Bias, Variance and Optimal Experimental Design: Some Comments on Closed Loop Identification

Lennart Ljung and Urban Forssell

Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden
WWW: `http://www.control.isy.liu.se`
Email: `ljung,ufo@isy.liu.se`

March 3, 1999

REGLERTEKNIK

AUTOMATIC CONTROL

LINKÖPING

Report no.: LiTH-ISY-R-2100
Submitted to "Perspectives in Control, a tribute to I.D. Landau, Paris, June 1998"

Bias, Variance and Optimal Experiment Design: Some Comments on Closed Loop Identification

Lennart Ljung and Urban Forssell

Division of Automatic Control
Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden
E-mail: ljung@isy.liu.se, ufo@isy.liu.se
URL: http://www.control.isy.liu.se/

March 3, 1999

Abstract

In this contribution we shall describe a rather unified way of expressing bias and variance in prediction error estimates. The emphasis is on systems operating in closed loop. We shall describe the identification criterion function in the frequency domain. The crucial entity is the joint spectrum of input and noise source. Different factorizations of this spectrum give different insights into the bias mechanisms of closed loop identification. It will be shown that so called *indirect identification* is the answer to the question of how to obtain consistent estimates of the dynamics part, even with an erroneous noise model. We also consider optimal design of experiments that seek to minimize the weighted variance of the dynamics estimate. It is shown that open loop experiments are optimal if the input power is constrained. However for any criteria that involve any kind of constraints on the output power, closed loop experiments will be optimal. The optimal regulator does not depend on the weighting function in the criterion to be minimized.

1 Introduction and Setup

Identification of systems operating in closed loop have long been of interest. The reasons are that many systems are not allowed to operate in open loop during an identification experiment. Adaptive control is another situation, where closed loop identification issues naturally arise. See among many references also [15]. The recent interest in so called identification for control has also spurred new methods and results, [4, 13, 18],[6] and [17].

See, among many general references on closed loop identification, [8],[16],[1], [9],[3],[2, 5, 11] and [12].

We shall consider identification of a linear system in a traditional prediction error setup. See [14] for all technical details. The true system is supposed to be described by

$$y(t) = G_0(q)u(t) + H_0(q)e(t) \tag{1}$$

where q is the delay operator, u is the input, y is the output and e is white noise with covariance matrix Λ_0.

The system is operating under arbitrary feedback, but we assume that all signals are quasi-stationary, so that the spectrum of

$$\zeta = \begin{bmatrix} u \\ e \end{bmatrix} \tag{2}$$

is well defined, and denoted by

$$\Phi_\zeta(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \Lambda_0 \end{bmatrix} \tag{3}$$

The system is identified within the model structure

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{4}$$

G will be called the dynamics model and H the noise model. The parameter θ is estimated by

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{5}$$

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta) \tag{6}$$

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{7}$$

Here Λ is a symmetric, positive definite weighting matrix.

We shall discuss the asymptotic properties of $\hat{\theta}_N$ in the sequel.

2 Expressions for the Data Spectrum

The data spectrum Φ_ζ plays an important role in the analysis, and we shall therefore collect some results on it here. We shall introduce the following spectrum

$$\Phi_u^r = \Phi_u - \Phi_{ue}\Lambda_0^{-1}\Phi_{eu} \tag{8}$$

where we suppress the argument ω. This is the spectrum of that part of the input u, that cannot be estimated from e by a linear, time-invariant filter.

