# Robust Identification of Process Models from Plant Data

Graham C. Goodwin [a], Juan C. Agüero [a], James S. Welsh [a],
Juan I. Yuz [b], Greg J. Adams [a], and Cristian R. Rojas [a]

[a] *The University of Newcastle, NSW 2308 Australia*

[b] *Electronics Department, Universidad Técnica Federico Santa María, Valparaíso, Chile.*

## Abstract

A precursor to any advanced control solution is the step of obtaining an accurate model of the process. Suitable models can be obtained from phenomenological reasoning, analysis of plant data or a combination of both. Here, we will focus on the problem of estimating (or calibrating) models from plant data. A key goal is to achieve *robust* identification. By *robust* we mean that small errors in the hypotheses should lead to small errors in the estimated models. We argue that, in some circumstances, it is essential that special precautions, including discarding some part of the data, be taken to ensure that robustness is preserved. We present several practical case studies to illustrate the results.

*Key words:* Closed loop identification; robust identification.

## 1 Introduction

The word *robust* has, in recent years become inextricably linked to advanced control. This has been an important step with significant practical consequences. However, by way of contrast, little has been explicitly written about robust identification, although the idea is implicit in much of the previous literature [32]. Nonetheless, robustness issues play a central role in successful identification experiments.

One way of thinking about robustness is that *small* changes in the working hypotheses should lead to *small* changes in the end result. This may seem rather obvious, but surprisingly, robustness issues are sometimes overlooked. For example, we know of an industrial case study where a $40^{th}$ order non-minimum phase model was fitted to a particular plant. It subsequently turned out that a more robust procedure showed that a much better description was provided by a simple integrator plus time delay.

Robust Identification has been previously studied in Econometrics and Statistics (see e.g. [24,21,13]). For example, estimation algorithms that consider the presence of outliers in the data have been proposed (see e.g. [12]). On the other hand, linear regression in the frequency domain has been proposed in Time Series Analysis literature (see [22] and the references therein). Moreover, the idea of using the data in a determined frequency range has been used in [23] and [37] to deal with seasonal noise, trends, and aliasing. In [14] the term *Frequency band estimation* was coined and employed in the testing of economic hypotheses. Extension to the identification of static Errors in Variables models with colored noise was presented in [38].

Here, we will suggest various strategies for *robustifying* solutions. Often, this amounts to discarding some of the data or detuning the algorithm in some way such that it is not overly sensitive to the hypotheses. Inevitably a price is paid in terms of nominal performance. However, this is often essential to achieve robustness.

The specific problems that we will study include

- selecting model class,
- sampling,
- experiment design,
- closed loop identification, and
- errors in variables identification.

## 2  Motivating practical case study

The system is a continuous metal caster which requires very good level control in the mould to ensure that product quality is maintained at a high standard. The system input is the position of a slide gate valve, and the output is the height of molten metal in the mould.

The company had applied (non-robust) identification methods, to data collected during normal operation, to obtain a model of the system for the purpose of improving their control system. Their analysis recommended models

2

of 4th and 5th order. When these models were used to design a controller, very poor control was experienced. The problem was that the data was collected under closed loop conditions with poor external excitation and significant output disturbances. Under these conditions, there is a strong chance that identification methods will simply estimate the negative inverse of the controller (or, at least, a causal approximation to this transfer function).

To understand why this occurs, consider the simple feedback loop shown in Figure 1.

One relationship between the input, $u_t$ and the output, $y_t$ is:

$$y_t = G_o u_t + v_t \tag{1}$$

However, there exists a second relationship between input and output via the feedback path i.e.

$$u_t = C[r_t - y_t] \tag{2}$$

or

$$y_t = -\frac{1}{C}u_t + r_t \tag{3}$$

Thus, it appears that there are two possible models i.e. (1) or (3). Special techniques exist that may rule out one of these possibilities. Some of these techniques will be discussed in detail later in the paper. However, in the absence of special precautions, one will tend to estimate $G_o$ when $r_t$ dominates $v_t$, and $-1/C$ when $v_t$ dominates $r_t$ and somewhere between $G_o$ and $-1/C$ when neither $r_t$ nor $v_t$ dominates.

In the continuous caster identification experiment, the estimates were indeed strongly biased toward the negative inverse of the controller. This can be clearly seen in Figure 2 were a Bode magnitude plot of the Empirical Transfer Function Estimate (ETFE) [32] and $-1/C$ are presented. In a subsequent experiment, external excitation was applied to the system such that a reasonably large signal to noise ratio was obtained at a small number of frequencies. In this particular case the external signal was added to the controller output (plant input). The Bode magnitude response of the system to the test signal can be seen in Figure 2 as indicated by the $\times$ symbols.

Paradoxically, by discarding most of the data and utilising only the data associated with the externally applied test frequencies in the identification algorithm it is clearly seen by the dash-dot line in Figure 2 that the estimated model is very similar to that of the theoretical model (dotted line).

This case study shows that thought must be given to the experiment and data used. It is not sufficient to collect normal operating data and expect

3

to model anything other than the negative inverse of the controller without special care. Also, the example suggests that robustness can be enhanced by discarding parts of the data. These ideas will be explained more fully below.

## 3   Selecting a model class

### 3.1   The role of physics

The first issue to be addressed in a system identification exercise is the specification of the class of models to be fitted. Here the physics of the problem plays a central role in achieving robustness. This does not mean that one necessarily has to develop a large scale, distributed parameter model. Indeed, quite to the contrary, a large scale model will typically contain far too many free parameters to be calibrated. The key thing is to be able to capture the *essential physics* of the problem. There are many examples, where a simple physical model would have saved a lot of subsequent difficulties in system identification. Thus, our recommendation is to always begin with simple physical reasoning to suggest a model structure. This will typically take the form of a set of ordinary differential equations (possibly nonlinear and with time delays).

### 3.2   Which operator?

The most common operator used for discrete time models is the shift operator. However, this operator can lead to robustness issues. This is because:

(1) shift operator parameters typically lack physical significance,
(2) shift operator models are usually associated with nontrivial numerical problems. (The source of these difficulties is that a near perfect model, with fast sampling, is invariably $y_{t+1} = y_t$). More generally, with moderately fast sampling, shift operator models have coefficients that approach the Binomial coefficients,
(3) it is difficult to subsequently change the sampling period once a shift operator model has been obtained.

By way of illustration of the above difficulties, we point to the following two discrete time second order systems expressed in terms of the shift operator

(equivalently, the $z-$ Transform variable $z$):

$$G_1(z) = \frac{0.048(z + 0.9672)}{z^2 - 1.9025z + 0.9048} \tag{4}$$

$$G_2(z) = \frac{0.048(z + 0.9672)}{z^2 - 1.9120z + 0.9048} \tag{5}$$

Notice that the denominator coefficients are very near to $(1, -2, 1)$ i.e. the Binomial coefficients–see point (2) above. It may surprise the reader, that these two models exhibit very different behaviour. Indeed, (4) is stable whilst (5) is unstable. Yet, the only difference in these two discrete time models is a subtle (0.5%) change in one of the coefficients!. Actually, (4) and (5) are the exact discrete equivalents of the following two continuous time systems, having zero order hold input and sampling period 0.1 seconds:

$$G_1(s) = \frac{1}{s^2 + s + 0.25} = \frac{1}{(s + 0.5)^2} \tag{6}$$

$$G_2(s) = \frac{1}{s^2 + s - 0.75} = \frac{1}{(s + 1.5)(s - 0.5)} \tag{7}$$

In continuous time models we clearly see that the coefficients differ by 400% and that (7) is unstable! More will be said about this example later. Our recommendation is to either use continuous time descriptions or delta operator based models. The latter has a close connection to continuous models as we show in section 5.

## 4  Choice of sampling strategy

The next step is to choose a suitable sample period. A simple rule of thumb is to sample as fast as possible and certainly one tenth of the dominant time constant. One important point that is sometimes overlooked is that sampling should always be preceded by low-pass (anti-aliasing) filtering to avoid folding of high frequency noise back into the bandwidth of interest. Also, these Anti-aliasing filters form part of the system description which needs to be accounted for if one seeks high fidelity for rapidly changing (i.e. high frequency) inputs. With high speed electro-mechanical systems anti-aliasing filtering is easily carried out with analogue filters. However, filtering is more difficult for systems with long time constants (as are typical for chemical process models). In the latter case, anti-alias filtering can be performed digitally by sampling at a higher rate than finally needed and then low pass filtering via digital filtering techniques. (Of course, there will also be an analogue anti-aliasing filter at the fast sampling rate) Note that, for robustness reasons, anti-aliasing filtering should be conducted well above the maximum frequency of interest to avoid introducing phase shifts or other contaminations in the range of interest.

