COMBINING MODELS IN SYSTEM IDENTIFICATION

Eşref Eşkinat
Department of Mechanical Engineering, Boğaziçi University, Istanbul, Turkey
E-mail: eskinat@boun.edu.tr

Abstract: Patching different kinds of models obtained from experimental data is considered in this paper. The problem of combining different experiments and models is analysed. As a related issue, iteratively improving models and combining models for control oriented identification is investigated. A simulated example is given to illustrate the approach. *Copyright © 2000 IFAC*

Keywords: Identification, Iterative model improvement, Internal model control

1. INTRODUCTION

The idea of iterative closed loop identification and control design has been exploited in several recent papers. A review is given in (Gevers, 1993). In (Aström, 1993) pole placement controller design is applied to models obtained from ARX and ARMAX model structures. In (Zang et.al., 1995), LQG design combined with output error (OE) model structure is used. Iterative identification, control, and performance improvement using IMC control design and closed loop identification using the fractional representation of the model is pursued in (Lee et.al., 1995), and is called the 'Windsurfer Approach' to adaptive control. An interesting approach which bypasses the modeling phase altogether and identifies controllers from closed loop data is given in (Hjalmarsson et.al. 1994).

Behind these methodologies lies the intuitively appealing notion that the data for model identification should be obtained under the conditions of the actual use, i.e., closed loop control. Although it is impossible to have the ideal controller in the loop (because the true plant is never known), by iterative identification and control, steps are taken towards reaching the ideal situation. In these schemes the criterion used in identification is merged to the overall performance criterion through the use of the pre-filters which are designed using the information obtained about the system up that time instant. The pre-filters act as a weighing function on the data, and therefore weight the frequency regions that would be more important for control. A criticism of these schemes is that (Ljung, 1997), the data obtained from previous experiments is not utilized in obtaining the current model and the controller. This is partly because, these methods put more emphasis on bias errors, and variance errors are not directly taken into account. If the means of obtaining data from the system is easy, then the previous experiments may be discounted. However, this is usually not the case. Given enough time, theoretically it is possible to identify any linear system with arbitrarily small error. However when the time and input power is limited, it becomes important to utilize the resources effectively to obtain the 'best' possible model for our needs.

In order to overcome the limitations imposed by the bias error in system identification, we might develop local models around certain frequency regions, which have negligible bias error. These models may be combined to obtain a final model. An extreme case of local modeling is sine wave testing. Since sine wave testing too much time, models estimated with inputs concentrated in certain important frequency regions might be very useful. This is the approach followed in this paper. Combining of the models is done by weighting of them in the frequency domain. The selection of the weights reflects our belief in the model, and may be done in an ad-hoc manner. However, using some approximations, it can be shown that models can be weighted in the frequency domain inversely proportional to their variances. Such weighting minimizes the variance of the final model. The issue of iterative experiments is also discussed. Using IMC parameterization, the design of external inputs for further closed and open-loop experiments are done such that the variance of the model error is minimized. Combining the models with the ones obtained from previous experiments has some advantages in terms of the variance of the model error, compared to the iterative identification-control methodologies.

2. COMBINING MODELS

Suppose M estimates of the transfer function $G(e^{i\omega})$ are available. These estimates may be obtained from separate experiments or may be known a-priori considerations, such as steady state balances. In addition, certain a-priori information about the system in terms of isolated frequency response points might also be available. Such knowledge may be obtained from sine wave tests. The method used in obtaining the estimates is immaterial; the individual estimates may be considered as translation of the knowledge we may have about the system into frequency domain. A natural way to combine these estimates is to obtain a weighted average of them as:

$$\overline{G}(e^{i\omega}) = \sum_{j=1}^{M} W_j(\omega) G_j(e^{i\omega}) \quad (1)$$

In order to minimize the variance of the final estimate $Var(\overline{G}(e^{i\omega}))$, it is well known that the weights $W_j(\omega)$ have to be selected as inversely proportional to the variance of the individual transfer function estimates $G_j(e^{i\omega})$ (Ljung, 1987). This can easily be shown by expressing the variance of $\overline{G}(e^{i\omega})$ in equation (1) in terms of variances of $G_j(e^{i\omega})$, and differentiating the variance expression with respect to $W_j(\omega)$. In that case, the individual weights $W_j(\omega)$ will be given by:

$$W_j(\omega) = \frac{\left(Var(G_j(e^{i\omega})\right)^{-1}}{\sum_{j=1}^{M}\left(Var(G_j(e^{i\omega})\right)^{-1}} \quad (2)$$

where, variance of the transfer function is:

$$Var\left[G(e^{i\omega})\right] = E\left|\hat{G}(e^{i\omega}) - G^*(e^{i\omega})\right|^2 \quad (3)$$

with $G^* = E\{\hat{G}\}$ being the limit model. Note that, the model set does not have to contain the true system in it, for equation (2) to be true. Therefore, the results are correct even for biased estimates. The variance of the combined transfer function estimate will be given by:

$$\left[Var(\overline{G}(e^{i\omega}))\right]^{-1} = \sum_{j=1}^{M}\left[Var(G_j(e^{i\omega}))\right]^{-1} \quad (4)$$

and:

$$\overline{G}^*(e^{i\omega}) = \sum_{j=1}^{M} W_j(\omega) G_j^*(e^{i\omega}) \quad (5)$$

It is clear that $Var(\overline{G}(e^{i\omega})) \leq Var(G_j(e^{i\omega}))$. If need be, a parametric model can be fitted to the transfer function estimate $\overline{G}(e^{i\omega})$ using a frequency domain parameter estimation algorithm (Schoukens and Pintelon, 1991). In this case, the weights at each frequency can be taken to be proportional to the estimated variance $Var(\overline{G}(e^{i\omega}))$.