Similarly we introduce

$$\Phi_e^r = \Lambda_0 - \Phi_{eu}\Phi_u^{-1}\Phi_{ue} \tag{9}$$

The data spectrum Φ_ζ can now be written as

$$\begin{bmatrix} \Phi_u & \Phi_{ue} \\ \Phi_{eu} & \Lambda_0 \end{bmatrix} = \begin{bmatrix} I & \Phi_{ue}\Lambda_0^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Phi_u^r & 0 \\ 0 & \Lambda_0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \Lambda_0^{-1}\Phi_{eu} & I \end{bmatrix} \tag{10}$$

$$= \begin{bmatrix} I & 0 \\ \Phi_{eu}\Phi_u^{-1} & I \end{bmatrix} \begin{bmatrix} \Phi_u & 0 \\ 0 & \Phi_e^r \end{bmatrix} \begin{bmatrix} I & \Phi_u^{-1}\Phi_{ue} \\ 0 & I \end{bmatrix} \tag{11}$$

$$\tag{12}$$

From these factorization results we also find an expression for the inverse

$$\Phi_\zeta^{-1} = \begin{bmatrix} \Phi_u & \Phi_{ue} \\ \Phi_{eu} & \Lambda_0 \end{bmatrix}^{-1} = \begin{bmatrix} (\Phi_u^r)^{-1} & -(\Phi_u^r)^{-1}\Phi_{ue}\Lambda_0^{-1} \\ -\Lambda_0^{-1}\Phi_{eu}(\Phi_u^r)^{-1} & (\Phi_e^r)^{-1} \end{bmatrix} \tag{13}$$

In case the regulator is linear, time-invariant, we have

$$u(t) = r(t) - K(q)y(t) \tag{14}$$

where $K(q)$ is a linear regulator of appropriate dimensions and where the reference signal $\{r(t)\}$ is independent of $\{e(t)\}$. We then have the following expressions. Let S and S^i the the output and input sensitivity functions:

$$S_0 = (I + G_0K)^{-1}, \quad S_0^i = (I + KG_0)^{-1} \tag{15}$$

Then

$$\Phi_u = S_0^i\Phi_r(S_0^i)^* + KS_0\Phi_v S_0^* K^* \tag{16}$$

where Φ_r is the spectrum of the reference signal and $\Phi_v = H_0\Lambda_0 H_0^*$ the noise spectrum. Superscript $*$ denotes complex conjugate transpose. We shall denote the two terms in (16)

$$\Phi_u^r = S_0^i\Phi_r(S_0^i)^* \tag{17}$$

and

$$\Phi_u^e = KS_0\Phi_v S_0^* K^* = S_0^i K\Phi_v K^*(S_0^i)^* \tag{18}$$

The cross spectrum between u and e is

$$\Phi_{ue} = -KS_0H_0\Lambda_0 = -S_0^i KH_0\Lambda_0 \tag{19}$$

3 The Main Expression

From standard asymptotic theory we know that

$$\hat{\theta}_N \to \arg\min_\theta \overline{V}(\theta) \quad \text{w.p.1 as} \quad N \to \infty \tag{20}$$

where

$$
\begin{aligned}
\overline{V}(\theta) &= \overline{E} V_N(\theta, Z^N) \\
&= \int_{-\pi}^{\pi} \operatorname{tr}\left\{\Lambda^{-1} H_\theta^{-1} \begin{bmatrix}\Delta G_\theta & \Delta H_\theta\end{bmatrix} \Phi_\zeta \begin{bmatrix}\Delta G_\theta^* \\ \Delta H_\theta^*\end{bmatrix} (H_\theta^{-1})^*\right\} d\omega
\end{aligned}
\tag{21}
$$

Here we have introduced the simplified notation

$$
\begin{aligned}
\Delta G_\theta &= G_0(e^{i\omega}) - G(e^{i\omega}, \theta) \\
\Delta H_\theta &= H_0(e^{i\omega}) - H(e^{i\omega}, \theta)
\end{aligned}
$$

To see (21) we rewrite (7) using (1) as

$$
\begin{aligned}
\varepsilon &= H_\theta^{-1}(y - G_\theta u) = H_\theta^{-1}(G_0 - G_\theta)u + H_\theta^{-1} H_0 e \\
&= H_\theta^{-1}(G_0 - G_\theta)u + (H_\theta^{-1} H_0 - I)e + e \\
&= H_\theta^{-1}[(G_0 - G_\theta)u + (H_0 - H_\theta)e] + e
\end{aligned}
$$

We then use that H_0 and H_θ are both monic (so that the difference starts with a delay) and that $e(t)$ is independent of $(G_0 - G_\theta)u(t)$, which is the case if either the regulator or the model/system contains a delay. Parseval's relationship then gives (21).