## 5 Sampled data models for continuous time systems

We have argued in section 4 that one should choose sampling periods which are relatively small compared to the dynamics of interest. However, this leads to numerical issues as discussed in Section 3.2. These issues are mitigated by the simple transformation of variables[1] :

$$\delta = \frac{q-1}{\Delta} \tag{8}$$

This operator is known as the *Delta Operator* [34]. Moreover, one can readily obtain an approximate discrete time model by replacing derivatives in the underlying continuous time model by divided differences i.e. $\delta$. The approximate (derivative replacement) discrete time models corresponding to (4), (5) are

$$[\delta^2 + \delta + 0.25]y_t = u_t \tag{9}$$
$$[\delta^2 + \delta - 0.75]y_t = u_t \tag{10}$$

We see that these models inherit the informative large parameter difference seen in the continuous model. This should be contrasted with the alternative shift operator models given in (4) and (5). This suggest that numerical properties will be much improved in the delta form compared with shift operator models. For example, it can be shown that the conditioning number for the least squares parameter estimation goes to $\infty$ (as $\Delta \to 0$) in shift form whereas it goes to 1 in delta form [34] (see also [20] regarding relative errors due to zero dynamics.)

Thus, we strongly advocate the use of the Delta operator especially at fast sampling rates. However, there has been some reluctance to use Delta domain models because it is (incorrectly) believed that difference operation will be sensitive to noise. This is actually false since, Delta models are simply a linear re-parameterization of shift operator models achieved by an origin change. Thus, the alternative operator only affects numerical issues and has no impact on sensitivity to noise. Indeed, the delta operator makes explicit the dependence on differences which is always implicit in discrete models (but alas hidden) in shift operator models.

Another confusion is that use of Delta operator is equivalent to use of "derivative replacement" as in (9) and (9). One should not confuse the model type with the operator. Indeed, simple derivative replacement models will always give a poor description if used in the vicinity of the sampling frequency. This is because sampling inevitably involves folding (i.e. aliasing) of high frequency

_____
[1]  $\Delta$: sampling period.

components back onto the range $(0, \pi/\Delta)$ in the frequency domain. For example if we assume that the continuous-time system frequency response $G(\mathrm{j}\omega)$ goes to zero as $|\omega| \to \infty$, then the corresponding discrete-time frequency is:

$$G_q(e^{\mathrm{j}\omega\Delta}) = \sum_{\ell=-\infty}^{\infty} \left[ \frac{(1 - e^{-s\Delta})}{s\Delta} G(s) \right]_{s=\mathrm{j}\omega+\mathrm{j}\frac{2\pi}{\Delta}\ell} \tag{11}$$

The impact of the folding described in (11) is illustrated in Figure 3. This figure shows a comparison of the Bode magnitude diagrams corresponding to a second order system and the exact sampled-data model obtained for 3 different sampling frequencies.

The figure clearly illustrates the fact that, **no matter how fast we sample**, there is always a difference (near the folding frequency) between the continuous-time model and the corresponding discretised model. A consequence of the folding of high frequency dynamics back onto lower frequencies is that additional zero dynamics are introduced into the corresponding sampled data model. Thus, for example (9), (10) are more accurately described as in the delta models presented earlier in equations (4) and (5). This suggest that we should actually use models of the following form in delta domain (for a second order system with zero order hold input)

$$[\delta^2 + \delta + 0.25]y_t = [\beta\delta + 1]u_t \tag{12}$$
$$[\delta^2 + \delta - 0.75]y_t = [\gamma\delta + 1]u_t \tag{13}$$

The additional zero corresponds to the asymptotic sampling zero near $z = -1$ in the discrete time models (4) and (5). Moreover, provided one samples relatively quickly (say ten times the transients of interest), then $\beta$ and $\gamma$ in (12) and (13) can be approximately fixed at the asymptotic value of $-1$ (in the $z-$ domain) or $0.5\Delta$ (in the delta domain)[15,20].

The above ideas can be readily generalized. Indeed, there exists a comprehensive theory describing additional *sampling zeros* of the type shown in (12) and (13). Specifically, for small $\Delta$, it can be shown that, for a system of relative degree $r$, then $(r-1)$ additional sampling zeros appear which asymptotically tend to the roots of the Euler Frobenius polynomials (see e.g. [11], [44], [45]). This leads to an interesting follow up robustness question, namely, *Should we incorporate the sampling zeros in our model for identification purposes?* Certainly, if one wishes to obtain small (relative) model errors in the vicinity of the sampling frequency, then one has no choice but to include the sampling zeros [47,19].

This is illustrated in the following example

**Example 5.1** *Consider a second order linear system:*

$$G_o(s) = \frac{\beta_0}{s^2 + \alpha_1 s + \alpha_0} \tag{14}$$

*where the continuous-time parameters are given by $\alpha_1 = 3$, $\alpha_0 = 2$, $\beta_0 = 2$. System identification was carried out assuming three different model structures:*

**SDRM:** *Simple Derivative Replacement Model. This corresponds to the structure, where continuous-time derivatives are simply replaced by divided differences.*

**MIFZ:** *Model Including Fixed Zero. This model considers the presence of the asymptotic zeros.*

**MIPZ:** *Model Including Parameterised Zero. This model also includes a sampling zero, whose location has to be estimated.*

*The three discrete-time models can be represented in terms of the $\delta$ operator as:*

$$G_\delta(\gamma) = \frac{B_\delta(\gamma)}{\gamma^2 + \alpha_1 \gamma + \alpha_0} \tag{15}$$

*where:*

$$B_\delta(\gamma) = \begin{cases} \beta_0 & \textit{(SDRM)} \\ \beta_0(1 + \frac{\Delta}{2}\gamma) & \textit{(MIFZ)} \\ \beta_0 + \beta_1\gamma & \textit{(MIPZ)} \end{cases} \tag{16}$$

*We use a sampling period $\Delta = \pi/100\,[s]$ and choose the input $u_k$ to be a random Gaussian sequence of unit variance. Note that the output sequence $y_k = y(k\Delta)$ can be obtained by either simulating the continuous-time system and sampling its output, or, alternatively, by simulating the exact sampled-data model in discrete-time. Also note that the data is free of any measurement noise. The parameters are estimated in such a way to minimise the following equation error cost function:*

$$J(\theta) = \frac{1}{N}\sum_{k=0}^{N-1} e_k(\theta)^2 = \frac{1}{N}\sum_{k=0}^{N-1}(\delta^2 y_k - \phi_k^T\theta)^2 \tag{17}$$

*where:*

$$\phi_k = \begin{cases} [-\delta y_k, \; -y_k, \; u_k]^T \\ [-\delta y_k, \; -y_k, \; (1 + \frac{\Delta}{2}\delta)u_k]^T \\ [-\delta y_k, \; -y_k, \; \delta u_k, \; u_k]^T \end{cases} \tag{18}$$

*and*

$$\theta = \begin{cases} [\alpha_1, \ \alpha_0, \ \beta_0]^T & \text{(SDRM)} \\ [\alpha_1, \ \alpha_0, \ \beta_0]^T & \text{(MIFZ)} \\ [\alpha_1, \ \alpha_0, \ \beta_1, \beta_0]^T & \text{(MIPZ)} \end{cases} \tag{19}$$

*Note that, in this example, we are estimating the parameters using unfiltered "equation errors". In practice this is not recommended as it is generally extremely non robust. The reason is that equation error models emphasize fitting at high frequencies and are thus very sensitive to high frequency model errors including the issue of having the correct sampling zeros. To explain the origin of the difficulty, consider a general transfer function model of the form:*

$$A_o(q^{-1})y_t = B_o(q^{-1})u_t + w_t \tag{20}$$

*where $w_t$ is zero mean white noise with variance $\sigma_w^2$. Using Parseval's Theorem it is readily seen [32] that the equation error cost function tends to*

$$J = \int |A|^2 |(G_o - G)\Phi_u|^2 d\omega + \int \left| \frac{A}{A_o} \right|^2 \sigma_w^2 d\omega$$

*where $\Phi_u$, $G_o = \frac{B_o}{A_o}$, and $G = \frac{B}{A}$ are the input spectrum ,the "true" system and the model respectively. Now $A(q^{-1})$ is, typically, high pass. Thus, we see that the estimator fits the error between $G_o$ and $G$ emphasizing high frequencies. Consequently, if one wants to use an equation error estimator then one should always use an appropriate filter, $E(q^{-1})$, to focus the fit on the frequency range of interest. With pre-filtering of $y_t$ and $u_t$ by $E(q^{-1})$ the cost function becomes*

$$J = \int \left| \frac{A}{E} \right|^2 |(G_o - G)\Phi_u|^2 d\omega + \int \left| \frac{A}{EA_o} \right|^2 \sigma_w^2 d\omega \tag{21}$$

*The filter $E(q^{-1})$ can thus focus the estimator in the frequency range of interest. In the remainder of this example we will continue to use unfiltered equation errors so as to emphasize the importance of using the correct sampling zeros if a model with high fidelity at high frequency is desired. Table 1 shows the estimation results in the delta domain. Note that the system considered is linear, thus, the exact discrete-time parameters (exact DT) can be computed for the given sampling period. These are also given in Table 1. We see that, while both models incorporating a sampling zero (MIFZ and MIPZ) are able to recover the correct discrete model 9and hence the approximate continuous-time parameters), when using SDRM the estimate $\hat{\beta}_0$ is clearly biased (by a factor of almost $2:1$).*

The result in the previous example may be surprising since, even though the SDRM converges to the continuous-time system as the sampling period goes to zero, the estimate $\hat{\beta}_0$ does not converge to the underlying continuous-time parameter. This estimate is asymptotically biased [45,47]. Specifically, we see

that $\beta_0$ is incorrectly estimated by a factor of 2 by the SDRM. This illustrates the impact of not considering sampling effects on the sampled-data models used for continuous-time system identification.

