we simulated a new realization of the process (25) under the actions of this controller and calculated the sample variance of this simulated output. A sample size of 200 was used for the variance calculations.

The estimated optimal controllers and the corresponding controlled process variances were computed for the parameter estimates obtained under the constraint scenarios that include the delay constraint and under the unconstrained case. The benefit of adding constraints is evaluated by comparing the variance provided by the estimated optimal controller to the variance value 1.49, provided by the true optimal controller. It is assumed that the controllers that provide variances in the range (0,3) are approximately equal to the true optimal controller and hence they are "good" controllers; and that the controllers that provide variances in the range $(100, \infty)$ are "unstable". Figure 3 gives the percentages of the estimated controllers (out of 1000) that have good control performance and those that are unstable.

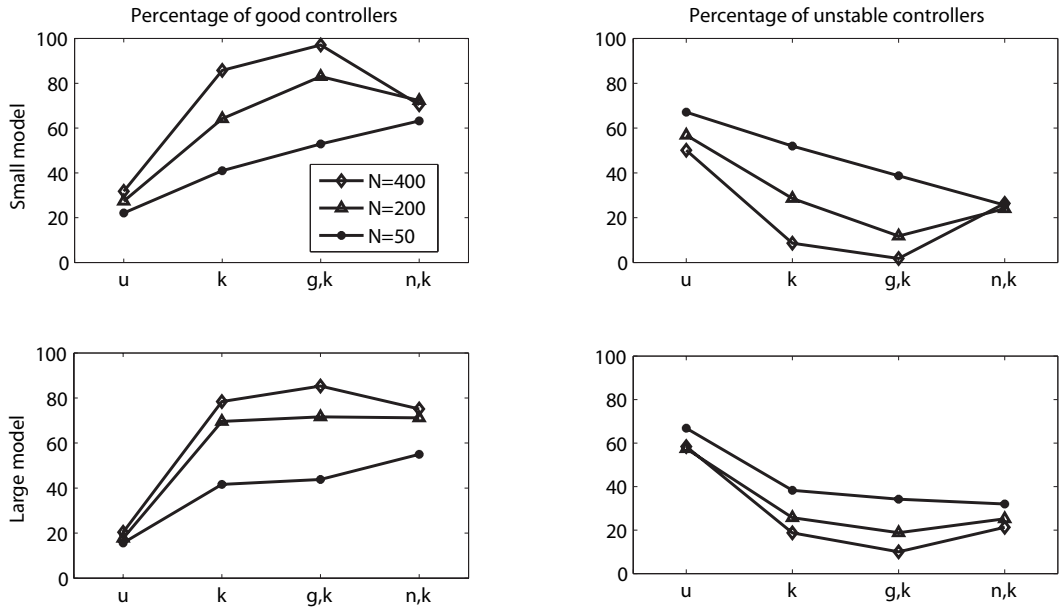

Figure 3: Percentages of the estimated controllers that have good control performance (left panels) and those that are unstable (right panels). The estimated controllers were computed from the models identified in the 1000 simulation runs under different constraints and sample sizes. The labels of the constraints are given in the $x-$ axis of the plots. **Top panels**: identification by using a small model, **bottom panels**: identification by using a large model.

It can be seen that by having a more certain prior knowledge on the true model (i.e. using a smaller assumed model) or by using a larger sample size we can improve the control performance. We can also see the benefits of adding constraints. With the knowledge of the delay the control performance can be significantly improved with respect to the unconstrained case regardless of the sample size and the model size. With the knowledge of the gain, the improvement in the control performance is larger when a relatively large sample size is used

(e.g. when $N = 200$ or 400) and the knowledge of the noise model is more advantageous when a relatively small sample size is used (e.g. when $N = 50$).

Among all constraint scenarios, the best control performance is obtained with the knowledge of the delay and the gain (i.e. the constraints k, g) when the sample size is relatively large and with the knowledge of the delay and the noise model (i.e. the constraints k, n) when the sample size is relatively small regardless of the model size used.

6 Case Study: Gas Furnace Process

In this section we compare the performance of the proposed approach and the dither signal approach by applying them on the gas furnace data set given by Box, Jenkins and Reinsel (1994).

