On the Frequency Domain Accuracy of Closed Loop Estimates

Brett Ninness*
School of of Electrical Engineering &
Computer Science, University of Newcastle,
Australia.
brett@ee.newcastle.edu.au

Håkan Hjalmarsson**
Department of Sensors, Signals and Systems,
The Royal Institute of Technology, S-100 44
Stockholm, Sweden.
hakan.hjalmarsson@s3.e.kth.se

***Abstract*— It has been argued that the frequency domain accuracy of high model-order estimates obtained on the basis of closed loop data is largely invariant to whether direct or indirect approaches are used. This paper revisits this study in light of new variance quantification results that apply for low model order and establishes that, under certain assumptions, there can be significant differences in the accuracy of frequency response estimates that are dependent on what type of direct, indirect or joint input-output identification strategy is pursued.**

I. Introduction

Over the last several years there has been significant interest in the interplay between system identification and subsequent control design [2], [4], [9]. In particular, a variety of estimation techniques adapted to these closed and open loop scenarios have been developed. They can be broadly divided into 'direct', 'indirect' and 'joint input-output' estimation methods, and of significant relevance to this paper has been recent work [1], [3], [9] examining the relative accuracy of estimates obtained via the different approaches.

In particular [1], [3] uses certain approximations derived in [6] for the mean square error in frequency response estimates, and whose accuracy depends on data length and model order being large. Via these [1], [3] then argues that, in the limit as the model order tends to infinity, direct, indirect and joint input-output methods offer the same accuracy in the frequency domain. However, in [3] the authors point out that an assumption of large model order '*apparently diminishes possible differences*'. Furthermore, in [9] it is noted that '*this asymptotic variance analysis tool is also quite crude*', and '*for finite model orders, the variance results will likely become different over the several methods*'. Indeed this expected difference is established in [1] for parameter estimates, but not for the associated frequency response estimates.

These prior works and observations by other authors suggest the need for further study that quantifies the frequency domain accuracy of direct, indirect and joint input-output estimates that are valid for finite, and possibly low model order. Such is the purpose of this and the companion paper [7], which addresses the problem by using new expressions for the mean square estimation error that have been recently developed in [8], and whose accuracy is not dependent on the model order being large; they are exact for finite model order.

As will be seen, these exact expressions can, depending on the estimation conditions, be very different from those arising from [6] and employed in [1], [3], and this will imply new conclusions on the relationship between direct, indirect and joint input-output estimation approaches.

*This work was supported by the Australian Research Council.

**Work by this author done while visiting University of Newcastle.

For example, under certain assumptions on reference spectrum and control design, we establish that for finite model order, the variance of the estimated frequency response arising from a direct method can be very different from that associated with indirect and joint input-output methods.

Furthermore, we establish that, again in the frequency domain and for finite model order, the variance of these flavours of indirect and joint-input output estimates are none of them equivalent to one another. All proofs of the results presented here are provided in [7] which, relative to this paper, provides an expanded discussion and extended results.

II. Preliminaries

This paper addresses the following closed loop control scenario

$$\mathcal{S}: \quad y_t \quad = \quad G(q)u_t + e_t, \; u_t = K(q)\left[r_t - y_t\right] \quad (1)$$

for some underlying true system $\mathcal{S}$ characterised by the transfer function $G(q)$ and which is under the influence of a linear time invariant controller $K(q)$ and an external set point signal $\{r_t\}$. Here $G(q)$ and $K(q)$ are both rational in the backward shift operator q^{-1} and $\{e_t\}$ is a zero-mean white noise sequence that satisfies $\mathbf{E}\{e_t^2\} = \sigma^2, \mathbf{E}\{|e_t|^8\} < \infty$.

In this case, the relationship between the signals in (1) may also be expressed as

$$y_t \quad = \quad Tr_t + e_t, \; u_t = SKr_t - SKe_t \quad (2)$$

where $S(q)$ and $T(q)$ are the sensitivity and complementary sensitivity functions given (respectively) as

$$S(q) = \frac{1}{1 + G(q)K(q)}, \qquad T(q) = \frac{G(q)K(q)}{1 + G(q)K(q)}. \quad (3)$$

It will be assumed that the reference signal $\{r_t\}$ is a quasi-stationary process with associated spectral density $\Phi_r(\omega)$, in which case provided the poles of $S(q)K(q)$ are in the open unit disk, then other spectral densities will exist. In particular the spectral density $\Phi_u^r(\omega)$ of the component of $\{u_t\}$ that derives solely from $\{r_t\}$ will be given as

$$\Phi_u^r(\omega) = |K(e^{j\omega})S(e^{j\omega})|^2 \Phi_r(\omega). \quad (4)$$

A prime focus of this paper will be the provision of quantifications that are exact with respect to a finite model order of interest. For this purpose, it will be necessary to restrict the class of possible controllers. Namely, if $G(q)$ and $K(q)$ are of rational form

$$G(q) = \frac{B(q)}{A(q)}, \qquad K(q) = K\frac{P(q)}{L(q)} \quad (5)$$

where $K \in \mathbf{R}$ and $A(q), B(q), L(q), P(q)$ are all polynomials in q^{-1}, then it will be required that the numerator $P(q)$

is formed as a subset of the open loop poles $A(q)$. That is, for some polynomial $\widetilde{A}(q)$ in q^{-1}

$$\widetilde{A}(q)P(q) = A(q). \tag{6}$$

This, and further specialised requirements that are necessary to impose in the interests of dealing with finite model order are collected in the following set of standing assumptions.