4 Identifiability, Bias and the Indirect Method

4.1 Consistency and Identifiability

Identifiability essentially means that the estimate is consistent and will converge to the true system, when the system is contained in the model structure. This is a joint requirement on the model structure and on the experiment condition, i.e., the data spectrum Φ_ζ.

The basic expression (20) with (21) shows that

> *consistency and identifiability follows when no model in the structure* $\begin{bmatrix}\Delta G_\theta & \Delta H_\theta\end{bmatrix}$ *lies in the null space of* Φ_ζ*.*

A sufficient condition for identifiability is thus that the data spectrum is positive definite (non-singular) almost everywhere. From (10) we see that this happens if and only if (we assume Λ_0 to be non-singular)

$$
\Phi_u^r(\omega) > 0, a.e.\omega \tag{22}
$$

In [14] this is termed that the experiment is *informative enough.* With the definition (8) this means that there should be a full rang part of u that cannot be estimated from e by a linear, time invariant filter. It essentially means that the regulator should not be linear, time-invariant with no extra signals (disturbances or setpoints). A persistently exciting setpoint (reference signal), a non-linear regulator or a time-varying regulator will basically secure (22).

Constraining the model structure may however secure identifiability even when (22) does not hold. Suppose, for example that the noise model H_θ is fixed to the true value H_0 ($\Delta H_\theta = 0$), and that the dynamics model structure G_θ contains the true system G_0. Then we see directly from (21) that consistency follows as soon as $\Phi_u(\omega) > 0$, $a.e.\, \omega$. However, it is essential that the noise model is the true one, otherwise the limit of G_θ will still be unique, but biased.

4.2 Bias Distribution

To more clearly see bias issues and lack of identifiability we can use the factorization (11) to rewrite the convergence expression as

$$\hat{\theta}_N \to \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\big[\Lambda^{-1}\big\{H_\theta^{-1}[(G_0 + B_\theta - G_\theta)\Phi_u(G_0 + B_\theta - G_\theta)^* \\ + (H_0 - H_\theta)\Phi_e^r(H_0 - H_\theta)^*]H_\theta^{-*}\big\}\big]\, d\omega \quad \text{w. p. 1 as } N \to \infty \quad (23)$$

where

$$B_\theta = (H_0 - H_\theta)\Phi_{eu}\Phi_u^{-1} \quad (24)$$

In the SISO, linear regulator case, we can characterize the "size" of B as

$$|B_\theta|^2 = \frac{\Lambda_0}{\Phi_u}\frac{\Phi_u^e}{\Phi_u}|H_0 - H_\theta|^2 \quad (25)$$

(see (16) and (19).) For a fixed noise model $H_\theta = H_*$, we see that the dynamics model will approximate the biased frequency function $G_0 + B_*$ in a frequency domain norm determined by H_*^{-1} and Φ_u (The ratio $\Phi_u/|H_*|^2$ in the SISO case). We obtain an approximation of the correct dynamics in case $B = 0$, which means that the noise model is correct $H_* = H_0$, or the system operates in open loop: $\Phi_{eu} = 0$.

4.3 The Indirect Method

We saw above that the noise model has to be correct, in order to ensure consistency of the dynamics part. Let us now ask the question:

Suppose the dynamics model structure G_θ contains the true dynamics G_0, and that we are not interested in the noise characteristics. Is is then possible to obtain a consistent estimate of G?