The data set contains 296 input and output observations which are collected during open-loop operation of the process. The process input is the gas rate of a gas furnace and the output is the $\%CO_2$ concentration at the furnace outlet. The transfer function model estimated from these open-loop data is:

$$Y_t = \frac{-0.4976\mathcal{B} - 0.39156\mathcal{B}^2 - 0.56424\mathcal{B}^3}{1 - 0.51213\mathcal{B}}\mathcal{B}^2 U_t + \frac{1}{(1 - 0.54505\mathcal{B})(1 - \mathcal{B})}\epsilon_t \qquad (29)$$

and $\hat{\sigma}_\epsilon^2 = 0.06398$.

In order to illustrate closed-loop identification we need a data set for the closed-loop operation of the process. We generate the closed-loop operating data using the reconstructed disturbance data of the process and by assuming that a PI controller was in operation. To reconstruct the disturbance data from the open-loop input and output measurements we assume that the model (29) is the true description of the process; that is, (29) gives the polynomials $\tilde{b}(\mathcal{B})$, $\tilde{a}(\mathcal{B})$, $\tilde{\theta}(\mathcal{B})$, $\tilde{\phi}(\mathcal{B})$, the delay $\tilde{k}$, and the degree of integration $\tilde{d}$ of the process. Denoting the open-loop input and output observations by U_{ot} and Y_{ot}, the disturbance data can be obtained by using $N_t = Y_{ot} - \frac{\tilde{b}(\mathcal{B})}{\tilde{a}(\mathcal{B})}\mathcal{B}^2 U_{ot}$.

For generating the closed-loop output of the process we assumed that the transfer function numerator polynomial in (29) is $\tilde{b}(\mathcal{B}) = -0.4976\mathcal{B}$ (i.e. the higher order terms were ignored). This simplification is made so that the resulting closed-loop model has a lower order and thus is easier to estimate. From the disturbance data, the controlled output sequence $\{Y_t\}_{t=1}^{296}$ was generated by using $Y_t = \frac{\tilde{b}(\mathcal{B})}{\tilde{a}(\mathcal{B})}\mathcal{B}^2 U_t + N_t$ where $\{U_t\}$ is the input sequence according to the PI control law $U_t = \frac{c_1 + c_2\mathcal{B}}{1-\mathcal{B}} Y_t'$. Y_t' is the deviation of the output from target T (i.e. $Y_t' = Y_t - T$) where the target is assumed to be 53 (the average of the Y_{ot} sequence).

Two PI controllers were used; the first one is with $(c_1, c_2) = (1, -0.7)$ and the second one is with $(c_1, c_2) = (0.05, -0.03)$. The first one, with its larger proportional and integral

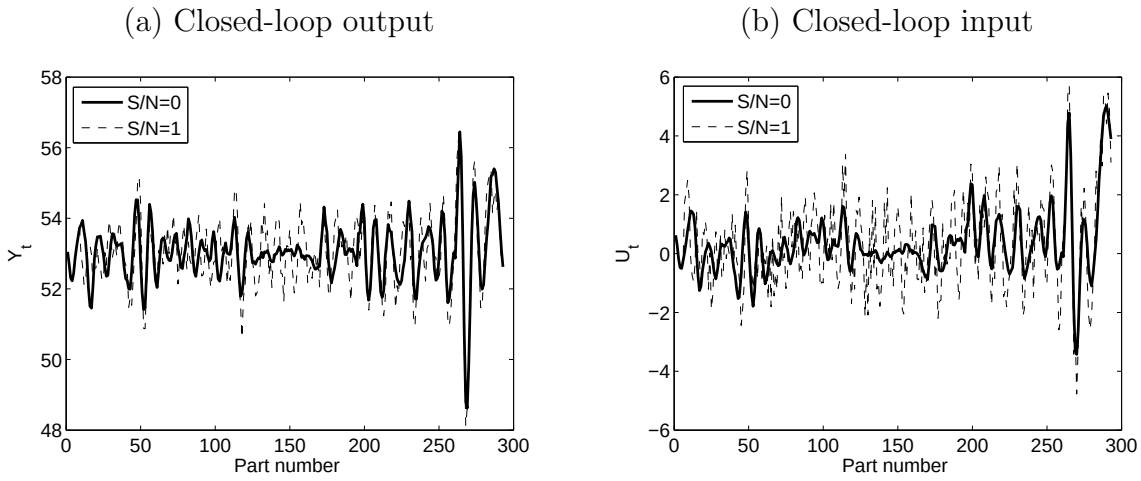

Figure 4: Closed-loop identification of the gas furnace process data set. (a) Closed-loop output under the first PI controller including an added dither signal with $S/N = 0$ (i.e. no dither applied) and $S/N = 1$. (b) PI controller actions including an added dither signal; with $S/N = 0$ and $S/N = 1$.

constants, varies the input more frequently and hence would result in more precise estimates of the closed-loop model than the second one. Figure 4 shows the closed-loop output and input data generated with the first PI controller. This figure also shows the controlled input and output values with dither signals (i.e. the data that correspond to $S/N \neq 0$). The dither signal approach will be explained later.