Standing Assumptions 2.1: The following standing assumptions will be repeatedly imposed:

1) $\Phi_r(\omega) = \mu$ a constant;
2) The controller $K(q)$ is of the form

$$K(q) = K\,\frac{P(q)}{L(q)}, \quad L(q) = \prod_{k=1}^{m_\ell}(1-\ell_k q^{-1}), \quad K \in \mathbf{R}, \tag{7}$$

and where $P(q)$ satisfies (6);
3) The closed loop poles are the zeros of

$$A_c(z) = \widetilde{A}(z)L(z) + K\,B(z);$$

4) There is no undermodelling and there are no pole-zero cancellations in any of the asymptotictransfer function estimates.

III. Direct Identification

With the preceding assumptions in mind, this paper addresses the issue of estimating the dynamics $G(q)$ on the basis of observing N observations of input-output response via the use of a model structure

$$\mathcal{M}: \quad y_t = G(q,\theta)u_t + e_t = \frac{B(q,\theta)}{A(q,\theta)}u_t + e_t \tag{8}$$

where

$$A(q,\theta) = 1 + a_1 q^{-1} + a_2 q^{-1} + \cdots + a_{m_a} q^{-m_a}, \tag{9}$$
$$B(q,\theta) = b_1 q^{-1} + b_2 q^{-1} + \cdots + b_{m_b} q^{-m_b}, \tag{10}$$

for some integers m_a, m_b, m_c, m_d and where $\theta \in \mathbf{R}$ is a vector parameterising the above polynomials. Under a scheme of so-called ‘direct’ identification, one works directly with the signals $\{y_t\}$ and $\{u_t\}$ so that the dynamics in (8) are estimated by forming $\widehat{\theta}_N$ according to

$$\widehat{\theta}_N = \arg\min_{\theta\in\mathbf{R}} V_N(\theta), \qquad V_N(\theta) \triangleq \frac{1}{2N}\sum_{t=1}^{N}\varepsilon_t^2(\theta) \tag{11}$$

where

$$\varepsilon_t(\theta) = y_t - \widehat{y}_t(\theta)$$

is the error between the observed y_t and the mean square optimal one step ahead prediction $\widehat{y}_t(\theta)$ of y_t, according to the model (8) given as

$$\widehat{y}_t(\theta) = H^{-1}(q,\theta)G(q,\theta)u_t + \left[1 - H^{-1}(q,\theta)\right]y_t.$$

IV. Main Results

A. Direct Identification

The first main result is for the case of direct estimation as just described. In what follows, a direct estimate formed from the model structure (8) is denoted as $G_{\rm di}(q,\widehat{\theta}_N)$.

Theorem 4.1: Suppose that the prediction error estimation method outlined previously is employed and that the standing assumptions 2.1 are satisfied. Suppose further that the polynomial $A(z)A_c(z)$ is of order less than or equal to the sum of the numerator and denominator model orders $m_a + m_b$. Then

$$\lim_{N\to\infty} N\,\mathit{Var}\left\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\right\} = \frac{\sigma^2}{\Phi_u(\omega)}\cdot\kappa_{\rm di}(\omega) \tag{12}$$

where $\kappa_{\rm di}(\omega)$ is given by the general form

$$\kappa(\omega) = \sum_{k=1}^{m}\frac{1-|\xi_k|^2}{|e^{j\omega}-\xi_k|^2} \tag{13}$$

with the associated $\{\xi_1,\cdots,\xi_{m_a+m_b}\}$, $m = m_a + m_b$ being defined as the zeros of the polynomial $z^{m_a+m_b}A(z)A_c(z)$. ∎

This result suggests the following quantification whose accuracy depends on data length N being large, but which is exact for the indicated finite model order

$$\mathrm{Var}\left\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2}{N\Phi_u(\omega)}\,\kappa_{\rm di}(\omega). \tag{14}$$

B. Indirect Identification

A unifying feature of indirect identification methods arises via the common consideration of the relationship (2) to suggest the use of the model structure

$$y_t = T(q,\theta)r_t + S(q,\theta)e_t. \tag{15}$$

The rationale for this choice is that since $\{e_t\}$ is typically uncorrelated with $\{r_t\}$, then the estimation of $T(q)$ on the basis of the observations $\{y_t\}$, $\{r_t\}$ can be achieved with a fixed noise model/prefilter while still resulting in a consistent estimate. Consequently, and in accord with how $\widehat{\theta}_N$ is formed according to (11), then with an indirect approach $\widehat{\theta}_N$ is also formed via (11) but with associated prediction error

$$\varepsilon_t(\theta) = SH^{-1}(q,\theta)\left[y_t - T(q,\theta)r_t\right]. \tag{16}$$

The estimate $T(q,\widehat{\theta}_N)$ can then be used to derive an ‘indirect’ estimate $G(q,\widehat{\theta}_N)$ according to (3) as

$$G(q,\widehat{\theta}_N) = \frac{T(q,\widehat{\theta}_N)}{K(q)[1-T(q,\widehat{\theta}_N)]}. \tag{17}$$

Therefore, aside from the advantages attendant to being an open loop scenario, the indirect approach suffers from the need to know the controller $K(q)$.

