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

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

Technical reports from the Automatic Control group in Linköping are available by anonymous ftp at the address ftp.control.isy.liu.se.
This report is contained in the compressed postscript file 2100.ps.Z.
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],[8],[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

\[ y(t) = G_0(q)u(t) + H_0(q)e(t) \]  

(1)

where \( q \) is the delay operator, \( u \) is the input, \( y \) is the output and \( e \) is white noise with covariance matrix \( \Lambda_0 \).

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

\[ \zeta = \begin{bmatrix} u \\ e \end{bmatrix} \]  

(2)

is well defined, and denoted by

\[ \Phi_\zeta(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \Lambda_0 \end{bmatrix} \]  

(3)

The system is identified within the model structure

\[ y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \]  

(4)

\( G \) will be called the dynamics model and \( H \) the noise model. The parameter \( \theta \) is estimated by

\[ \hat{\theta}_N = \arg \min_{\theta \in \mathcal{D}_N} V_N(\theta, Z^N) \]  

(5)

\[ V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t, \theta) \Lambda^{-1}\varepsilon(t, \theta) \]  

(6)

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

(7)

Here \( \Lambda \) 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 \( \Phi_\zeta \) plays an important role in the analysis, and we shall therefore collect some results on it here. We shall introduce the following spectrum

\[ \Phi^r_u = \Phi_u - \Phi_{ue}\Lambda_0^{-1}\Phi_{eu} \]  

(8)

where we suppress the argument \( \omega \). 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

\[ \Phi^r_e = \Lambda_0 - \Phi_{eu}\Phi_u^{-1}\Phi_{ue} \]  

(9)
The data spectrum $\Phi_\zeta$ can now be written as
\[
\begin{bmatrix}
\Phi_u & \Phi_{ue} \\
\Phi_{eu} & \Lambda_0
\end{bmatrix} = 
\begin{bmatrix}
I & \Phi_{ue}\Lambda_0^{-1} \\
0 & I
\end{bmatrix} 
\begin{bmatrix}
\Phi_u^r & 0 \\
0 & \Lambda_0^{-1}\Phi_{eu}
\end{bmatrix} 
\begin{bmatrix}
I & 0 \\
0 & \Lambda_0^{-1}
\end{bmatrix} = 
\begin{bmatrix}
I & \Phi_{ue}^{-1} \\
\Phi_{eu} & I
\end{bmatrix} 
\begin{bmatrix}
I & 0 \\
0 & \Phi_{eu}^{-1}
\end{bmatrix}
\] (10)

From these factorization results we also find an expression for the inverse
\[
\Phi_\zeta^{-1} = 
\begin{bmatrix}
\Phi_u & \Phi_{ue} \\
\Phi_{eu} & \Lambda_0
\end{bmatrix}^{-1} = 
\begin{bmatrix}
(\Phi_u^r)^{-1} & -\Phi_{ue}^{-1}\Phi_{ue}\Lambda_0^{-1} \\
-\Lambda_0^{-1}\Phi_{eu}(\Phi_u^r)^{-1} & (\Phi_u^r)^{-1}
\end{bmatrix}
\] (13)

In case the regulator is linear, time-invariant, we have
\[
u(t) = r(t) - K(q)y(t)
\] (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 $S^i$ be the output and input sensitivity functions:
\[
S_0 = (I + G_0K)^{-1}, \quad S_0^i = (I + KG_0)^{-1}
\] (15)

Then
\[
\Phi_u = S_0^i \Phi_r(S_0^i)^*, \quad K S_0 \Phi_r S_0^* K^*
\] (16)

where $\Phi_r$ is the spectrum of the reference signal and $\Phi_r = H_0\Lambda_0 H_0^*$ the noise spectrum. Superscript * denotes complex conjugate transpose. We shall denote the two terms in (16)
\[
\Phi_u^r = S_0^i \Phi_r(S_0^i)^*
\] (17)

and
\[
\Phi_u^e = K S_0 \Phi_r S_0^* K^* = S_0^i K \Phi_r K^*(S_0^i)^*
\] (18)

The cross spectrum between $u$ and $e$ is
\[
\Phi_{ue} = -K S_0 H_0 \Lambda_0 = -S_0^i K H_0 \Lambda_0
\] (19)

3 The Main Expression

From standard asymptotic theory we know that
\[
\hat{\theta}_N \to \arg\min_{\theta} \nabla(\theta) \quad \text{w.p.1 as} \quad N \to \infty
\] (20)
where
\[
\nabla(\theta) = \mathbb{E} V_N(\theta, Z^N)
\]
\[
= \int_{-\pi}^{\pi} \text{tr} \left\{ \Lambda^{-1} H_\theta^{-1} [\Delta G_\theta \quad \Delta H_\theta] \Phi_\zeta \left[ \frac{\Delta G_\theta^*}{\Delta H_\theta^*} (H_\theta^{-1})^* \right] \right\} d\omega \quad (21)
\]