The parameter identifiability of the process is checked by using the condition in (19) that corresponds to $\tilde{d} = 1$ and a PI controller; that is: $\max\{\tilde{n}_a, \tilde{k} + \tilde{n}_b\} - \tilde{n}_p \geq \tilde{n}_\phi + \tilde{n}_a + \tilde{n}_b$. It can be seen that for this process $\tilde{n}_a = 1, \tilde{n}_b = 1, \tilde{k} = 2, \tilde{n}_\phi = 1, \tilde{n}_\theta = 0$ and $\tilde{n}_p = 0$ and that this condition is satisfied.

Closed-loop identification using the CLS approach

It can be shown that the closed-loop equation of the true process (29) with a PI controller is an ARMA(5,1). From the generated output data $\{Y_t\}_{t=1}^{296}$ shown in figure 4 (a) we have (correctly) identified this model. The ARMA(5,1) model fitted to this data set is:

$$\begin{array}{ccccccccccc} (1 & -1.9515\mathcal{B} & +1.2099\mathcal{B}^2 & +0.4557\mathcal{B}^3 & -0.8476\mathcal{B}^4 & +0.3341\mathcal{B}^5 &)Y_t' = (1 & -0.6526\mathcal{B} &)\epsilon_t \\ & (0.1618) & (0.2254) & (0.1592) & (0.1406) & (0.0623) & & (0.1659) & \end{array} \tag{30}$$

where all estimates are highly significant (the numbers in the parentheses are the standard errors of the estimates). An ARMA(5,1) model was also fitted to the closed-loop data of

the second PI controller; in this model the standard errors of the estimates were relatively larger.

We assumed a BJ model (1) with parameter vector $\boldsymbol{\beta} = (a_1, b_1, b_2, b_3, \phi_1, \theta_1)$ and delay $k = 0$ for the identification and estimation of the process model. The degree of integration of the disturbance is identified as $d = 1$ since the MA part of (30) does not contain a unit root and the controller is PI. The parameters $\boldsymbol{\beta}$ of the model are estimated by minimizing the sum of squared errors function (10) that is evaluated with the closed-loop model parameter estimates given in (30). This corresponds to the unconstrained closed-loop identification.

To illustrate the CLS approach, we assume knowledge of the delay, the gain and the noise model of the process. The true values for these constraints are $k = 2$ for the delay, $g = -0.4976/(1 - 0.51213) = -1.0199$ for the gain, and $\phi_1 = 0.54505$ and $\theta_1 = 0$ for the noise model. Similarly to Section 5, we use the symbols k, g and n to refer to the constraints. The constraint scenarios $k, (k, g), (k, n)$ and (k, n, g) were considered.

Closed-loop identification using the dither signal approach

In closed-loop identification with dither signals an external dither signal d_t is added to the feedback input. Provided that the added dither is signal is uncorrelated with ϵ_t and is persistently exciting the system, the transfer function and the disturbance models can be identified from the estimated cross-correlation function between the input and the output sequences. This approach is sometimes referred to as the "direct" method for closed-loop identification (Ljung, 1974).

After the addition of the dither signal, the input adjustment of the PI controller is:

$$\nabla U_t = (c_1 + c_2\mathcal{B})Y_t' + \nabla d_t$$

where d_t is a pseudo random binary sequence (PRBS). The signal to noise ratio, or S/N, with the dither signal is calculated as $S/N = var(d_t)/var(N_t)$. We considered the signal-to-noise ratios $S/N = 0.1, 1$ and 5. For identification under each S/N, we added the dither d_t to the actions of the PI controllers and generated the corresponding closed-loop input and output sequences $\{U_t\}$ and $\{Y_t\}$.

Similarly to the CLS approach, we assumed a BJ model (1) with a parameter vector $\boldsymbol{\beta} = (a_1, b_1, b_2, b_3, \phi_1, \theta_1)$, a delay $k = 0$ and a degree of integration $d = 1$ to identify and estimate the process model. The process transfer function model was estimated from the generated closed-loop output and input data with the added dither signals (by using the Proc ARIMA function in SAS). Figure 4 shows the generated closed-loop output and input data using the first PI controller and a dither signal with $S/N = 1$.