With these essential features in mind, the various indirect estimation approaches then differ in their choice of parameterisation for $T(q,\theta)$.

1) Basic Indirect Identification: What will here be called 'basic' indirect identification involves the following obvious choice of model structure

$$T(q,\theta)=\frac{\beta_0+\beta_1 q^{-1}+\cdots+\beta_{m_\beta}q^{-m_\beta}}{1+\alpha_1 q^{-1}+\cdots+\alpha_{m_\alpha}q^{-m_\alpha}} \qquad (18)$$

where $S(q,\theta)$ is parametrized independently of $T(q,\theta)$ and in such a manner to encompass the true underlying $S(q)$.

It is natural to also question the likely accuracy of $G(e^{j\omega},\widehat{\theta}_N)$ found via (17) in relation to that of a direct approach. This question has been addressed in [1], [3] which concluded that the accuracy of $G(e^{j\omega},\widehat{\theta}_N)$ found via this *indirect* approach can be quantified as [5]

$$\mathrm{Var}\{G(e^{j\omega},\widehat{\theta}_N)\} \;\approx\; \frac{m}{N}\frac{\sigma^2}{\Phi_u^r(\omega)}. \qquad (19)$$

The first main result of this paper in the context of indirect identification methods employs the results of [8] to examine a different but related question of frequency domain accuracy for finite, and possibly small model order.

Theorem 4.2: Suppose that the Standing Assumptions 2.1 are satisfied. Suppose further that the polynomial $A_c(z)\widetilde{A}(z)L(z)$ is of order less than or equal to the sum of the numerator and denominator model orders $m_\alpha+m_\beta$. Then in the case of the indirect identification procedure (15)- (17)

$$\lim_{N\to\infty} N\,\mathit{Var}\left\{G_{\mathrm{id}}(e^{j\omega},\widehat{\theta}_N)\right\}=\frac{\sigma^2}{\Phi_u(\omega)}\cdot\kappa_{\mathrm{id}}(\omega) \qquad (20)$$

where $\kappa_{\mathrm{id}}(\omega)$ is given by (13) with the associated zeros $\{\xi_1,\cdots,\xi_{m_a+m_b}\}$ being defined as those of the polynomial $z^{m_\alpha+m_\beta}A_c(z)\widetilde{A}(z)L(z)$.

■

As per the case of the previous Theorem 4.1, this result suggests a variance quantification whose accuracy depends only on N being sufficiently large, and in that sense is exact for finite model order:

$$\mathrm{Var}\left\{G_{\mathrm{id}}(e^{j\omega},\widehat{\theta}_N)\right\}\approx\frac{\sigma^2}{N\Phi_u^r(\omega)}\kappa_{\mathrm{id}}(\omega). \qquad (21)$$

There are two important conclusions to be drawn from this quantification and the associated one (14) pertaining to direct identification.

First, and most important, they indicate that from an estimation accuracy point of view, direct and indirect identification are *not* equivalent. Both quantifications (21) and (14) depend on the closed loop poles defined by $A_c(z)$.

However, according to (21) the indirect method estimation error $\mathrm{Var}\{G_{\mathrm{id}}(e^{j\omega},\widehat{\theta}_N)\}$ also depends on *a subset* of the open loop poles defined by $\widetilde{A}(z)$ and the controller poles defined by $L(z)$. On the other hand (14) indicates that, modulo effects on $\Phi_u(\omega)$, the direct method estimation error $\mathrm{Var}\{G_{\mathrm{di}}(e^{j\omega},\widehat{\theta}_N)\}$ does not depend on these poles, but does depend on *all* the open loop poles of the system $G(q)$.

Of course, it is entirely possible for the controller poles defined and open loop system poles to be very different and hence for the estimation errors to be very different. This will shortly be demonstrated via simulation example.

However, one situation in which the controller poles and open loop poles are the same is if the controller (7) is in fact simply a proportional controller $K(q)=K$ for some constant K, and in this case when $\mu\gg\sigma^2$ so that $\Phi_u^r(\omega)\approx\Phi_u(\omega)$ then $\mathrm{Var}\{G_{\mathrm{id}}(e^{j\omega},\widehat{\theta}_N)\}\approx\mathrm{Var}\{G_{\mathrm{di}}(e^{j\omega},\widehat{\theta}_N)\}$.