The above discussion and example suggest that one cannot, in general, ignore the effect of folding in sampled data models. Indeed, one should include the *sampling zero dynamics* which result from the folding of high frequency components back into the range $(0, \pi/\Delta)$. These ideas can also be extended to the nonlinear case [46]. However, there is a further robustness issue to be considered. Specifically, the high frequency components that are folded back by sampling are likely to be ill-defined or non-stationary in practice due to one's inability to exactly describe systems at high frequencies. Hence, it seems desirable to not place too much confidence in folded artifacts. For example, the asymptotic sampling zero of $(0.5\Delta\delta + 1)$ described in relation to (12), (13) holds only under very precise conditions i.e. (i) that the input is generated by a zero order hold and (ii) no under-modelled poles or zeros lie above $\pi/\Delta$. If these assumptions do not hold then this can be a source of problems [47]. A simple procedure for robustifying against this issue is discussed in the next section.


## 6    Limited bandwidth estimation


Identification procedures can be robustified to the effects of high frequency folding (including the presence of sampling zeros and anti-aliasing filters) by simply avoiding estimators which focus on frequencies near the folding frequency. This can be readily achieved in the frequency domain (for linear systems) as we show below. Similar ideas can be develop in the time domain.

One can readily perform robust Frequency Domain Maximum Likelihood (FDML) estimation for linear models. The core idea is to convert the data to the frequency domain and then carry out the identification over a limited range of frequencies. This amounts to not using some parts of the data. This will reduce nominal performance but can lead to substantial improvements in robustness. Note, however, that one needs to carefully define the likelihood function in this case. For example, the following result Note this result has been derived in [31] for the scalar case, and in [33] for the multi-variable case.

**Lemma 1** *Assume a given set of input-output data* $\{u_k = u(k\Delta), y_k = y(k\Delta)\}$, $k = 0 \dots N$ , *is generated by the following discrete-time model:*

$$y_k = G_q(q, \theta)u_k + H_q(q, \theta)w_k \tag{22}$$

*where* $w_k$ *is Gaussian discrete-time white noise (DTWN) sequence,* $w_k \sim$

$N(0, \sigma_w^2)$.

*The data is transformed to the frequency domain (ignoring transients) yielding the discrete Fourier transforms $U_\ell$ and $Y_\ell$ of the input and output sequences, respectively. Then the maximum likelihood estimate of $\theta$, when considering frequency components up to $\omega_{\max}$, is given by:*

$$\hat{\theta}_{ML} = \arg \min_\theta L(\theta) \tag{23}$$

*where $L(\theta)$ is the negative logarithm of the likelihood function of the data given $\theta$, i.e.*

$$\begin{aligned} L(\theta) &= -\log p(Y_0, \ldots, Y_{n_{\max}} | \theta) \\ &= \sum_{\ell=0}^{n_{\max}} \frac{|Y_\ell - G_q(e^{\mathrm{j}\omega_\ell \Delta}, \theta) U_\ell|^2}{\lambda_w^2 |H_q(e^{\mathrm{j}\omega_\ell \Delta}, \theta)|^2} + \log(\pi \lambda_w^2 |H_q(e^{\mathrm{j}\omega_\ell \Delta}, \theta)|^2) \end{aligned} \tag{24}$$

*where $\lambda_w^2 = \Delta N \sigma_w^2$, and $n_{\max}$ is the index associated with $\omega_{\max}$.*

□□□

**Remark 2** *Within the above framework, the logarithmic term must be included in the log-likelihood function since this plays a key role in obtaining consistent estimates of the true system. This term can be neglected only under special circumstances e.g. if [31]:*

- *The noise model is assumed to be known. In this case $H_q$ does not depend on $\theta$ and, thus, plays no role in the minimisation (23); or*
- *The frequencies $\omega_\ell$ are equidistantly distributed over the full frequency range $[0, \frac{2\pi}{\Delta})$. This is equivalent to considering the **full bandwidth** case in (24), i.e. $n_{\max} = \frac{N}{2}$ (or $N$, because of periodicity). This yields:*

$$\frac{2\pi}{N} \sum_{\ell=0}^{N-1} \log |H_q(e^{\mathrm{j}\omega_\ell \Delta}, \theta)|^2 \xrightarrow{N \to \infty} \int_0^{2\pi} \log |H_q(e^{\mathrm{j}\omega}, \theta)|^2 d\omega \tag{25}$$

*A standard result from Complex analysis (The Bode integral [17]) ensures that the last integral is equal to zero for any monic, stable and inversely stable transfer function $H_q(e^{\mathrm{j}\omega}, \theta)$.*

▽▽▽

Using the cost function (24) it is straightforward to estimate the parameters. Moreover, tests, see e.g. those reported in [45,47], show this to be a robust strategy. The above ideas can also be extended to closed loop identification (see [35]), and state-space multivariable systems [9].

## 7 Robust experiment design

Having decided on the class of models and a suitable sampling strategy, then the next step is to design a suitable robust experiment. We saw in the motivating example of Section 2 that estimation under closed loop conditions is a potential source of robustness problems. On the other hand, colleagues from industry often tell us that the only acceptable experiment (to them) is one performed in closed loop since this ensures that all safety and feedback mechanism are in place. Indeed in extreme circumstances, it is often said that the experiment should not be detectable on the plant output records (at least as far as plant operators are concerned). This goal can be given a mathematical description by requiring that the plant output satisfy certain production oriented constraints during the experiment. We will argue below that this type of constraint invariably implies that the experiment should be conducted in closed loop.

For simplicity of exposition we will develop the mathematical support for our conclusions based on linear single input single output models. However, one can reasonably expect that similar qualitative conclusions hold more generally (i.e. for nonlinear models, etc.). Thus, consider a single input - single output linear system of the form

$$\mathcal{S} = \{(G_o, H_o) \in \mathcal{C} : y(t) = G_o(q^{-1})u(t) + H_o(q^{-1})w(t)\} \qquad (26)$$

where $\mathcal{C}$ is the set of causal linear systems, $q^{-1}$ is the unit delay operator and $G_o(q^{-1}) = q^{-d}\bar{G}_o(q^{-1})$ $(\bar{G}_o(0) = b_0 \neq 0, d \in \mathbb{N})$ [2] and $\{w(t)\}$ is zero mean white noise sequence with variance $E\{w(t)^2\} = \sigma_w^2$ (note that $\sigma_w^2$ is also the noise spectral density). We take $H_o(q^{-1})$ to be the stable minimum phase spectral factor, and $H_o(0) = 1$. We consider Box-Jenkins models of the form

$G(q^{-1}, \theta) = G(q^{-1}, \rho), H(q^{-1}, \theta) = H(q^{-1}, \eta)$ where $\theta = \begin{bmatrix} \rho \\ \eta \end{bmatrix}$. Notice that the

sub-parameter vector $\rho$ refers only to $G(q^{-1})$.

Under mild conditions, it is well known that, when using the Prediction Error Method (PEM), [18, page 99], [32, page 282] we have:

$$\sqrt{N}(\hat{\theta}_N - \theta_o) \xrightarrow{d} N(0, P_\theta) \qquad (27)$$

where the matrix $P_\theta$ (assumed non-singular) is given by:

$$P_\theta = \sigma_w^2 \left[ \bar{E}\{\Psi(t, \theta_o)\Psi(t, \theta_o)^T\} \right]^{-1} \qquad (28)$$

---

[2] We take $\bar{G}_o(q^{-1}) = \frac{\bar{B}_o(q^{-1})}{A_o(q^{-1})}$.

and $\Psi(t, \theta) = -\frac{\partial \epsilon(t,\theta)}{\partial \theta}$, $\epsilon(t, \theta) = H(q^{-1}, \theta)^{-1}[y(t) - G(q^{-1}, \theta)u(t)]$.

The covariance of the parameter $\hat{\theta}_N$ is usually approximated as $\text{cov}\{\hat{\theta}_N\} \approx \frac{1}{N}P_\theta$ for experiment design [32, Chapter 9] ($N$ is the number of data points).

It is well known [32] that, for BJ models, the following is satisfied:

$$\Psi(t, \theta_o) = \begin{bmatrix} H_o(q^{-1})^{-1}\Lambda_{\rho_o}(q^{-1})u(t) \\ H_o(q^{-1})^{-1}\Lambda_{\eta_o}(q^{-1})w(t) \end{bmatrix}$$

where

$$\Lambda_{\rho_o}(q^{-1}) = \left.\frac{\partial G(q^{-1}, \rho)}{\partial \rho}\right|_{\rho=\rho_o}$$

$$\Lambda_{\eta_o}(q^{-1}) = \left.\frac{\partial H(q^{-1}, \eta)}{\partial \eta}\right|_{\eta=\eta_o}$$

The information matrix for the full parameter vector is given by:

$$M_\theta = NP_\theta^{-1} = \frac{N}{\sigma_w^2}\bar{E}\{\Psi(t, \theta_o)\Psi(t, \theta_o)^T\}$$

This can also be re-written using Parseval's Theorem as [3] [4] [32, page 291]:

$$M_\theta = \frac{N}{\sigma_w^2}\frac{1}{2\pi}\begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \tag{29}$$

where

$$A = \int G_1 G_1^H \Phi_u, \ B = \int G_1 G_2^H \Phi_{uw},$$

$$D = \int G_2 G_2^H \sigma_w^2, \ G_1 = H_o^{-1}\Lambda_{\rho_o}, \ G_2 = H_o^{-1}\Lambda_{\eta_o}$$

From (29) and utilizing standard matrix algebra we have that the inverse of

---

[3] Here, and in the sequel, we omit the limits of integration and the integration variable. Unless, otherwise stated the limits of integration are from $-\pi$ to $\pi$, and the integration variable is $\omega$.