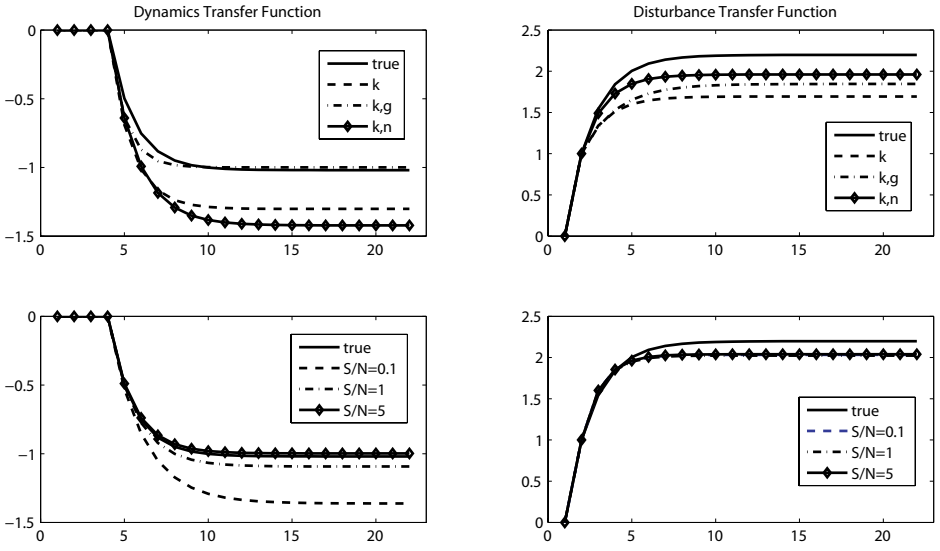

Figure 5: The unit step response of the dynamics and disturbance transfer functions of the true and the estimated models. The estimated models are obtained by using the CLS and dither signal approaches. The step input is applied at time 1. **Left panels**: the dynamics transfer function, **right panels**: the disturbance transfer function.

Comparison of the CLS and the dither signal approaches

Model parameter estimates $\hat{\boldsymbol{\beta}}$ were computed from the closed-loop output data (with $S/N = 0$) using the CLS method and from the closed-loop input and output data under $S/N = 0.1, 1$ and 5 using the dither signal method. We summarized these estimates by considering the corresponding estimated dynamics transfer function $\frac{b_1\mathcal{B}+b_2\mathcal{B}^2+b_3\mathcal{B}^3}{1-a_1\mathcal{B}}$ and the estimated disturbance transfer function $\frac{1-\theta\mathcal{B}}{1-\phi\mathcal{B}}$. Figure 5 shows the unit step response of the estimated and true transfer functions and illustrates how the model estimates improve as more constraints are added or as the signal to noise ratio of the dither signal is increased.

The two identification approaches were compared on the basis of the costs they incur during identification and during operation. The cost of identification and the cost of operation were measured by the variability of the output around the target (the mean squared deviation from target) $MSE(Y_t)$ and by the variability of the input $Var(U_t)$ during the identification experiment and during the operation. During the operation, an optimal controller re-designed based on the identified model adjusts the process.

Similarly to section 5, we considered MMSE controllers as the optimal controller. Using the expression (27), we computed an estimated MMSE controller $\hat{U}_t^*$ for each of the models estimated by the two approaches. As a point of reference, we also computed the true MMSE

Controller design using closed-loop data using the first PI controller with $c_1 = 1$ and $c_2 = -0.7$ and open-loop data:

Cost		Open loop (True MMSE)	Closed-loop w/ Dither Signal			Closed-loop w/ Constraints			
			$S/N = 0.1$	$S/N = 1$	$S/N = 5$	k	k, g	k, n	k, n, g
Ident.	$MSE(Y_t)$	10.478	0.804	1.247	2.769	0.759	0.759	0.759	0.759
($N = 296$)	$Var(U_t)$	1.151	1.3706	2.522	6.620	1.264	1.264	1.264	1.264
Oper.	$MSE(Y_t)$	0.432	0.529	0.460	0.451	0.520	0.484	0.472	0.436
($N = 296$)	($s.e.$)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	$Var(U_t)$	14.820	14.197	14.446	14.544	14.380	14.732	14.473	14.662
	($s.e.$)	(0.401)	(0.400)	(0.401)	(0.401)	(0.401)	(0.402)	(0.401)	(0.401)