Second, in comparing direct and indirect methods by comparing $\kappa_{\mathrm{di}}(\omega)$ and $\kappa_{\mathrm{id}}(\omega)$, note that both these terms will be of the generic form

$$\kappa(\omega)=m_\star+\sum_{k=1}^{2m-m_\star}\frac{1-|\xi_k|^2}{|e^{j\omega}-\xi_k|^2} \qquad (22)$$

where m is the McMillan degree of the true underlying system, and $m_\star$ is the number of numerator lags in the model structure that are excess of this degree, while the $\{\xi_k\}$ are the zeros of $A_c(z)A(z)$ or $A_c(z)\widetilde{A}(z)L(z)$ as appropriate.

For cases of large model order with many lags in the numerator, the constant $m_\star$ will dominate (22) and hence $\kappa_{\mathrm{di}}(\omega)$ and $\kappa_{\mathrm{id}}(\omega)$ will be essentially equivalent.

On the other hand, while (22) illustrates that the expressions (21) and (14) both imply (in the high SNR case where $\Phi_u\approx\Phi_u^r$) the asymptotic result

$$\lim_{m_\star\to\infty}\lim_{N\to\infty}\frac{N}{m_\star}\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\}=\frac{\sigma^2}{\Phi_u(\omega)} \qquad (23)$$

it should be clear from the afore-mentioned potential differences between (14) and (21) that there can be serious pitfalls in concluding the approximate quantification

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\}\approx\frac{m_\star}{N}\frac{\sigma^2}{\Phi_u(\omega)} \qquad (24)$$

on the basis of the asymptotic result (23).

2) Identification via Dual-Youla Parameterisation: The second indirect estimation method to be considered here is one based on the so-called dual-Youla parameterisation of all systems stabilised by a given controller. Namely, suppose that the controller $K(q)$, and a nominal, user chosen system $G^0(q)$ which is stabilised by $K(q)$, have co-prime representation with respect to stable $M(q),N(q),X(q),Y(q)$ as

$$G^0(q)=N(q)M^{-1}(q), \qquad K(q)=X(q)Y^{-1}(q). \quad (25)$$

Then a further system $G(q,\theta)$, parametrized by $\theta\in\mathbf{R}^n$, is also stabilised by $K(q)$ if and only if there exists a stable and proper $R(q,\theta)$ such that

$$G(q,\theta)=\frac{N(q)+Y(q)R(q,\theta)}{M(q)-X(q)R(q,\theta)}. \qquad (26)$$

Substituting this parameterisation into (3) then implies a parameterisation of $T(q,\theta)$ in terms of $R(q,\theta)$ according to

$$T(q,\theta) \;=\; T^0(q)+M^{-1}(q)X(q)S^0(q)R(q,\theta) \quad (27)$$

where $T^0(q)$ and $S^0(q)$ arise from (3) with the substitution $G(q)=G^0(q)$. Further substitution of (27) into (15) then leads to

$$z_t=R(q,\theta)x_t+W(q,\theta)e_t \qquad (28)$$

where

$$z_t\triangleq y_t-T^0r_t,\; x_t=M^{-1}XS^0r_t,\; W(q,\theta)=S(q,\theta).$$

Here, if $\{r_t\}$ and $\{e_t\}$ are uncorrelated, then so are $\{x_t\}$ and $\{e_t\}$ and hence finding an estimate $\widehat{\theta}_N$ via (11) and the prediction error residual suggested by (28) of

$$\varepsilon_t(\theta) = SH^{-1}(q,\theta)\,[z_t - R(q,\theta)x_t]$$

is an open loop estimation problem [1], [9]. Furthermore, it is essentially a re-parameterisation of the indirect estimation model structure (15) in which an estimate $G_{\rm dy}(q,\widehat{\theta})$ may be derived from $R(q,\widehat{\theta}_N)$ via (26).

In order to examine the accuracy of $G_{\rm dy}(q,\widehat{\theta}_N)$, it will be expedient to restrict attention to the following class of co-prime factorisations

$$N = \frac{B^0}{E_G},\ M = \frac{A^0}{E_G},\ X = K\frac{P}{E_K},\ Y = \frac{L}{E_K} \qquad (29)$$

where in $X(q), K \in \mathbf{R}$ and all the terms on the right hand side of equals signs in (29) are polynomial in q^{-1}, $E_G(q)$ and $E_K(q)$ are user-chosen Schur polynomials, and $G^0(q) = B^0(q)/A^0(q)$, $K(q) = P(q)/L(q)$. Furthermore, it will be assumed that the model structure $R(q,\theta)$ is of the form

$$R(q,\theta) = \frac{\lambda^0 + \lambda_1 q^{-1} + \cdots + \lambda_{m_\lambda} q^{-m_\lambda}}{1 + \gamma_1 q^{-1} + \cdots + \gamma_{m_\gamma} q^{-m_\gamma}} \qquad (30)$$

and that $W(q,\theta)$, while being parametrized independently of $R(q,\theta)$, is also of sufficient order to encompass $Y^{-1}(q)S(q)$. With this in mind, the following theorem provides the main result of this subsection.