[4] We use the following notation: $\Phi_{xy} = \sum_{\tau=-\infty}^{\infty} R_{xy}(\tau)e^{-j\omega\tau}$, and $R_{xy}(\tau) = \bar{E}\{x(t)y(t-\tau)\} = \frac{1}{2\pi}\int \Phi_{xy}e^{j\omega\tau}dw$ for any pair of signals $x(t)$ and $y(t)$. We also use $\Phi_x = \Phi_{xx}$.

the covariance for $\rho$ and $\eta$ are given by:

$$\frac{2\pi\sigma_w^2}{N}P_\rho^{-1} = \int G_1 G_1^H \Phi_u - \beta(\Phi_{uw})$$

$$\frac{2\pi\sigma_w^2}{N}P_\eta^{-1} = \int G_2 G_2^H \sigma_w^2 - \gamma(\Phi_{uw}) \qquad (30)$$

where

$$\beta(\Phi_{uw}) = BD^{-1}B^T, \quad \gamma(\Phi_{uw}) = B^T A^{-1} B \qquad (31)$$

**Remark 3** *Notice that $\beta(\Phi_{uw}) \geq 0$ since $\int G_2 G_2^H \sigma_w^2$ (and its inverse) is a positive definite matrix. Similarly, $\gamma(\Phi_{uw}) \geq 0$ since $\int G_1 G_1^H \Phi_u$ (and its inverse) is a positive definite matrix.*

$$\nabla\nabla\nabla$$

For future use we next describe bounds on $P_\rho^{-1}$.

**Lemma 4** *The inverse of the covariance for $\rho$ (for any experiment $\xi$) is bounded as follows: $\frac{2\pi\sigma_w^2}{N}P_\rho^{-1}\{\xi\} \geq \int G_1 G_1^H \left[\Phi_u - \frac{|\Phi_{uw}|^2}{\sigma_w^2}\right]$. Moreover, equality holds if and only if there is a non-frequency dependent matrix $\Gamma$ (of appropriate dimensions) such that $G_1 \Phi_{uw} = \Gamma G_2$.*

**PROOF.** *This is a direct consequence of the Cauchy Schwarz inequality (see [4], [5] for details).*

$$\square\square\square$$

Our goal in the sequel will be to compare open loop and closed loop experiments. We will use $\Xi_{ol}$ and $\Xi_{cl}$ to denote open and closed loop experiments respectively. We define these classes below.

Consider a general class of experiments carried out with any linear time invariant feedback control law of the form

$$u(t) = F_o(q^{-1})r(t) - C_o(q^{-1})y(t) \qquad (32)$$

where $r(t)$ is a reference signal. This class includes open loop experiments when we take $C_o(q^{-1}) = 0$.

Under the control law (32), the closed loop satisfies

$$y(t) = G_o F_o S_o r(t) + S_o H_o w(t)$$
$$u(t) = F_o S_o r(t) - C_o H_o S_o w(t)$$

where $S_o(q^{-1})$ is the sensitivity function given by [17, page 125]:

$$S_o(q^{-1}) = \frac{1}{1 + G_o(q^{-1})C_o(q^{-1})}$$

The corresponding output, input and cross spectrum are given by

$$\Phi_y = |G_o F_o S_o|^2 \Phi_r + |S_o H_o|^2 \sigma_w^2$$
$$\Phi_u = |F_o S_o|^2 \Phi_r + |C_o H_o S_o|^2 \sigma_w^2$$
$$\Phi_{uw} = -C_o H_o S_o \sigma_w^2 \tag{33}$$

We can then define the classes of experiments of interest as follows:

**Definition 1**

$$\textbf{Open loop: } \Xi_{ol} = \{\xi \in \Xi : C_o(q^{-1}) = 0\}$$
$$\textbf{Closed loop: } \Xi_{cl} = \{\xi \in \Xi : C_o(q^{-1}) \neq 0\} \tag{34}$$

$$\triangledown\triangledown\triangledown$$

With the above background we can now define what we mean by a *good* experiment. Specifically:

**Definition 2** *The class $\Xi_{cl}$ of experiments is said to be better (not worse) than the class $\Xi_{ol}$ if and only if $\forall \xi_c \in \Xi_{cl}, \exists \xi_o \in \Xi_{ol}$ such that*[5]

$$P_\rho^{-1}\{\xi_c\} > P_\rho^{-1}\{\xi_o\} \quad \left(P_\rho^{-1}\{\xi_c\} \geq P_\rho^{-1}\{\xi_o\}\right) \tag{35}$$

$$\triangledown\triangledown\triangledown$$

The above definition uses a strong notion of optimality called *Loewner optimality* [36, chapter 4] due to the association with the ordering of symmetric matrices called *The Loewner partial ordering* (see e.g. [42]). This definition can be extended to define a partial ordering (for the reduced vector $\rho$) amongst experiments in the sense that [25]

$$\xi_1 \overset{\rho}{\succeq} \xi_2 \Leftrightarrow P_\rho^{-1}\{\xi_1\} \geq P_\rho^{-1}\{\xi_2\} \quad \xi_1, \xi_2 \in \Xi \tag{36}$$

where $\rho$ is the vector of parameters of interest and $\Xi$ is the allowable set of experiments. When strict inequality holds in (36) we write $\xi_1 \overset{\rho}{\succ} \xi_2$.

**Remark 5** *Notice that $\xi_1 \overset{\rho}{\succeq} \xi_2$ implies that $\xi_1$ is preferable under any order preserving (isotonic) criterion such as $det\{P_\theta\}$, $\lambda_{max}\{P_\theta\}$, $trace\{P_\theta\}$, etc.*
$$\triangledown\triangledown\triangledown$$

---

[5] $A \geq B$ and $A > B$ signify $A - B$ positive semidefinite and positive definite respectively.

Next, to obtain a meaningful design problem, it is necessary to place constraints on the allowable set of experiments. Based on our stated goal that the experiment should have minimal impact on nominal production, we will require that the experiment satisfy an output power constraint i.e. we require that

$$\mathcal{P}_y : \int \Phi_y \leq K \tag{37}$$

Note that this requirement is heuristically related to keeping the perceived output variations small during the experiment. The constant $K$ in (37) is assumed to be strictly greater than the minimum achievable output variance. In order to find an optimal solution we use the scalar function $J(P_\rho^{-1})$ which is any *isotonic* (order preserving) function.

Our key conclusion is then summarized in the following recent result [4,5] which generalizes results in [32]:

**Theorem 6** *For the system described in equation* (26) *and provided that a BJ model is used and that the associated minimum variance controller is different from zero then the class of experiments $\Xi_{cl}$ is better than the class of experiments $\Xi_{ol}$. Moreover, for any isotonic scalar design criterion, $J(\cdot)$ for $\rho$, the optimal experiment is necessarily in the class $\Xi_{cl}$.* □□□

The above result is very strong since it shows that, if one constrains the output power during an experiment, then one should always perform the experiment in closed loop (for a general class of systems as stated in the Theorem). The proof of this Theorem is based on the construction of a closed loop experiment which is better than any given open loop experiment. Indeed, the closed loop experiment needs only to satisfy the following:

$$\int |S_o H_o|^2 < \int |H_o|^2 \tag{38}$$

where $S_o$ is the closed loop sensitivity function.

Actually, we can gain a little more insight by examining this latter requirement for the closed loop experiment. What this says is that the (mean square) impact of disturbances during the (closed loop) experiment should be less in closed loop than it would have been had the experiment been performed in open loop. Indeed, the more we can reduce the impact of disturbances on the output via feedback, then the greater "room" we make for the output variations caused by the experimental test signal (within the limits imposed by production constraints). This seems heuristically reasonable.

One might actually believe by examining (38) that the best experiment coincides with making $\int |S_o H_o|^2$ as small as possible. Indeed, the controller that

16

makes the output variance as small as possible is commonly known as a minimum variance controller. The design of such a controller depends on having a detailed plant model. However, this introduces a further robustness problem since, presumably, the purpose of the experiment is to learn (more) about the plant model. (More will be said on this point later). However, this apparent paradox is readily resolved (at least heuristically) by utilizing ideas from traditional control theory. Specifically we know that $|S_o|$ can be made small over the frequencies where the relative model error is less than one. Hence, prior knowledge about the system dictates the bandwidth over which we can robustly obtain significant sensitivity reduction in experiment design. Indeed, an iterative design procedure can be employed where one successively fits a model (over a given bandwidth), then uses that model to design a controller to achieve a slightly wider bandwidth and so on. This also is related to the *wind-surfer* approach to adaptive control [10,30,29,26].