Controller design using closed-loop data by using the second PI controller with $c_1 = 0.05$ and $c_2 = -0.03$ and open-loop data:

Cost		Open loop (True MMSE)	Closed-loop w/ dither signal			Closed-loop w/ CLS			
			$S/N = 0.1$	$S/N = 1$	$S/N = 5$	k	k, g	k, n	k, n, g
Ident.	$MSE(Y_t)$	10.478	0.705	0.956	2.119	0.683	0.683	0.683	0.683
($N = 296$)	$Var(U_t)$	1.151	0.106	0.778	3.812	0.037	0.037	0.037	0.037
Oper.	$MSE(Y_t)$	0.432	0.467	0.434	0.432	4.935	3.261	1.587	0.527
($N = 296$)	($s.e.$)	(0.002)	(0.002)	(0.002)	(0.002)	(0.091)	(0.121)	(0.018)	(0.002)
	$Var(U_t)$	14.820	14.683	14.903	14.816	32.909	40.772	18.571	15.495
	($s.e.$)	(0.401)	(0.401)	(0.401)	(0.401)	(0.776)	(1.217)	(0.414)	(0.401)

Table 3: Cost of identification and cost of operation for closed-loop identification of the gas furnace process using the CLS and dither signal approaches.

controller of this process. The true MMSE controller uses the parameter estimates obtained from the open-loop data.

The cost of operation was calculated by simulating the process 1000 more times for 296 periods each under the action of these controllers. Table 3 summarizes the cost of identification and the cost of operation of the controllers designed with the two approaches and those of the true MMSE controller. During identification (performed only once) the cost for the estimated controllers is calculated from the closed-loop data of the PI controllers. With the dither signal approach the input varies more due to the dither signal, while with the CLS approach no dither signal is added hence the output MSE and the input variance remain at the corresponding PI control operation levels. During operation, the $MSE(Y_t)$ and $Var(U_t)$ are the averages of the values computed from the simulations and the numbers in the parentheses are the corresponding standard errors. The open-loop column gives the corresponding costs when applying the true MMSE controller.

The two PI controllers represent rather extreme closed-loop experimental conditions for

identification with constraints: the first one varies the input more frequently (i.e. it incurs in a larger cost of identification) but results in better parameter estimates which in turn provide a lower cost of operation. In contrast, the second controller varies the input less frequently and results in relatively poor parameter estimates which in turn provide a higher cost of operation.

However, the CLS approach gave a smaller overall cost than the dither signal approach under both controllers. When the process was identified from the closed-loop data using the second PI controller, the CLS approach (under the constraints k, n, g) gave only a slightly larger operation cost than the dither signal approach (under $S/N = 1$), while incurring in a considerably smaller cost of identification (in identification, $S/N = 1$ requires 20 times as large input variance as that required by k, n, g). When the process was identified from the closed-loop data using the first PI controller, the CLS approach achieved a smaller cost of operation than the dither signal approach (under $S/N = 0.1$) and incurred in a still smaller identification cost when only the delay (k) was assumed known. As more constraints were added, the CLS approach (under the constraints k, n, g) achieved a smaller cost of operation than the dither signal approach (under $S/N = 1$ and $S/N = 5$) while incurring in a much lower identification cost (in identification, $S/N = 1$ requires almost twice as large input variance as that required by k, n, g). We also note that the cost of operation with the constraints k, n, g is almost equal to that of the true MMSE controller.

The results from using either of the PI controllers illustrate that, by using the CLS approach and when prior knowledge about the process is available, one can identify models with better quality (i.e. with lower cost of operation) than the dither signal approach by introducing much less extra variability into the process (i.e. by incurring in a lower cost of identification).

7 Conclusions

This paper presented a closed-loop system identification methodology, CLS, that uses prior process knowledge to obtain improved model parameter estimates. We presented a simulation example and a case study that demonstrated the benefits of using various forms of prior process knowledge with different model and sample sizes over alternative methods, in particular, over the dither signal method. All investigated cases assumed a model that contains the true process model. The results of the simulation example indicated that, regardless of the model and sample size used, the knowledge of the input-output delay (k) is crucial in closed-loop identification. It was further shown that when a small model is used, prior knowledge of the delay and the gain (k, g) give best results. Likewise, when a large model is used, knowing the delay and the noise model (k, n) give the best identified models. This is

true with respect to both the bias and variance of the estimates and from a control performance perspective. Therefore, if the objective in modelling the process is to better control it, it is suggested that more effort be spent in identifying the delay and the asymptotic gain of the process. The case study based on real process data shows that the proposed CLS method provides better quality models than the dither signal approach without having the need to introduce extraneous variability into the process.