Here we have introduced the simplified notation
\[
\Delta G_\theta = G_0(e^{i\omega}) - G(e^{i\omega}, \theta) \\
\Delta H_\theta = H_0(e^{i\omega}) - H(e^{i\omega}, \theta)
\]

To see (21) we rewrite (7) using (1) as
\[
\epsilon = H_\theta^{-1}(y - G_\theta u) = H_\theta^{-1}(G_0 - G_\theta)u + H_\theta^{-1}H_0\epsilon \\
= H_\theta^{-1}(G_0 - G_\theta)u + (H_\theta^{-1}H_0 - I)\epsilon + \epsilon \\
= H_\theta^{-1}[(G_0 - G_\theta)u + (H_0 - H_\theta)\epsilon] + \epsilon
\]

We then use that \( H_0 \) and \( H_\theta \) 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 \( \Phi_\zeta \).

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

\[\text{consistency and identifiability follows when no model in the structure } [\Delta G_\theta \quad \Delta H_\theta] \text{ lies in the null space of } \Phi_\zeta.\]

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 \( \Lambda_0 \) to be non-singular)
\[
\Phi_\zeta^u(\omega) > 0, a.e.\omega \quad (22)
\]

In [14] this is termed that the experiment is \textit{informative enough}. With the definition (8) this means that there should be a full range 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 \( H_\theta \) is fixed to the true value \( H_0 = 0 \) \((\Delta H_\theta = 0)\), and that the dynamics model structure \( G_\theta \) 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 \( G_\theta \) 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

\[
\hat{\theta}_N \to \arg \min_{\theta \in \mathcal{D}_M} \int_{-\pi}^{\pi} \text{tr} \left[ \Lambda^{-1} \left\{ (G_0 + B_\theta - G_\theta) \Phi_u (G_0 + B_\theta - G_\theta) \right\}^* \right. \\
+ \left. (H_0 - H_\theta) \Phi_e (H_0 - H_\theta)^* H_\theta^{-1} \right] d\omega \quad \text{w. p. 1 as } N \to \infty \tag{23}
\]

where

\[
B_\theta = (H_0 - H_\theta) \Phi_e \Phi_u^{-1} \tag{24}
\]

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

\[
|B_\theta|^2 = \frac{\Lambda_0}{\Phi_u} \frac{\Phi_e}{\Phi_u} |H_0 - H_\theta|^2 \tag{25}
\]

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

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 \( G_\theta \) contains the true dynamics \( G_0 \), and that we are not interested in the noise characteristics. Is it then possible to obtain a consistent estimate of \( G \)?

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

\[ \nabla(\theta) = \int_{-\pi}^{\pi} \text{tr} \left[ \Lambda^{-1} H^{-1}_\theta (\Delta G_\theta \Phi^*_u (\Delta G_\theta)^* H^{-*}_\theta \right] d\omega \\
+ \int_{-\pi}^{\pi} \text{tr} \left[ \Lambda^{-1} H^{-1}_\theta (\Delta G_\theta \Phi_{ue} \Lambda^{-1}_0 + \Delta H_\theta) \Lambda_0 \right] \\
\times (\Delta G_\theta \Phi_{ue} \Lambda^{-1}_0 + \Delta H_\theta)^* H^{-*}_\theta] d\omega \]

(26)

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

\[ H^{-1}_\theta(G_\theta K S_0 H_0 - G_\theta K S_0 H_0 + H_0) - I \\
= H^{-1}_\theta(G_\theta K - G_\theta K + I + G_\theta K) S_0 H_0 - I \\
= H^{-1}_\theta(I + G_\theta K) S_0 H_0 - I = \tilde{H}_\eta^{-1} S_0 H_0 - I \\
\]

where we introduced the noise model parameterization

\[ H_\theta = (I + G_\theta K) \tilde{H}_\eta = S^{-1}_\theta \tilde{H}_\eta \]

(27)

where \( \eta \) is a parameterization, independent of \( \theta \). Here \( S_\theta \) 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 \( G_\theta \). The first integral will thus determine to what the dynamics model converges: If \( \tilde{H} \) is a fixed model we have

\[ \theta_N \rightarrow \arg \min_{\theta \in \Omega_N} \int_{-\pi}^{\pi} \text{tr} \left[ \Lambda^{-1} \left\{ \tilde{H}^{-1}_\theta(S_\theta \Delta G_\theta) \Phi^*_u (S_\theta \Delta G_\theta)^{*} H^{-*}_\theta \right\} \right] d\omega \]

(28)

With (17) we find that

\[ (S_\theta \Delta G_\theta) \Phi^*_u (S_\theta \Delta G_\theta)^* = (S_\theta \Delta G_\theta S_0^{i}) \Phi_r (S_\theta \Delta G_\theta S_0^{i})^* \]