The key conclusion from the above reasoning is that, subject to the output variations being constrained, we should always design the experiment so that it is conducted in closed loop. We also conclude from the proof of Theorem 6 (see [4,5]) that the reference should be injected via $\bar{S}_o$ where $\bar{S}_o S_o$ is all pass. However, so far, we said nothing about the reference signal perturbations themselves. Examination of the proof of Theorem 6 shows that a good option is to use the best "open loop" test signal in conjunction with tight feedback control. Here again we are confronted by a robustness issue, namely the best open loop test signal typically depends on the nature of the system i.e. the very thing that the experiment is aimed at learning. We are thus led to consider more heuristic (and practical) issues. Practitioners who carry out experiments often report that step type test signals are good, but typically do not excite high frequencies terms adequately. On the other hand random signals such as PRBS are also considered good, but typically have wasted energy at high frequencies.

In the frequency domain, step type inputs have power that decays as $1/(freq.)^2$ $(1/f^2)$ whereas random signals have power that is constant over frequency. The above line of reasoning implies that a signal having power that lies somewhere between $1/f^2$ and a constant might be a good open-loop test signal. This suggests that a test signal with power that decays as $1/f$ (over a limited bandwidth) could be a good choice. Indeed, recent research has shown that test signals having energy that decays as $1/(f)$ have remarkable robustness properties in system identification. This claim is illustrated below for a particular example.

As a specific illustration of the properties of (band-limited) $1/f$ noise, we refer to an illustrative first order system having transfer function

$$G(s) = \frac{1}{s/\theta + 1} \tag{39}$$

The optimal nominal (open loop) test signal for identification of this system is well known to be a single sinusoid of frequency $w^* = \theta_o$ where $\theta_o$ is the a-priori estimate of $\theta$. This is an intuitively pleasing result, i.e. one places the test signal at the (nominal) 3dB break point. However, the result reinforces the fundamental robustness difficulty associated with *nominal* experiment design, namely, the optimal (open loop) test signal depends on the very thing that the experiment is aimed at estimating.

An alternative, robust input design strategy is to assume that the a-priori distribution of $\theta$ is anywhere in a compact set $\Theta$. Then, robust experiment design may be formulated as

$$\xi = \min_{\xi \in \Xi} \max_{\theta \in \Theta} J(P(\theta), \theta) \tag{40}$$

where $\xi$ is the experimental conditions (here the test signal), $J(\cdot)$ a suitable scalar function of $P(\theta)$ the parameter variance matrix.

For the one parameter problem (39), we choose $\Theta = [\theta_{min}, \theta_{max}]$ and $J(P(\theta), \theta)$ as the relative error $P(\theta)/\theta^2$. For $\theta_{min} = 1$ and $\theta_{max} = 10$.

For the above problem, the following key properties have been established in [40]:

 (i) Existence: There exists at least one optimal input.
 (ii) Uniqueness: The optimal input is unique, and $\underline{\theta}$ and $\overline{\theta}$ do not belong to the input spectrum.
 (iii) Compact support: Every optimal input should have all its energy inside $[\underline{\theta}, \overline{\theta}]$.
 (iv) Finite support: The optimal input has finite support in the frequency domain, and thus can be realised as a finite sum of sinusoids.

The robust optimal test signal obtained via the cost function (40) when the a-priori range for $\theta$ is $(0.1, 10)$ is given in reference [40].

A remarkable property (established in [40]) is that bandlimited '$1/f$' noise, defined by the spectrum

$$\phi_u^{1/f}(\omega) \triangleq \begin{cases} \dfrac{1/\omega}{\ln \overline{\omega} - \ln \underline{\omega}}, & \omega \in [\underline{\omega}, \overline{\omega}] \\ 0, & \text{otherwise} \end{cases}, \tag{41}$$

is near optimal. Here we take the frequency range as $\underline{\omega} = \underline{\theta}$ and $\overline{\omega} = \overline{\theta}$. In fact, it has been proven that the performance of bandlimited '$1/f$' noise is (at most) a factor of 2:1 away from the performance of the true robust optimal test signal. This is confirmed in Table 2 where different experiments are compared.

Note that the single sinusoid at $\omega = 1$ is the nominal test signal if we take the nominal parameter value as the geometric mean of $\theta_{min}$ and $\theta_{max}$.

We see from Table 2 that $1/f$ noise is indeed an excellent input for robust experiment design. This is further supported by recent research reported in [39]. In practice, it is also desirable to keep the amplitude of the test signal small. Thus, one may be interested in generating binary signals having a (band-limited) $1/f$ spectrum. Methods for designing such test signals are described in [41,8].

## 8    Estimation Procedures

We have argued above that the best experiment (when the ouput variations are constrained) is a closed loop one. This leads to the obvious follow up question "How should we estimate the parameters from closed loop data?". This issue has led to considerable consternation by practitioners. Indeed, some believe that closed loop identification is extremely difficult and often impossible. We will show below that, in fact, closed loop identification, when carried out properly, is no more difficult than open loop identification.

Thus, consider the closed loop shown in Figure 1. To illustrate the ideas we assume that the process of interest has the following linear model (extensions to the nonlinear case are also possible):

$$
\begin{aligned}
y_t &= G_o u_t + v_t \\
v_t &= H_o w_t
\end{aligned}
\tag{42}
$$

where $G_o$ and $H_o$ are linear transfer functions, and $w_t$ is a sequence of independent random variables with zero mean value, variance $\sigma_w^2$. We also, assume that $y_t$ and $u_t$ are jointly quasi-stationary, that the model $G$ for $G_o$ belongs to the family of models $G(\theta)$, and that $H$ (the model for $H_o$) belongs to a family of models $H(\theta)$.

The equations describing the system in Figure 1 are

$$
u_t = C(q^{-1})[r_t - y_t] \tag{43}
$$
$$
y_t = G_o(q^{-1})u_t + v_t \tag{44}
$$
$$
v_t = H_o(q^{-1})w_t \tag{45}
$$

where $w_t$ is a white noise sequence. The closed loop system is assumed to be stable. We further assume that at least one of the following two conditions holds:

(1) There is a delay in both the process and the model $(G_o(0) = G(0) = 0)$ and
(2) the true controller is strictly proper $(C(0) = 0)$.

Also, we normalize the true noise transfer function $H_o(q^{-1})$, and the model $H(q^{-1})$ by requiring that

$$H_o(0) = H(0) = 1 \tag{46}$$

The literature on identification basically offers two choices for closed loop identification [6]

- Direct: Here one treats $u_t$ and $y_t$ as if they were in open loop and estimates $G$ directly.
- Indirect [7]: Here the relationship between $y_t$ and $r_t$ is modeled and then $G_o$ is obtained from this model. In the linear case, we use the model

$$y_t = \frac{G_o C}{1 + G_o C} r_t + \frac{1}{1 + G_o C} v_t \tag{47}$$

and then extract an estimate of $G_o$. Indeed (47) can be simply thought as a potential model structure with some known parts (the parameters in $C$) and some parts to be estimated (the parameters in $G_o$). This is known as a Taylor made parameterization.

Each of the above approaches has advantages and disadvantages. Specifically,

(1) Direct identification is impossible with open loop unstable systems (in Box Jenkins form) since one has no way to ensure that the unstable initial condition response remains bounded. Also, direct identification is sensitive to being able to accurate specification of the noise model. Indeed, we will show below that errors in the noise model lead to bias.
(2) Indirect identification is, on the other hand, sensitive to the fidelity of the controller $C$. Thus, errors in the controller, e.g. due to saturation, will cause bias errors.

---

[6] There are some other alternatives in the available literature such as joint input/output identification. However, all of them assume perfect knowledge of the controller. Moreover, in most cases the controller must be linear.

[7] Note that *indirect* identification is equivalent to *direct* identification of a (closed loop) model having a particular parameterization.