8 Acknowledgements

The authors would like to thank the Editor, Associate Editor and a referee for their constructive comments. Thanks are also due to Professor Rong Pan (Arizona State University) for reviewing and commenting on an earlier version of the paper.

References

[1] Åström, K., J. (1970), *Introduction to Stochastic Control Theory*, Academic Press, Inc.

[2] Bazaraa, M. S., Sherali, H. D., Shetty, C. M. (1993), *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons.

[3] Björklund, S., Ljung, L. (2003), "A Review of Time-Delay Estimation Techniques," *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii USA.

[4] Box, G. E. P., Jenkins, G. W., Reinsel, G. C. (1994), *Time Series Analysis, Forecasting and Control*, Prentice Hall, Inc.

[5] Box, G. E. P., Luceño, A. (1997), *Statistical Control by Monitoring and Feedback Adjustment*, Wiley.

[6] Box, G. E. P., MacGregor, J. F. (1974), "Analysis of Closed-Loop Dynamic-Stochastic Systems", *Technometrics*, 16, No. 3, 391-398.

[7] Box, G. E. P., MacGregor, J. F. (1976), "Parameter Estimation with Closed-Loop Operating data", *Technometrics*, 18, No. 4, 371-384.

[8] Forssell, U., Ljung, L. (2000), "A Projection Method for Closed-Loop Identification", *IEEE Transactions on Automatic Control*, 45, No. 11, 2101-2106.

[9] Ljung, L., Gustavsson, I., Söderström, T. (1974), "Identification of Linear Multivariable Systems Operating Under Linear Feedback Control", *IEEE Transactions on Automatic Control*, 19, No. 6, 836-840.

[10] Ljung, L. (1978), "Convergence Analysis of Parametric Identification Methods", *IEEE Transactions on Automatic Control*, 23, No. 5, 770-783.

[11] Ljung, L., Forssell, U. (1998), "Bias, Variance and Optimal Experiment Design: Some Comments on Closed Loop Identification". *Perspectives in Control*, editor Norm, D., Springer-Verlag.

[12] Ljung, L. (2002), *MATLAB System Identification Toolbox User's Guide*.

[13] Luceño, A. (1997), "Parameter Estimation With Closed Loop Operating Data Under Time Varying Discrete Proportional-Integral Control", *Communications in Statistics-Simulation*, 26, No. 1, 215-232.

[14] MacGregor, J. F., Fogal, D. T. (1995), "Closed-Loop Identification: The role of the noise model and prefilters", *Journal of Process Control*, 5, No. 3, 163-171.

[15] Marriott, J., Ravishanker, N., Gelfand, A., Pai, J. (1996). "Bayesian Analysis of ARMA Processes: Complete Sampling Based Inference under Exact Likelihoods". *Bayesian Analysis in Statistics and Econometrics*, Edited by Berry, D. A., Chaloner, K. M., Geweke, J. K., John Wiley & Sons, Inc.

[16] Myers, R. H. (1990), *Classical and Modern Regression with Applications*, Duxbury Classic Series.

[17] Pan, R., Del Castillo E. (2001), "Identification and Fine Tuning of Closed-Loop Processes Under Discrete EWMA and PI Adjustments", *Quality and Reliability Engineering International*, 17, 419-427.

[18] Saikkonen, P., and Luukkonen, R. (1993), "Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models", *Journal of American Statistical Association*, 88, 596-601.

[19] Söderström, T., Gustavsson, I., Ljung, L. (1975), "Identifiability conditions for linear systems operating in closed loop", *International Journal of Control*, 21, No. 2, 243-255.

[20] Söderström, T., Ljung, L., Gustavsson, I. (1976), "Identifiability Conditions for Linear Multivariable Systems Operating Under Feedback", *IEEE Transactions on Automatic Control*, Short Papers, 837-840

[21] Van Den Hof, P.M.J., Schrama, R.J.P. (1993), "An Indirect Method for Transfer Function Estimation from Closed-Loop Data", *Automatica*, 29, No. 6, 1523-1527.