Theorem 4.3: Suppose that the Standing Assumptions 2.1 are satisfied. Suppose further that the polynomial $E_G(z)E_K(z)P(z)A_c(z)A_c^0(z)$ is of order less than or equal to the sum of the numerator and denominator model orders $m_\gamma + m_\lambda$, where $A_c(z)$ was defined in (8) and

$$A_c^0(z) \triangleq A^0(z)L(z) + K\,P(z)B^0(z).$$

Then in the case of indirect identification using the dual-Youla parameterisation (25)-(30)

$$\lim_{N\to\infty} N\,\mathit{Var}\left\{G_{\rm dy}(e^{j\omega},\widehat{\theta}_N)\right\} = \frac{\sigma^2}{\Phi_u^r(\omega)}\,\kappa_{\rm dy}(\omega)$$

where $\kappa_{\rm dy}(\omega)$ is as defined in (13) with the associated zeros $\{\xi_1,\cdots,\xi_{m_\lambda+m_\gamma}\}$ being those of

$$z^{m_\gamma+m_\lambda}E_G(z)E_K(z)A(z)A_c(z)A_c^0(z). \qquad (31)$$

■

As in previous cases, this new result invites comment. Firstly, and most obviously, the associated quantification, exact for finite model order, of

$$\mathrm{Var}\left\{G_{\rm dy}(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2}{N\Phi_u^r(\omega)}\kappa_{\rm dy}(\omega) \qquad (32)$$

and compared to (21), (14) and (37), indicates that in general, the accuracy of estimates $G_{\rm dy}(q,\widehat{\theta}_N)$ obtained by the dual Youla method are *not* equivalent to those obtained by direct, or basic indirect methods. This is due to the dependence of (32) on the zeros of the nominal closed loop denominator $A_c^0(q)$, and on the zeros of $E_G(z)$ and $E_K(z)$ used in the definition of the co-prime factorisations of $G^0(z)$ and $K(z)$. Since $A_c^0(q), E_G(z)$ and $E_K(z)$ can be rather arbitrary (modulo the requirement they be Schur), then $\mathrm{Var}\{G_{\rm dy}(e^{j\omega},\widehat{\theta}_N)\}$ can also be very different from the variance of other indirect, and also direct system estimates.

Indeed, for $\mu \gg \sigma^2$ so that $\Phi_u^r(\omega) \approx \Phi_u(\omega)$ then comparison with (14) indicates that with the choices $E_K(z) = E_G(z) = 1$ and with $\{\rho_k\}$ being the zeros of $A_c^0(z)$, and $K(q)$ being a pole cancelling design with $P(q) = KA(q)$, then

$$\mathrm{Var}\{G_{\rm dy}\} \ \approx\ \mathrm{Var}\{G_{\rm di}\} + \frac{\sigma^2}{N\Phi_u^r(\omega)}\sum_{k=1}^{m_\rho}\frac{1-|\rho_k|^2}{|e^{j\omega}-\rho_k|^2} \quad (33)$$

and hence, under the conditions of the preceding theorems, it could be expected that regardless of the choice of $G^0(q)$, the accuracy of the dual Youla method will be inferior to that of a direct method.

Secondly, in the preceding cases it was illustrated that in the special case of proportional control $K(q) = K \in \mathbf{R}$, the variance of the basic indirect method became equal to that of the direct method. Such is not the case for the dual-Youla approach. Regardless of the controller being proportional, there is a further degree of freedom in the dual-Youla approach, namely the choice of nominal system $G^0(q)$, which introduces nominal closed loop poles $\{\rho_k\}$ into the variance quantification (33) and hence precludes $\mathrm{Var}\{G_{\rm dy}(e^{j\omega},\widehat{\theta}_N)\}$ from ever being equal to $\mathrm{Var}\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\}$, regardless of the choice of controller.

C. Joint Input-Output Identification

The final closed loop identification strategy to be studied here is the class of so-called ‘joint input-output’ methods, in which two model structures

$$\begin{aligned} y_t &= T(q,\theta)r_t + S(q,\theta)e_t && (34)\\ u_t &= SK(q,\beta)r_t - SK(q,\beta)e_t && (35) \end{aligned}$$

are used. These structures imply the following one-step ahead prediction errors

$$\begin{aligned} \varepsilon_t^y(\theta) &= S^{-1}(q,\theta)\,[y_t - T(q,\theta)r_t]\\ \varepsilon_t^u(\beta) &= SK^{-1}(q,\beta)\,[u_t - SK(q,\beta)r_t] \end{aligned}$$

which are used to find estimates $\widehat{\theta}_N$, $\widehat{\beta}_N$ according to

$$[\widehat{\theta}_N,\widehat{\beta}_N] = \arg\min_{\theta,\beta} V_N(\theta,\beta),$$

$$V_N(\theta,\beta) \triangleq \frac{1}{2N}\sum_{t=1}^{N}[\varepsilon_t^y(\theta)]^2 + [\varepsilon_t^u(\beta)]^2. \qquad (36)$$