We take a short diversion to review known results [16,32] on identification via PEM's. We assume that the system under study has input $\bar{u}_t$ and output $y_t$. (The specific form of $\bar{u}_t$ and its relationship to $u_t$ will be described later.) We conceptually model the relationship between $\bar{u}_t$ and $y_t$ by

$$y_t = G\bar{u}_t + Hw_t \tag{48}$$

where $w_t$ is notionally "white noise" and $G$ and $H$ are (independently) parameterized transfer functions. The PEM typically uses a cost function of the form [32]:

$$V = \frac{1}{N}\sum_{t=1}^{N}\epsilon_t^2 \tag{49}$$

where $\epsilon_t$ denotes the *prediction error* given by:

$$\epsilon_t = H^{-1}[y_t - G\bar{u}_t] \tag{50}$$

The following results are standard for PEM identification (see [16,32]):

**Lemma 7** *The cost function $V(\theta)$ converges, almost surely, to*

$$\bar{V}(\theta) = \bar{E}\{\epsilon_t^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_\epsilon(w)dw \tag{51}$$

□□□

**Lemma 8** *The prediction error spectrum satisfies the following:*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_\epsilon = \sigma_w^2 + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{|H|^2}\begin{bmatrix}\tilde{G} & \tilde{H}\end{bmatrix}\Phi_X\begin{bmatrix}\tilde{G}^* \\ \tilde{H}^*\end{bmatrix} \tag{52}$$

*where $*$ denotes complex conjugate and*

$$\tilde{G} = G_o - G \tag{53}$$

$$\tilde{H} = H_o - H \tag{54}$$

$$\Phi_X = \begin{bmatrix}\Phi_{\bar{u}} & \Phi_{\bar{u}w} \\ \Phi_{w\bar{u}} & \sigma_w^2\end{bmatrix} \tag{55}$$

□□□

Using the result (52) it is easy to see that if the spectrum $\Phi_X$ is a positive definite matrix for all frequencies, and that the true model ($G_o$ and $H_o$) are contained in the family of models ($G(\theta)$ and $H(\theta)$) a consistent estimate is obtained. This condition has been called an *informative* experiment [16,32]. We also have the following result [16,32]:

**Lemma 9** *The prediction error spectrum satisfies the following:*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_\epsilon = \sigma_w^2 + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{|H|^2}\left[|\tilde{G} + \tilde{H}\Phi_{w\bar{u}}\Phi_{\bar{u}}^{-1}|^2\Phi_{\bar{u}} + |\tilde{H}|^2\Phi_w^r\right] \tag{56}$$

*where $\tilde{G}$, $\tilde{H}$ are as in (53), (54), and*

$$\Phi_w^r = \sigma_w^2 - \Phi_{w\bar{u}}\Phi_{\bar{u}}^{-1}\Phi_{\bar{u}w} \tag{57}$$

□□□

Using (56), it can be readily seen that, when we have an erroneous noise model, the resulting estimate of $G_o$ will have an asymptotic bias given by:

$$B_\theta = \tilde{H}\Phi_{w\bar{u}}\Phi_{\bar{u}}^{-1} \tag{58}$$

where $\tilde{H}$ is as in (54). Here, we have assumed that $H$, and $G$ are independently parameterized. This, also covers the case when a fixed noise model is used. Using equation (58) we see that it is possible to obtain a consistent estimate in the following two cases:

- When a sufficiently rich family of noise models $H(\theta)$ is used such that $H = H_o$ is achievable.
- When the cross-spectrum $\Phi_{w\bar{u}}$ is (near) zero. (This is always the case with open loop data but will generally be false for closed loop data when a direct identification method is used).

Equation (56) also gives insight when the system model is misspecified i.e. when there does not exist a $G$ in the model class such that $\tilde{G}$ can be zero. In this case, minimization of (51) can be viewed as an approximation problem where the "weighted distance" between $G$ and $G_o + \tilde{H}\Phi_{wu}\Phi_u^{-1}$ is minimized over the given model class.

## 9 Virtual Closed Loop Identification

We have seen in section 8 that both direct and indirect closed loop identification have potential robustness issues. In this section, we describe a robustified procedure that combines the best features of direct and indirect identification.

To explain the idea, consider a hypothetical controller, $\bar{C}$, which is linear and has the transfer function

$$\bar{C}(q^{-1}) = \frac{P(q^{-1})}{L(q^{-1})} \tag{59}$$

We use a fractional representation and express (59) in the equivalent form:

$$\bar{C}(q^{-1}) = \frac{P(q^{-1})/E(q^{-1})}{L(q^{-1})/E(q^{-1})} \tag{60}$$

where $E(q^{-1})$ is any stable polynomial. We also write

$$N(q^{-1}) = E(q^{-1}) - L(q^{-1}) \tag{61}$$

Next we form a filtered version of the reference signal via the following stable operations on the measured signals $u_t$ and $y_t$:

$$\bar{u}_t = u_t - \frac{N(q^{-1})}{E(q^{-1})} u_t + \frac{P(q^{-1})}{E(q^{-1})} y_t \tag{62}$$

We can now use the *virtual input*, $\bar{u}_t$, as a simple mechanism for choosing an identification algorithm that lies anywhere in the range from direct to indirect. To illustrate, we consider two extreme cases.

(1) When $\bar{C} = \frac{P}{L}$ is chosen as the true control law, then

$$u_t = \frac{P(q^{-1})}{L(q^{-1})} [y_t - r_t] \tag{63}$$

$$\bar{u}_t = u_t - \frac{N}{E} u_t + \frac{P}{E} y_t \tag{64}$$

Substituting (63) into (64) gives

$$\bar{u}_t = - \left(1 - \frac{N}{E}\right) \left(\frac{P}{L}\right) [y_t - r_t] + \frac{P}{E} y_t \tag{65}$$

$$= -\frac{P}{E} [y_t - r_t] + \frac{P}{E} y_t \tag{66}$$

$$= \frac{P}{E} r_t \tag{67}$$

Hence, in this case, $\bar{u}_t$ is simply a filtered version of the reference input. Thus, if we use direct identification the system linking $\bar{u}_t$ to $y_t$, we are, in effect, using *indirect identification* between the filtered reference signal $\frac{P}{E} r_t$ and the output $y_t$.
(2) Alternativelly, say we choose $E(q^{-1}) = L(q^{-1})$ (i.e. $N(q^{-1}) = 0$) and $P(q^{-1}) = 0$, then from (62) we have that $\bar{u}_t = u_t$. Hence, in this case, $\bar{u}_t$ is simply the true plant input. Thus, fitting a model between $\bar{u}_t$ and $y_t$ corresponds to the usual *direct identification* method.

In the sequel, we will consider alternatives that lie between the above extremes. Our principal interest here is when $\bar{C}(q^{-1})$ is **not** the true controller. Indeed,

all we require is that $\bar{C}(q^{-1})$ stabilizes the system if it were to be applied to the system (even if it is, in fact, not in the real closed loop!). When $\bar{C}(q^{-1})$ is *not* the true controller, it is still easy to compute the equations linking $\bar{u}_t$ to $y_t$. Specifically, we have from (62) that $y_t$ and $\bar{u}_t$ are related by the following *Virtual Closed Loop* (where we treat $\bar{u}_t$ as a given signal).

$$
\begin{aligned}
&\textbf{Plant model:} \qquad y_t = G(q^{-1})u_t + v_t \\
&\textbf{Virtual feedback:}\ u_t = \frac{N(q^{-1})}{E(q^{-1})}u_t - \frac{P(q^{-1})}{E(q^{-1})}y_t + \bar{u}_t
\end{aligned}
\tag{68}
$$

where $\{v_t\}$ denotes the noise sequence and is assumed to satisfy $v_t = H(q^{-1})w_t$ for $\{w_t\}$ a white sequence.

**Remark 10** *The plant model $G(q^{-1})$ in (68) could be unstable, nonlinear etc. However, for (68) to be suitable for identification it is necessary that $\bar{C}(q^{-1})$ stabilizes the plant model.* $\triangledown\triangledown\triangledown$

**Remark 11** *Equations in (68) simply represent a particular parameterized model linking $\bar{u}_t$ to $y_t$ where parts of the model are fixed and known (namely $N(q^{-1})$, $E(q^{-1})$ and $P(q^{-1})$) and parts are unknown (namely the parameters in $G(q^{-1})$). See (71) below.* $\triangledown\triangledown\triangledown$

Finally, one can ask where do we get $\bar{u}_t$ from to drive (68). The answer is again provided by (62) i.e.

$$
\bar{u}_t = u_t - \frac{N(q^{-1})}{E(q^{-1})}u_t + \frac{P(q^{-1})}{E(q^{-1})}y_t
\tag{69}
$$

**Remark 12** *When $\bar{C}(q^{-1})$ is **not** the true controller, then there will exist residual correlations between the noise and $\bar{u}_t$. This will, in turn, lead to residual bias errors if the associated noise model is erroneous. However, we can see that one gains the advantages of both direct identification (i,e. $\bar{u}_t$ can be treated as a known input into a particularly parameterized plant) and indirect identification can be performed since the virtual controller is **exactly** known.* $\triangledown\triangledown\triangledown$

We next show that the use of a virtual closed loop has a beneficial effect on estimation accuracy provided $\bar{C}$ is chosen appropriately. In particular, we wish to study the effect of the virtual closed loop on the correlation between $w$ and $\bar{u}_t$ and hence on the bias expression given in (58). The key difference between direct identification and virtual closed loop identification is that direct identification is based on the model

$$
y_t = G_o u_t + H_o w_t
\tag{70}
$$

whereas, the virtual closed loop identification is based on (68) which can also

be expressed as

$$y_t = \frac{G\frac{E}{L}}{1 + G\bar{C}}\bar{u}_t + \frac{H}{1 + G\bar{C}}w_t = T^y\bar{u}_t + \bar{H}w_t \tag{71}$$

The key problems with (70) are

(1) One cannot run the model $G$ in open loop if it is unstable.
(2) The presence of feedback means that $u_t$ and $w_t$ are potentially highly correlated which is a source of robustness and bias problems.