To answer that question, we rewrite the basic expression (21) using the alternative factorization (10):

$$
\begin{aligned}
\overline{V}(\theta) = & \int_{-\pi}^{\pi} \operatorname{tr}\left[\Lambda^{-1} H_\theta^{-1}(\Delta G_\theta \Phi_u^r (\Delta G_\theta)^* H_\theta^{-*}\right] d\omega \\
& + \int_{-\pi}^{\pi} \operatorname{tr}\left[\Lambda^{-1} H_\theta^{-1}(\Delta G_\theta \Phi_{ue}\Lambda_0^{-1} + \Delta H_\theta)\Lambda_0 \right. \\
& \left. \times(\Delta G_\theta \Phi_{ue}\Lambda_0^{-1} + \Delta H_\theta)^* H_\theta^{-*}\right] d\omega
\end{aligned} \tag{26}
$$

To assure consistency even if the noise model is not correct, we should make the second term (the second integral) independent of ΔG_θ. We then expand the factor in the integral as follows (using the expression (19) for Φ_{ue})

$$
\begin{aligned}
& H_\theta^{-1}(G_\theta K S_0 H_0 - G_0 K S_0 H_0 + H_0) - I \\
& \quad = H_\theta^{-1}(G_\theta K - G_0 K + I + G_0 K) S_0 H_0 - I \\
& \quad = H_\theta^{-1}(I + G_\theta K) S_0 H_0 - I = \tilde{H}_\eta^{-1} S_0 H_0 - I
\end{aligned}
$$

where we introduced the noise model parameterization

$$
H_\theta = (I + G_\theta K)\tilde{H}_\eta = S_\theta^{-1}\tilde{H}_\eta \tag{27}
$$

where η is a parameterization, independent of θ. Here S_θ is the model sensitivity function, compare (15). With (27) we have thus achieved that the second integral of (26) is independent of the parameterization of G_θ. The first integral will thus determine to what the dynamics model converges: If $\tilde{H}$ is a fixed model we have

$$
\hat{\theta}_N \to \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\left[\Lambda^{-1}\left\{\tilde{H}^{-1}(S_\theta \Delta G_\theta)\Phi_u^r (S_\theta \Delta G_\theta)^* \tilde{H}^{-*}\right\}\right] d\omega \tag{28}
$$

With (17) we find that

$$
(S_\theta \Delta G_\theta)\Phi_u^r (S_\theta \Delta G_\theta)^* = (S_\theta \Delta G_\theta S_0^i)\Phi_r (S_\theta \Delta G_\theta S_0^i)^* \tag{29}
$$

with

$$
\begin{aligned}
S_0 G_0 - S_\theta G_\theta &= G_0 S_0^i - S_\theta G_\theta \\
&= S_\theta((I + G_\theta K)G_0 - G_\theta(I + K G_0))S_0^i = S_\theta \Delta G_\theta S_0^i
\end{aligned}
$$

This shows that the noise model parameterization (27) will fit the closed loop model $S_\theta G_\theta$ to the closed loop system $S_0 G_0$ in a norm that is determined by the reference signal spectrum Φ_r and the fixed noise model $\tilde{H}$.

In fact the parameterization (27) corresponds to a well known method for dealing with closed loop identification data: The predictor for

$$
y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{30}
$$

is

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (I - H^{-1}(q,\theta))y(t) \tag{31}$$

Using $u = r - Ky$ and inserting (27) we get

$$\begin{aligned}\hat{y}(t|\theta) =& \tilde{H}^{-1}(q,\theta)(I + G(q,\theta)K(q))^{-1}G(q,\theta)r(t) \\ &+ (I - \tilde{H}^{-1}(q,\theta))y(t)\end{aligned} \tag{32}$$

But this is exactly the predictor also for the closed-loop model structure

$$y(t) = (I + G(q,\theta)K(q))^{-1}G(q,\theta)r(t) + \tilde{H}(q,\theta)e(t) \tag{33}$$

Identifying the closed loop, and then solving for the open loop dynamics is called *the indirect method.* We have here derived this approach as the answer to the question of how to obtain consistent dynamics models, without dealing with a noise model. Of course, in the open loop case ($K = 0$) (27) tells us that this is achieved by letting the noise model be parameterized independently from the dynamics.