Since $\{r_t\}$ is not correlated with $\{e_t\}$, then this implies the equivalent of two open loop estimation problems, with all the attendant advantages of this scenario as mentioned earlier. The estimate of the open loop dynamics is then found via this joint input-output method as

$$G_{\rm jio}(q,\widehat{\theta}_N,\widehat{\beta}_N) = \frac{T(q,\widehat{\theta}_N)}{SK(q,\widehat{\beta}_N)}. \qquad (37)$$

The so-called 'basic' joint input-output approach then involves using a model structure for $T(q,\theta)$ given by (18) and for $SK(q,\beta)$ given by

$$SK(q,\beta) = \frac{\rho_0 + \rho_1 q^{-1} + \cdots + \rho_{m_\rho} q^{-m_\rho}}{1 + \delta_1 q^{-1} + \cdots + \delta_{m_\delta} q^{-m_\delta}}.$$

The model structures for the noise models $S(q,\theta)$ and $SK(q,\beta)$ are assumed to be independently parametrized from the dynamics models $T(q,\theta)$, $SK(q,\beta)$, and such that they can completely describe the true noise models $S(q)(q)$, $S(q)K(q)(q)$ without any pole-zero cancellations.

The variance properties of the resulting basic joint input-output estimate $G_{\rm jio}(q,\widehat{\theta}_N,\widehat{\beta}_N)$ derived from (37), and under the same conditions considered in the previous sections, are now established via the following theorem.

Theorem 4.4: Suppose that the Standing Assumptions 2.1 are satisfied. Suppose further that the model orders chosen for $T(q,\theta)$ satisfy the conditions of Theorem 4.2 so that $A_c(z)\widetilde{A}(z)L(z)$ is a polynomial in z^{-1} of order less than $m_\alpha + m_\beta$. Finally, suppose that the model orders chosen for $SK(q,\beta)$ are such that $A_c(z)A(z)$ is a polynomial in z^{-1} of order less than $m_\delta + m_\rho$. Then using the joint input-output identification method it holds that

$$\lim_{N\to\infty} N\, \mathit{Var}\{G_{\rm jio}\} = |S|^2 \lim_{N\to\infty} N\,\mathit{Var}\{G_{\rm id}\} +$$

$$\left(\frac{\mu}{\mu+\sigma^2}\right)|T|^2 \lim_{N\to\infty} N\,\mathit{Var}\{G_{\rm di}\} + \frac{2\sigma^2}{\Phi_u^r}\mathit{Re}\left\{T\overline{S}\Delta\right\} \quad (38)$$

where $\Delta(\omega)$ will be specified in the discussion following this theorem and

$$\lim_{N\to\infty} N\,\mathit{Var}\left\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\right\},\ \lim_{N\to\infty} N\,\mathit{Var}\left\{G_{\rm id}(e^{j\omega},\widehat{\theta}_N)\right\}$$

are given by (20) with m_α, m_β the same as here, and (12) with $m_a = m_\delta$, $m_b = m_\rho$.

■

As per the previous theorem, this result suggests the following approximate quantification which is 'exact' for finite model order, but is of increasing accuracy with increasing data length N

$$\mathrm{Cov}\{G_{\rm co}\} \approx |S|^2 \mathrm{Var}\{G_{\rm id}\} + \left[\frac{\mu+\sigma^2}{\mu}\right]|T|^2\{G_{\rm di}\} + \frac{2}{N}\frac{\sigma^2}{\Phi_u^r}\mathrm{Re}\left\{T\overline{S}\Delta\right\}.$$

Some comments about this result are again clearly in order. Firstly, the quantification (39) clearly shows that, just as the previous section has established that the variance $\mathrm{Var}\{G_{\rm id}(e^{j\omega},\widehat{\theta}_N)\}$ for the basic indirect method estimate is not, in general, the same as the variance $\mathrm{Var}\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\}$ for a direct methods estimate, the new quantification (38) shows that the variance $\mathrm{Var}\{G_{\rm jio}(e^{j\omega},\widehat{\theta}_N,\widehat{\beta}_N)\}$ for basic joint input-output methods is also *not*, in general, equal to either of these.

In fact, consideration of only the first two terms in (39) illustrates that under the assumptions of Theorem 4.4, $\mathrm{Var}\{G_{\rm co}(e^{j\omega},\widehat{\theta}_N,\widehat{\beta}_N)\}$ is a pseudo-convex combination of $\mathrm{Var}\{G_{\rm id}(e^{j\omega},\widehat{\theta}_N)\}$ and $\mathrm{Var}\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\}$, where the epithet 'pseudo' is used since $S+T=1$ and hence if $\mu \gg \sigma^2$ then

$$|S(e^{j\omega})|^2 + \left(\frac{\mu+\sigma^2}{\mu}\right)|T(e^{j\omega})|^2 \approx 1.$$

Therefore, again ignoring the last $\Delta(\omega)$ term of (39) shows that while at some frequencies where $|S|\approx 1$ or $|T|\approx 1$ then the variance of joint input-output methods estimates might be approximately the same as either a direct method or basic indirect method estimate, it is certainly not in general equal to either of them.