On the other hand, the model $\frac{G\frac{E}{L}}{1+G\bar{C}}$ can be stable even if $G$ is unstable. Also, the correlations between $\bar{u}$ and $w_t$ are potentially less than between $u_t$ and $w_t$. The implications of the above observations are explained in detail in [3,2,1]. In particular we argue in [2,1], that the bias resulting from virtual closed loop identification has the form

$$\hat{G} = G_o\bar{\lambda} - \bar{C}^{-1}(1 - \bar{\lambda}) \tag{72}$$

where

$$\bar{\lambda} = \frac{1}{1 + \kappa C_\Delta H_\Delta} \tag{73}$$

where $\kappa$ is a frequency dependent parameter, $C_\Delta$ denotes the relative error in the virtual controller i.e. the true controller $C_l$ is expressed as

$$C_l = \bar{C}(1 + C_\Delta) \tag{74}$$

and where $H_\Delta$ denotes the error in the noise model, i.e.

$$H_\Delta = \bar{H} - \bar{H}_o \tag{75}$$

Also note that $\kappa$ is inversely proportional to the size of the reference signal spectrum.

**Remark 13** *We see from (72) that the bias in the estimate will be small if any of the following three conditions is satisfied*

*(1) $H_\Delta$ is small (i.e. small noise model errors)*
*(2) $C_\Delta$ is small (i.e. small errors between $\bar{C}$ and the true controller).*
*(3) The reference signal dominates disturbances.*

Thus, the virtual closed loop method achieves the best features of both direct and direct identification. We illustrate by a simple example:

Consider the following system:

$$G_o(q^{-1}) = \frac{bq^{-1}}{1 - aq^{-1}} \tag{76}$$

with $a = 0.8$, and $b = 0.2$. This system is operated in closed loop with the following nominal linear controller:

$$C_l(q^{-1}) = \frac{0.3}{1 - q^{-1}} \tag{77}$$

However, the true controller operates in such a fashion that the input signal saturates such that $|u_t| \leq 5$. Thus, the true controller, $C$, is actually nonlinear. The reference signal $r_t$ is taken to be zero mean Gaussian white noise process of variance $\sigma_r^2 = 30$. The noise $w_t$ is taken to be zero mean Gaussian white noise process of variance $\sigma_w^2 = 0.1$ ($\sigma_w^2 \approx 5$). The true noise model is given by

$$H_o(q^{-1}) = \frac{1}{1 + d_1 q^{-1} + d_2 q^{-2} + d_3 q^{-3} + d_4 q^{-4}} \tag{78}$$

where $d_1 = -1.992$, $d_2 = 2.202$, $d_3 = -1.841$, and $d_4 = 0.8941$.

In order to identify the process, we assume (incorrectly) that the noise model, $H$, is an autoregressive AR model of first order. (This is, after all, not unreasonable, since the system is a first order transfer function). We then identify the process by using direct identification, and also by using the virtual closed loop method with the virtual controller equal to the true controller (but without saturation). For the Virtual Closed Loop (VCL) case, we use a PEM to identify the virtual closed loop transfer function. (We use a filter $E = 1 - 0.95q^{-1}$). We then extract the estimate of the open loop model by using the known relationship between the open and closed loop parameters. We carried out 400 Monte-Carlo simulations for 3000 data points. The average models so obtained are presented in table 3. We can see that, for this example, the model obtained by using direct identification is not satisfactory – there is a significant model misfit due to a poor noise model. Also, the model found by indirect identification is poor. This is because the controller occasionally saturates. On the other hand, the model obtained by using the virtual closed loop method is extremely close to the true model. The reason for this improvement is that the use of a virtual closed loop, reduces the bias due to the reduction in the value of $\beta$. (see equation (72)).

## 10    Errors in variables

A problem which is closely related to that of closed loop identification is that of Errors In Variables (EIV)– see ([43,7] and the references therein). The core idea in EIV estimation is that one does not know the exact form of the system input but, maybe, has a noisy measurement of it. Thus, the system model may be as in (1) but we only measure $u_t^m$ where

$$u_t^m = u_t + \epsilon_t \tag{79}$$

To illustrate the difficulties arising from EIV's, say that $v_t$ in (1) is zero and that $G_o$ is a scalar. Then, clearly

$$G_o = \frac{y_t}{u_t} \tag{80}$$

However, if we replace $u_t$ by $u_t^m$ and use (79), then

$$\hat{G} = \frac{y_t}{u_t + \epsilon_t} = \frac{G_o}{1 + \frac{1}{\beta}} \tag{81}$$

where $\beta$ is a form of "signal to noise ratio" i.e.

$$\beta = \frac{u_t}{\epsilon} \tag{82}$$

The above idea can readily extended to general estimation procedures e.g. PEM's etc.

We may summarize the above discussion by noting that significant bias errors occur due to EIV issues whenever $\beta$ (roughly the input signal to input noise ratio) is small.

Special ways of analyzing data to minimize the impact of EIV's can be found in the literature (see e.g. [43,7,27,6]). Again, one should be aware of the potential of sensitivity to the assumptions (i.e. lack of robustness) in these procedures. One simple idea is to select those parts of the data where the EIV issue is known to be small. For example, frequency and time selectivity has been successfully used in the area of electromagnetic minerals exploration (see [28,27]). Note again the idea of discarding data to enhance robustness. b

## 11    Conclusions

This paper has argued that robustness (i.e. insensitive to assumptions) is a key aspect of system identification. This is not surprising in view of the fact

that robustness plays a central role in many other areas including control. We have shown that, if robustness is ignored, then models obtained via blond application of identification procedures can bear little resemblance to reality. On the other hand, we have argued that often robustness can be achieved by relatively small modification to the algorithm e.g. by discarding some parts of the data. The ideas have been illustrated with special emphasis on:

- selecting a model class,
- the effect of sampling on model aliasing,
- limited bandwidth estimation,
- robust experiment design,
- robust estimation for closed loop data, and
- robust estimation in the presence of errors in variables.

The ideas have been illustrated by several real world case studies. Finally, we note that the ultimate test of robustness of an estimated model is through validation (i.e. using the model in practice on data collected independently of the identification experiment).

# References

[1] J. C. Agüero. *System identification methodologies incorporating constraints.* PhD thesis, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2005.

[2] J. C. Agüero and G. C. Goodwin. Virtual closed loop identification. Technical report, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2004.

[3] J. C. Agüero and G. C. Goodwin. Virtual closed loop identification: A Subspace Approach. In *Proceedings of the 43rd IEEE Conference on Decision and Control, CDC*, pages 364–369, Atlantis, Paradise Island, Bahamas, 2004.

[4] J. C. Agüero and G. C. Goodwin. On the optimality of open and closed loop experiments in system identification. In *45th IEEE Conference on Decision and Control (CDC)*, San Diego, California, 2006.

[5] J. C. Agüero and G. C. Goodwin. Choosing between open and closed loop experiments in linear system identification. *IEEE Transactions on Automatic Control*, 52(8):1475–1480, 2007.

[6] J. C. Agüero and G. C. Goodwin. Identifiability of EIV dynamic systems with non-stationary data. In *IFAC World Congress*, Seoul, Korea, 2008.

[7] J. C. Agüero and G. C. Goodwin. Identifiability of errors in variables dynamic systems. *Automatica*, 44(2):371–382, 2008.

[8] J. C. Agüero, J. Welsh, and O. J. Rojas. On the optimality of binary-open-loop experiments to identify Hammerstein systems. In *American Control Conference*, New York, USA, 2007.

[9] J. C. Agüero, J. I. Yuz, and G. C. Goodwin. Frequency domain identification of MIMO state space models using the EM algorithm. In *European Control Conference ECC*, Kos, Greece, 2007.

[10] B. D. O. Anderson and R. L. Kosut. Adaptive robust control: On-line learning. In *Proceedings of the 30th Conference of Decision and Control, CDC*, Brighton, UK, 1991.

[11] K. J. Åström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.

[12] G. E. P. Box and G. C. Tiao. A bayesian approach to some outlier problems. *Biometrika*, 55(1):119–129, mar 1968.

[13] L. Davies and U. Gather. Robust statistics. In J. E. Gentle, W. Härdle, and Y. Mori, editors, *Handbook of Computational Statistics*. Springer, 2004.

[14] R. F. Engle. Band spectrum regression. *International Economic Review*, 15(1):1–11, feb 1974.

[15] A. Feuer and G.C. Goodwin. *Sampling in Digital Signal Processing and Control*. Birkhäuser, Boston, 1996.

[16] U. Forssell and L. Ljung. Closed loop identification revisited. *Automatica*, 35:1215–1241, 1999.

[17] G. C. Goodwin, S. F. Graebe, and M. E. Salgado. *Control System Design*. Prentice Hall, Upper Saddle River, NJ, 2001.

[18] G. C. Goodwin and R. Payne. *Dynamic System Identification: Experiment design and data analysis*. Academic Press, 1977.

[19] G. C. Goodwin, J. I. Yuz, and J. C. Agüero. Relative error issues in sampled data models. In *IFAC World Congress*, Seoul, Korea, 2008.

[20] G. C. Goodwin, J. I. Yuz, and M. E. Salgado. Insights into the zero dynamics of sampled-data models for linear and nonlinear stochastic systems. In *European Control Conference ECC*, Kos, Greece, 2007.