5 Asymptotic Variance

5.1 Parameter and Transfer Function Covariance

The classical result on the asymptotic distribution of the parameter estimates (cf [14], Chapter 9) is as follows. Suppose that the true system is contained in the model structure, and that the experimental conditions are such that identifiability is secured. Let the true parameters be denoted by θ_0. In the multi-output case we also assume that $\Lambda = \Lambda_0$. Then $\sqrt{N}(\hat{\theta}_N - \theta_0)$ converges in distribution to the normal distribution with zero mean and a covariance matrix

$$P_\theta = [\overline{V}''(\theta_0)]^{-1} \tag{34}$$

where $\overline{V}$ is defined by (21) and prime denotes differentiation w.r.t θ.

Let us from now on concentrate on the SISO case for notational simplicity and denote

$$T_\theta = \begin{bmatrix} G_\theta & H_\theta \end{bmatrix}, \quad \hat{T}_N = \begin{bmatrix} \hat{G}_N & \hat{H}_N \end{bmatrix} \tag{35}$$

Then, using (21) we can write

$$P_\theta = \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\Phi_v} T' \Phi_\zeta (T')^* d\omega \right]^{-1} \tag{36}$$

Here $\Phi_v = |H_0|^2 \Lambda_0$. From this expression and the factorizations (10) and (11) explicit expression can be found how Φ_r and K affect the parameter accuracy.

If we are interested in the covariance of $\hat{T}_N$, rather than the parameters, Gauss' approximation formula gives the expression

$$N \cdot \mathrm{Cov}\hat{T}_N(e^{i\omega}) = T'(e^{i\omega})^T$$
$$\times \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\Phi_v(\xi)} T'(e^{i\xi}) \begin{bmatrix} \Phi_u(\xi) & \Phi_{ue}(\xi) \\ \Phi_{eu}(\xi) & \Lambda_0 \end{bmatrix} T'(e^{-i\xi})^T d\xi \right]^{-1} T'(e^{-i\omega}) \tag{37}$$

5.2 Asymptotic Black Box Expressions

The expression (37) shows an intriguing symmetry, with the factors T' in "cancelling positions". In fact, suppose that parameterization of T has the following shift structure,

$$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$
$$\frac{d}{d\theta_k} T(q,\theta) = q^{-k+1} \frac{d}{d\theta_1} T(q,\theta)$$

which is satisfied by many black-box model parameterizations. Then, as n tends to infinity we have (see [14], Chapter 9):

$$\mathrm{Cov} \begin{bmatrix} \hat{G}_N(e^{i\omega}) & \hat{H}_N(e^{i\omega}) \end{bmatrix} \approx \frac{n}{N} \Phi_v(\omega) \begin{bmatrix} \Phi_u(\omega) & \Phi_{eu}(\omega) \\ \Phi_{ue}(\omega) & \Lambda_0 \end{bmatrix}^{-1} \tag{38}$$

n is here the "model order" and N is the number of data. From (13) we then find that

$$\mathrm{Cov} \left[\hat{G}_N(e^{i\omega}) \right] \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u^r(\omega)} \tag{39}$$

6 Optimal Experiment Design

In this section we will consider experiment design problems where the goal is to minimize a weighted norm of the covariance of G:

$$J = \int_{-\pi}^{\pi} \mathrm{Cov}\hat{G}_N(e^{-i\omega}) C(\omega) d\omega \tag{40}$$

The minimization shall be carried out with respect to the experiment design variables, which we take as K (the regulator) and Φ_r (the reference signal spectrum). Other equivalent choices are also possible, e.g., Φ_u and Φ_{ue} or K and Φ_u^r. To make the designs realistic we will also impose constraints on the input power or the output power, or both.