Turning now to the last component of (39), the term $\Delta(\omega)$ needs to be defined. For this purpose, denote by $\{\zeta_k\}$ the zeros of $z^{m_\alpha+m_\beta}A_c(z)\widetilde{A}(z)L(z)$ and by $\{\tau_k\}$ the zeros of $z^{m_\delta+m_\rho}A_c(z)A(z)$ to consider the space

$$V_\tau \triangleq \mathrm{Span}\left\{\frac{e^{j\lambda\ell}}{\prod_{k=1}^{m_\delta+m_\rho}(e^{j\lambda}-\tau_k)}\right\}_{\ell=0}^{m_\delta+m_\rho}. \quad (39)$$

and the bi-variate function

$$\Delta_1(\lambda,\omega) = \sum_{k=1}^{m_\alpha+m_\beta}\frac{1-|\zeta_k|^2}{(e^{j\lambda}-\zeta_k)(e^{-j\omega}-\overline{\zeta_k})}. \quad (40)$$

Finally, for any fixed ω, consider the above $\Delta(\lambda,\omega)$ as a function only of λ. Then $\Delta(\omega)$ is defined to be the L_2 optimal approximant of $\Delta(\lambda,\omega)$ that lies in V_τ.

With this in mind, consider the special case of proportional control in which $K(q) = K \in \mathbf{R}$ so that $\widetilde{A}(z) = A(z)$ and hence, provided that $\mu \gg \sigma^2$ so $\Phi_u^r(\omega)\approx\Phi_u^r(\omega)$ then

$$\mathrm{Var}\{G_{\rm id}\} = \mathrm{Var}\{G_{\rm di}\} \approx \frac{\sigma^2}{N\Phi_u(\omega)}\kappa_{\rm di}(\omega). \quad (41)$$

In this same situation, the sets $\{\tau_k\}$ and $\{\zeta_k\}$ are equal to that $\Delta(\lambda,\omega)$ lies in V_τ for any ω so that

$$\Delta(\omega) = \Delta_1(\omega,\omega) = \Bigg[(m_\beta - m_\alpha) + \sum_{k=1}^{m_\alpha}\frac{1-|\eta_k|^2}{|e^{j\omega}-\eta_k|^2} + \sum_{k=1}^{m_a}\frac{1-|\xi_k|^2}{|e^{j\omega}-\xi_k|^2}\Bigg] = \kappa_{\rm di}(\omega). \quad (42)$$

Consequently, substituting (41) and (42) into (39) indicates that in this special case of proportional control

$$\begin{aligned}\mathrm{Var}\{G_{\rm co}\} &= \mathrm{Var}\left\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\right\}\left[|S|^2+|T|^2+T\overline{S}\right]\\ &= \mathrm{Var}\left\{G_{\rm di}(e^{j\omega},\widehat{\theta}_N)\right\} = \mathrm{Var}\left\{G_{\rm id}(e^{j\omega},\widehat{\theta}_N)\right\}.\end{aligned}$$

V. Simulation Examples

In order to emphasise the significance and ramifications of the new variance quantifications developed in this paper, this section provides a simulation example which considers the system

$$y_t = G(q)u_t + e_t, \qquad G(q) = \frac{1.6177q^2 - 0.74q}{q^2 - 1.8q + 0.81}$$

which, according to (1) is under the influence of the pole cancelling controller

$$K(q) = \frac{q^2 - 1.8q + 0.81}{q^3 - 1.9801q^2 + 0.99q}. \quad (43)$$

This choice of $K(q)$ implies $\widetilde{A}(z) = 1$ according to (6) and delivers closed loop poles at $\eta_1,\eta_2 = 0.5e^{\pm j1.2}$. While this might seem a relatively benign situation, note that the controller poles are at $\ell_1, \ell_2 = 0.995e^{\pm j0.1}$, which are different from the open loop plant poles at $\xi_1, \xi_2 = 0.9$.

Therefore, according to (21) the estimation error involved with direct identification using a Box–Jenkins model structure to first estimate $T(q)$ and $S(q)$, should be different from the estimation error given by (14) for the case of direct identification of $G(q)$ using an output-error structure. Indeed, as shown in Figure 1, simulation confirms this. In that figure the solid lines are the true variability as computed via Monte–Carlo simulation over 1000 estimate realisations, each derived from $N = 1000$ data points with measurement noise variance $\sigma^2 = 10$ and input $\Phi_r(\omega) = 1$. Note that the actual mean square errors agree essentially perfectly with the quantifications (21) and (14) shown (respectively) as dashed and dash-dot lines.

Furthermore, note that as predicted by (21), the location of controller poles near the unit circle has a detrimental effect on indirect estimation accuracy at frequencies on the unit circle very close to those poles. In this case, the increased variability is at 0.1 radians/second which is near the plant crossover frequency.

By way of contrast, the lower (labelled) dash-dot line in Figure 1 is the existing quantification derived in [1], [3] which clearly provides a less informative quantification than (21), (14) and certainly does not expose the true accuracy differences between a direct and indirect approach.