[21] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust Statistics: The approach based on influence functions*. John Wiley and Sons, 1986.

[22] E. J. Hannan. *Multiple time series*. John Wiley and Sons, Inc., 1970.

[23] E. J.. Hannan and P. M. Robinson. Lagged regression with unknown lags. *Journal of the Royal Statistical Society. Series B (Methodological)*, 35(2):252–267, 1973.

[24] P. J. Huber. *Robust Statistics*. John Wiley and Sons, 1981.

[25] J. Kiefer. Optimum experiment designs. *Journal of the Royal Statistical Society, Series B*, 21(2):272–319, 1959.

[26] R. L. Kosut. Iterative adaptive control: Windsurfing with confidence. *Model identification and adaptive control. Edited by G. C. Goodwin*, pages 201–229, 2000.

[27] K. Lau, J. H. Braslavsky, J. C. Agüero, and G. C. Goodwin. Application of non-stationary EIV methods to transient electromagnetic mineral exploration. In *IFAC World Congress*, 2008.

[28] K. Lau, J. H. Braslavsky, and G. C. Goodwin. Errors-in-variables problems in transient electromagnetic mineral exploration. In *46th IEEE Conference on Decision and Control (CDC)*, New Orleans, LA, USA, 2007.

[29] W. S. Lee, B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut. On some key issues in the windsurfer approach to adaptive robust control;. *Automatica*, 31(11):1619–1636, 1995.

[30] W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y Mareels. A new approach to adaptive robust control. *International Journal of Adaptive Control and Signal Processing*, 7(5):183–211, 1993.

[31] L. Ljung. Some results on identifying linear systems using frequency domain data. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 3534–3538, New York, USA, 1993.

[32] L. Ljung. *System Identification: Theory for the user*. Prentice Hall, 2nd edition, 1999.

[33] T. McKelvey and L. Ljung. Frequency domain maximum likelihood identification. In *11th IFAC Symposium on System Identification*, pages 1741–1746, Fukuoka, Japan, 1997.

[34] R. Middleton and G.C. Goodwin. *Digital control and estimation: A unified approach*. Prentice-Hall International, 1990.

[35] R. Pintelon and J. Schoukens. ML identification of closed-loop systems in a specified frequency band. In *Proc. of 16th IFAC World Congress*, Prague, Czech Republic, July 2005.

[36] F. Pukelsheim. *Optimal Design of Experiments*. John Wiley, 1993.

[37] P. M. Robinson. Stochastic difference equations with non-integral differences. *Advances in Applied Probability*, 6(3):524–545, sep 1974.

[38] P. M. Robinson. On the errors-in-variables problem for time series. *Journal of Multivariate Analysis*, 19:240–250, 1986.

[39] C. R. Rojas, G. C. Goodwin, J. S. Welsh, and A. Feuer. Optimal experiment design with diffuse prior information. In *European Control Conference (ECC)*, Kos, Greece, 2007.

[40] C. R. Rojas, G. C. Goodwin, J. S. Welsh, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 2007.

[41] C. R. Rojas, J. S. Welsh, and G. C. Goodwin. A receding horizon algorithm to generate binary signals with a prescribed autocovariance. In *American Control Conference*, New York, USA, 2007.

[42] M. Siotani. Some applications of Loewner's ordering on symmetric matrices. *Annals of the Institute of Statistical Mathematics*, 19:245–259, 1967.

[43] T. Söderström. Errors-in-variables methods in system identification. *Automatica*, 43(6):939–958, 2007.

[44] S. R. Weller, W. Moran, B. Ninness, and A. D. Pollington. Sampling zeros and the euler-frobenius polynomials. *IEEE Transactions on Automatic Control*, 46(2):340–343, 2001.

[45] J. I.. Yuz. *Sampled-Data models for linear and non-linear systems*. PhD thesis, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2005.

[46] J. I. Yuz and G. C. Goodwin. Sampled-data models for stochastic nonlinear systems. In *14th IFAC symposium on system identification*, Newcastle, Australia, 2006.

[47] J.I. Yuz and G.C. Goodwin. *Continuous-Time System Identification. Edited by H. Garnier and L. Wang.*, chapter Robust Identification of Continuous-Time Systems from Sampled Data. Springer, 2008.
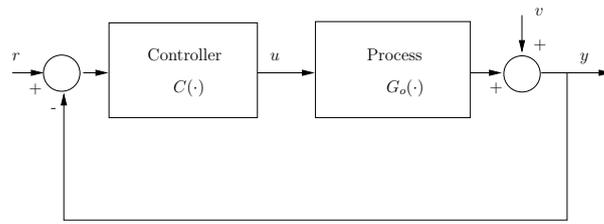
Fig. 1. Closed Loop used in Identification
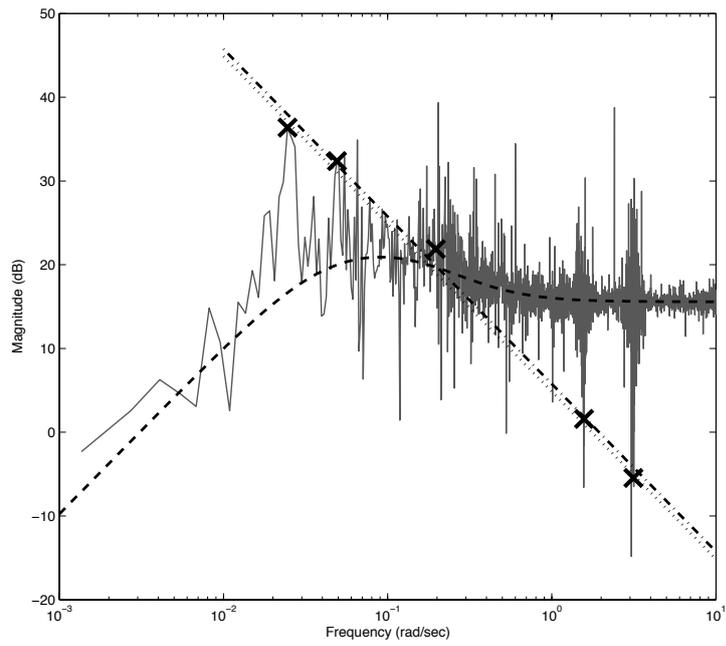
Fig. 2. Results of the experiment in the frequency domain. ETFE (solid line), $-1/C$ (dashed line), Theoretical model (dotted line), Estimated model with the reduced data (dash-dot line) and $\times$ - location of excitation frequencies.
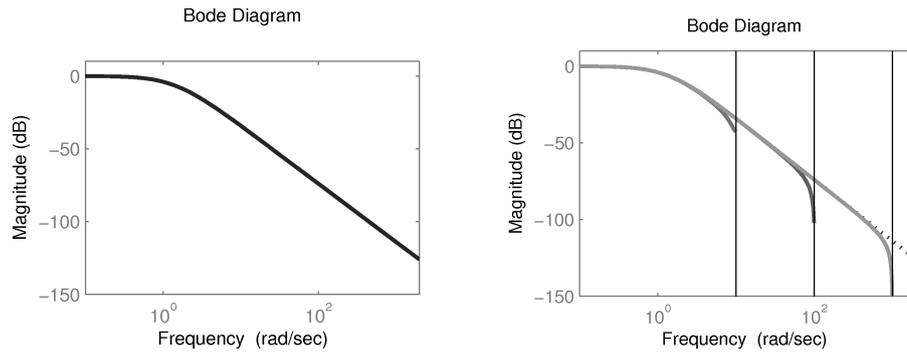
Fig. 3. Continuous (left)- and discrete-time (right) frequency response magnitudes (vertical lines denote folding frequency of $\pi/\Delta$ for 3 different values of $\Delta$).

Table 1

Parameter estimates for a linear system

| | Parameters | | Estimates | | |
|---|---|---|---|---|---|
| | CT | Exact DT | SDRM | MIFZ | MIPZ |
| $\alpha_1$ | 3 | 2.923 | 2.8804 | 2.9471 | 2.9229 |
| $\alpha_0$ | 2 | 1.908 | 1.9420 | 1.9090 | 1.9083 |
| $\beta_1$ | – | 0.0305 | – | $\frac{\beta_0 \Delta}{2} = 0.03$ | 0.0304 |
| $\beta_0$ | 2 | 1.908 | 0.9777 | 1.9090 | 1.9083 |

Table 2
Relative Values of Cost for the Different Input Signals

|  | $\max\limits_{\theta \in \Theta} [\theta^2 \overline{M}(\theta, \phi_u)]^{-1}$ |
|---|---|
| Sinusoid at $\omega = 1$ | 7.75 |
| Bandlimited white noise | 12.09 |
| Bandlimited '$1/f$' noise | 1.43 |
| Min-max optimal input | 1.00 |

Table 3
Monte-Carlo comparison between VCL (PEM) and Direct (PEM) method for $N = 3000$ data points

| $\sigma_w^2 = 0.1$ | $a$ | $b$ |
|---|---|---|
| True Values | 0.8 | 0.2 |
| Mean Value VCL Identification | 0.7955 | 0.1990 |
| Mean Value Direct Identification | 0.8634 | 0.3041 |
| Mean Value Indirect Identification | 0.7746 | 0.1515 |