Consider the problem to minimize J given by (40) and using the asymptotic expression (39). The minimization is to be carried out under the constraint

$$\int_{-\pi}^{\pi} \{\alpha \Phi_u + (1-\alpha)\Phi_y\} d\omega \leq 1, \ \alpha \in [0,1] \tag{41}$$

with respect to the design variables K and Φ_r. The solution is to select the regulator $u = -Ky$ that solves the standard LQG problem

$$K^{opt} = \arg\min_K [\alpha E u^2 + (1-\alpha) E y^2], \quad y = G_0 u + H_0 e \tag{42}$$

The reference signal spectrum shall be chosen as

$$\Phi_r^{opt}(\omega) = \mu \sqrt{\Phi_v(\omega) C(\omega)} \frac{|1 + G_0(e^{i\omega}) K^{opt}(e^{i\omega})|^2}{\sqrt{\alpha + (1-\alpha)|G_0(e^{i\omega})|^2}} \tag{43}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \{\alpha \Phi_u + (1-\alpha)\Phi_y\} d\omega = 1 \tag{44}$$

This result can be proved as follows:

Proof. Replace the design variables K and Φ_r by the equivalent pair K and Φ_u^r. Then, by using expressions for the input and output spectra in terms of K and Φ_u^r we can rewrite the problem as

$$\begin{aligned} &\min_{K, \Phi_u^r} \int_{-\pi}^{\pi} \frac{\Phi_v}{\Phi_u^r} C d\omega \\ &\quad \text{under the contstraint} \\ &\quad \int_{-\pi}^{\pi} \{(\alpha + (1-\alpha)|G_0|^2)\Phi_u^r + \frac{\alpha |K|^2 + (1-\alpha)}{|1 + G_0 K|^2} \Phi_v\} d\omega \leq 1 \\ &\quad \alpha \in [0,1] \end{aligned} \tag{45}$$

The criterion function is independent of K hence the optimal controller K^{opt} can be found by solving the LQ problem

$$\min_K \int_{-\pi}^{\pi} \frac{\alpha |K|^2 + (1-\alpha)}{|1 + G_0 K|^2} \Phi_v d\omega \tag{46}$$

(Here it is implicitly assumed that $y(t) = G_0(q)u(t) + v(t)$, $u(t) = -K(q)y(t)$, and $\alpha \in [0,1]$.) This proves (42). Define the constant γ as

$$\gamma = 1 - \int_{-\pi}^{\pi} \frac{\alpha |K^{opt}|^2 + (1-\alpha)}{|1 + G_0 K^{opt}|^2} \Phi_v d\omega \tag{47}$$

Problem (45) now reads

$$\min_{\Phi_u^r} \{ \int_{-\pi}^{\pi} \frac{\Phi_v}{\Phi_u^r} C d\omega : \int_{-\pi}^{\pi} (\alpha + (1-\alpha)|G_0|^2)\Phi_u^r d\omega \leq \gamma \} \tag{48}$$

This problem has the solution (cf. [14], p. 376)

$$\Phi_u^r = \mu \sqrt{\frac{\Phi_v C}{(\alpha + (1-\alpha)|G_0|^2)}} \tag{49}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi}((\alpha + (1-\alpha)|G_0|^2)\Phi_u^r d\omega = \gamma \tag{50}$$

or in other words so that

$$\int_{-\pi}^{\pi}\{\alpha\Phi_u + (1-\alpha)\Phi_y\}d\omega = 1 \tag{51}$$

Consequently the optimal Φ_r is

$$\Phi_r^{opt} = \mu\sqrt{\Phi_v C}\frac{|1+G_0 K^{opt}|^2}{\sqrt{\alpha + (1-\alpha)|G_0|^2}} \tag{52}$$

which ends the proof. □

We stress that the optimal controller K^{opt} in (42) can easily be found by solving the indicated discrete-time LQ problem (if G_0 and Φ_v were known). Among other things this implies that the optimal controller K^{opt} is guaranteed to stabilize the closed-loop system and be linear, of the same order as G_0. This is a clear advantage over the results reported in, e.g., [10]. Furthermore, the optimal controller is independent of C which also is quite interesting and perhaps somewhat surprising. This means that whatever weighting C is used in the design criterion, it is always optimal to use the LQ regulator (42) in the identification experiment.