(29)

with

\[ S_0 G_0 - S_0 G_\theta = G_0 S_0^{i} - S_\theta G_\theta \]

\[ = S_\theta ((I + G_\theta K)G_0 - G_\theta (I + K G_0)) S_0^{i} = S_\theta \Delta G_\theta S_0^{i} \]

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 \( \Phi_r \) 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

\[ y(t) = G(q, \theta) u(t) + H(q, \theta) e(t) \]

(30)
is

\[
\dot{y}(t|\theta) = H^{-1}(q, \theta)G(q, \theta)u(t) + (I - H^{-1}(q, \theta))y(t)
\]

(31)

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

\[
\dot{y}(t|\theta) = \tilde{H}^{-1}(q, \theta)(I + G(q, \theta)K(q))^{-1}G(q, \theta)r(t)
+ (I - \tilde{H}^{-1}(q, \theta))y(t)
\]

(32)

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

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

(33)

Identifying the closed loop, and then solving for the open loop dynamics is called the \textit{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 \( \theta_0 \). 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

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

(34)

where \( V \) is defined by (21) and prime denotes differentiation w.r.t \( \theta \).

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

\[
T_\theta = \begin{bmatrix} G_\theta & H_\theta \end{bmatrix}, \quad T_N = \begin{bmatrix} \hat{G}_N & \hat{H}_N \end{bmatrix}
\]

(35)

Then, using (21) we can write

\[
P_\theta = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} T^T \Phi_v(T') \, d\omega \right]^{-1}
\]

(36)

Here \( \Phi_v = |H_0|^2 \Lambda_0 \). From this expression and the factorizations (10) and (11) explicit expression can be found how \( \Phi_r \) 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

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

(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):

$$\text{Cov} \left[ \hat{G}_N(e^{i\omega}) \right] \approx \frac{n}{N} \Phi_v(\omega) \left[ \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}(\omega) & \Lambda_0 \end{bmatrix} \right]^{-1} = \frac{n}{N} \Phi_v(\omega)$$

(38)

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

$$\text{Cov} \left[ \hat{G}_N(e^{i\omega}) \right] \approx \frac{n}{N} \Phi_v(\omega)$$

(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 $G$:

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

(40)

The minimization shall be carried out with respect to the experiment design variables, which we take as $K$ (the regulator) and $\Phi_r$ (the reference signal spectrum). Other equivalent choices are also possible, e.g., $\Phi_u$ and $\Phi_{ue}$ or $K$ and $\Phi_r$. 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

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

(41)
with respect to the design variables $K$ and $r$. The solution is to select the regulator $u = -Ky$ that solves the standard LQG problem

$$K^{opt} = \arg \min_K \left[ \alpha Eu^2 + (1 - \alpha)Ey^2 \right], \quad y = G_0u + H_0e \quad (42)$$

The reference signal spectrum shall be chosen as

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

where $\mu$ is a constant, adjusted so that

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

This result can be proved as follows:

**Proof.** Replace the design variables $K$ and $\Phi_r$ by the equivalent pair $K$ and $\Phi_u$. Then, by using expressions for the input and output spectra in terms of $K$ and $\Phi_u$ we can rewrite the problem as

$$\min_{K, \Phi_u} \int_{-\pi}^{\pi} \frac{\Phi_v}{\Phi_u} C d\omega$$

under the constraint

$$\int_{-\pi}^{\pi} \{ (\alpha + (1 - \alpha)|G_0|^2)\Phi_u + \frac{\alpha|K|^2 + (1 - \alpha)}{1 + G_0K^2} \Phi_v \} d\omega \leq 1$$

$$\alpha \in [0, 1] \quad (45)$$

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

$$\min_K \int_{-\pi}^{\pi} \frac{\alpha|K|^2 + (1 - \alpha)}{1 + G_0K^2} \Phi_v d\omega \quad (46)$$

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

$$\gamma = 1 - \int_{-\pi}^{\pi} \frac{\alpha|K^{opt}|^2 + (1 - \alpha)}{1 + G_0K^{opt}^2} \Phi_v d\omega \quad (47)$$

Problem (45) now reads

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

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

$$\Phi_u = \mu \sqrt{\frac{\Phi_v C}{\alpha + (1 - \alpha)|G_0|^2}} \quad (49)$$
where $\mu$ is a constant, adjusted so that

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

(50)

or in other words so that

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

(51)

Consequently the optimal $\Phi_r$ is

$$\Phi_r^{opt} = \frac{\mu \sqrt{\Phi_v C} \left[1 + G_0 K^{opt} \right]^2}{\sqrt{\alpha + (1 - \alpha)|G_0|^2}}$$

(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 $\Phi_v$ 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

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

(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 $\Phi_x$. 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