To give a sense of the scale of estimation errors being quantified here, the first ten realisations of the estimate $G(e^{j\omega}, \widehat{\theta}_N)$ are shown in Nyquist plot form in Figure 2. Clearly, the errors involved in this example are non-trivial.

Turning now to the examination of joint input-output methods, consider the 'basic' joint input-output method, the Monte–Carlo based simulation results are shown as the solid line in Figure 3. Also shown on that Figure as a dashed line is the new quantification (38) which, via its very close match to the solid line showing the 'true' variability, is illustrated to be an accurate approximation.

A main point to notice is the qualitative and quantitative difference between these direct, and basic indirect variances when compared to the basic joint input-output method variability shown as the solid line, which further substantiates the theme of this paper that it is not generically true that all indirect estimation methods offer equivalent accuracy when compared to direct estimation methods.

VI. Conclusion

A key conclusion to arise from this work is that it is not generally true that there are negligible differences between the accuracy of the various methods. Furthermore, the accuracy of various indirect methods is not invariant to closed loop configuration, or to user choices such as that of pre-filter (co-prime factor method) or nominal system (dual Youla method).

VII. References

[1] U. Forssell and L. Ljung, *Closed-loop identification revisited*, Automatica, 53 (1999), pp. 1215–1241.

[2] M. Gevers, X. Bombois, et al, *Model validation for control and controller validation in a pred. error id. framework. part i& ii: Theory*, Automatica, 39 (2003), pp. 403–415.

[3] M. Gevers, L. Ljung, and P. Van den Hof, *Asymptotic variance expressions for closed loop identification*, Automatica, 37 (2001), pp. 781–786.

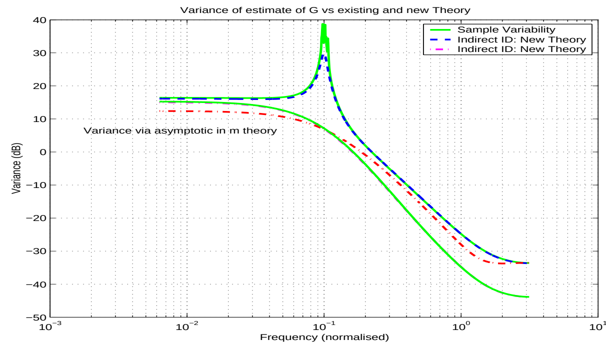
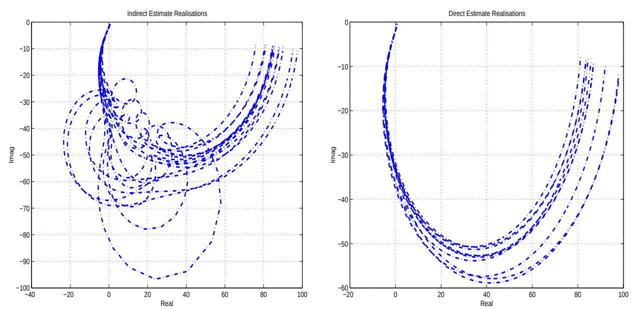


Fig. 1. *Variability of direct and indirect estimates. Solid lines:true variabilities, dashed and dash dot lines: new theoretical quantifications of those same variabilities. Lower dash dot line is pre-existing quantification.*



(a) Indirect Estimation (b) Direct Estimation

Fig. 2. *First twenty estimate realisations via direct and indirect id.*

[4] H. Hjalmarsson, M. Gevers, and F. De Bruyne, *For some model based control design criteria, closed loop identification gives better performance*, Automatica, 32 (1996), pp. 1659–1673.

[5] E. Lehmann, *Theory of Point Estimation*, John Wiley & Sons, 1983.

[6] L.Ljung, *Asymptotic variance expressions for identified black-box transfer function models*, IEEE Transactions on Automatic Control, AC-30 (1985), pp. 834–844.

[7] B. Ninness and H. Hjalmarsson, *On the freq. dom. accuracy of closed loop estimates,*, tech. rep., School of Elec. Eng, Uni. Newcastle, 2003.

[8] B. Ninness and H. Hjalmarsson, *Variance error quantifications that are exact for finite model order*, Submitted to IEEE Transactions on Automatic Control, (2003).

[9] P. Van den Hof, *Closed-loop issues in sys. id.*, Annual Reviews in Control, 22 (1998), pp. 173–186.

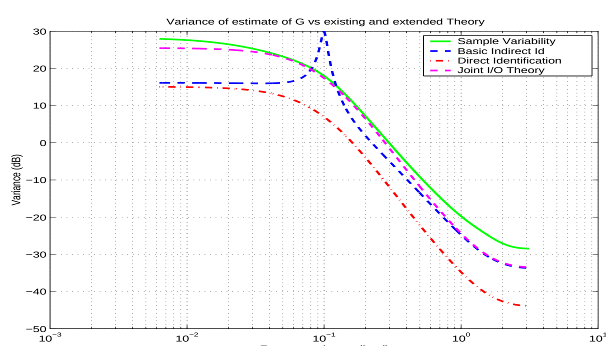


Fig. 3. *Joint input-output vs. direct and indirect id.*