From the result we also see that closed-loop experiments are optimal as long as there is a constraint on the output power, i.e., as long as $\alpha \neq 1$. If $\alpha = 1$ then $K^{opt} = 0$ and the optimal input spectrum $\Phi_u^{opt}(= \Phi_r^{opt})$ becomes

$$\Phi_u^{opt} = \mu\sqrt{\Phi_v C} \tag{53}$$

If the constraint is on the output power only, Ey^2, the regulator K is the minimum variance controller. We then have the simple result, that any experiment that aims at minimizing the covariance of the dynamics, under output power constraint should be a minimum variance controller. To this we should add a reference signal with a power distribution that reflects the weighting function in the criterion. This result ties nicely with the special case treated in [7].

7 Conclusions

We have re-examined some basic results on bias and variance in closed loop identification. The common denominator in this analysis has been the data spectrum Φ_ζ. We have shown how different factorizations of this matrix give direct insights into identifiability, and bias distribution. It also gives a pragmatic "derivation" of the indirect method for closed loop identification.

This data spectrum also directly determines the variance of the estimated parameters and its inverse gives a simple and explicit expression for the asymptotic, black-box variance of the frequency functions. This in turn can be used to solve rather general experiment design problems, aiming at minimizing the covariance of the estimated dynamics under various realistic constraints.

References

[1] B.D.O. Anderson and M. Gevers. Identifiability of linear stochastic systems operating under linear feedback. *Automatica*, 18(2):195–213, 1982.

[2] K. J. Åström. Matching criteria for control and identification. In *Proceedings of the 2nd European Control Conference*, pages 248–251, Groningen, The Netherlands, 1993.

[3] R. de Callafon and P. Van den Hof. Multivariable closed-loop identification: From indirect identification to Dual-Youla parametrization. In *Proceedings of the 35th Conference on Decision and Control*, pages 1397–1402, Kobe, Japan, 1996.

[4] R. de Callafon, P. Van den Hof, and M. Steinbuch. Control relevant identification of a compact disc pick-up mechanism. In *Proceedings of the 32nd Conference on Decision and Control*, volume 3, pages 2050–2055, San Antonio,TX, 1993.

[5] B. Egardt. On the role of noise models for approximate closed loop identification. In *Proceedings of the European Control Conference*, Brussels, Belgium, 1997.

[6] M. Gevers. Towards a joint design of identification and control. In H. L. Trentelman and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkhäuser, 1993.

[7] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model application. *Automatica*, 22:543–554, 1986.

[8] I. Gustavsson, L. Ljung, and T. Söderström. Identification of processes in closed loop — Identifiability and accuracy aspects. *Automatica*, 13:59–75, 1977.

[9] F. R. Hansen. *A fractional representation to closed-loop system identification and experiment design.* Phd thesis, Stanford University, Stanford, CA, USA, 1989.

[10] H. Hjalmarsson, M. Gevers, and F. De Bruyne. For model-based control design,closed loop identification gives better performance. *Automatica*, 32, 1996.

[11] I. D. Landau and K. Boumaiza. An output error recursive algorithm for identification in closed loop. In *Proceedings of the 13th IFAC World Congress*, volume I, pages 215–220, San Francisco, CA, 1996.

[12] I. D. Landau and A. Karimi. Recursive algorithms for identification in closed loop: A unified approach and evaluation. *Automatica*, 33(8):1499–1523, 1997.

[13] 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.

[14] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[15] L. Ljung and I. D. Landau. Model-reference adaptive systems and self-tuning regulators: Some connections. In *Proc 7th IFAC World Congress*, pages 1899–1906, Helsinki, Finland, 1978. Paper no 46 A.2.

[16] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall International, 1989.

[17] P. M. J. Van den Hof and R. J. P. Schrama. Identification and control — Closed-loop issues. *Automatica*, 31(12):1751–1770, 1995.

[18] Z. Zang, R. R. Bitmead, and M. Gevers. Iterative weighted least-squares identification and weighted LQG control design. *Automatica*, 31(11):1577–1594, 1995.