If the prediction error methods are used in obtaining the individual estimates, the covariance of the estimated parameters θ_j may be used to obtain the variance of the transfer function. In the prediction error framework, the model to be estimated has the following form:

$$y(t) = G(q)u(t) + H(q)e(t) \quad (6)$$

$$E\{e^2(t)\} = \lambda^2 \quad (7)$$

The parameters θ defining $G(q)$ and $H(q)$ may be estimated by minimizing the cost function:

$$V(\theta) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t,\theta)^2$$
$$\hat{\theta}_N = \arg\min V(\theta) \quad (8)$$

where, the prediction errors and the covariance of the parameter estimates is given by (Ljung, 1987):

$$\varepsilon(t,\theta) = \frac{1}{H(q)}\left(y(t) - G(q)u(t)\right) \quad (9)$$

$$Cov(\hat{\theta}_j) = \frac{\lambda_j^2}{N_j}\left[E(\psi_j(t)\psi_j^T(t)\right]^{-1} \quad (10)$$

where $\psi_j(t)$ is the negative gradient of the prediction error $\varepsilon_j(t,\theta)$ with respect to θ_j.

$$\psi_j(t) = -\frac{\partial \varepsilon(t,\theta)}{\partial \theta}$$

Then, if the bias error is negligible, the variance of the transfer function can be (approximately) expressed as:

$$Var\left[G(e^{i\omega})\right] \cong \left(\frac{\partial \hat{G}_j}{\partial \theta}\right) Cov[\hat{\theta}_j] \left(\frac{\partial \hat{G}_j}{\partial \theta}\right)^* \quad (11)$$

Alternatively, for large values of model order n and the number of data N, it has been shown that (Ljung,1987) the variance of the transfer function may be expressed as:

$$Var(\hat{G}(e^{i\omega})) \cong \frac{n}{N}\frac{\Phi_v(\omega)}{\Phi_u(\omega)} \quad (12)$$

This expression is valid even when there is considerable bias error in the model. However, it is asymptotic in both data and the model order. The expression (12) gives good approximation to the true variance of the model, even for finite data and finite model order, as demonstrated in (Ljung, 1987). The noise spectrum $\Phi_v(\omega)$ may be obtained from:

$$\Phi_v(\omega) = \lambda^2 \left| H(e^{i\omega}) \right|^2 \tag{13}$$

The noise variance λ^2 may be estimated from the data, as the sum of square of the prediction errors. Another possible approach is to estimate the variance considering the undermodeling, as done in (Hjalmarsson, 1993) and (Ninnes and Goodwin, 1994). An important point is, in order to use expressions as equation (12), good estimates of the noise model *H(q)* are needed, in addition to the system model *G(q)*. So, the model structures that do not have an adequate disturbance model, (i.e., output error structure) can not be used.

A large number of model types can be combined in such a framework. The different models to be combined do not have to have the same order, only necessary information is their frequency response and the estimated variance at these frequency response points. It is possible to combine time domain models with frequency domain experiments performed at certain frequencies. We may also design different experiments at higher and lower frequencies, fit models to them, and then combine them. Performing iterative experiments, in which input design for further experiments is done using the previous models, is another possibility. This is considered in the next section.

3. ITERATIVE MODEL IMPROVEMENT FOR CONTROL

The approach presented above may also be used for control oriented identification. Iterative experiments used to obtain individual models can then be combined to arrive at a final model. After one model is obtained, a controller may be designed, and then further experiments may be performed either in closed loop, or further input signals may be designed so that they are concentrated in the frequency region where input power needs to be large (a gain increase is desired, effect of disturbance are large etc.).

The above-explained ideas can be formalized using the Internal Model Control (Morari and Zafiriou, 1989) (IMC) parameterization of the closed loop system. In IMC formalism (Figure 1), the system output *y(t)* in closed loop is given by:

$$y(t) = \frac{GQ}{1+\tilde{G}Q} r(t) + \frac{1+\hat{G}Q}{1+\tilde{G}Q} v(t) \tag{14}$$

with $\tilde{G} = G - \hat{G}$. G is the true system and $\hat{G}$ is the model. The response given by the 'perfect' closed loop system is:

$$y^*(t) = GQr(t) + (1-GQ)v(t) \tag{15}$$

The error to be minimized in identification can be expressed as:

$$\tilde{y}(t) = y^*(t) - y(t) = \frac{T^*\tilde{G}Q}{1+\tilde{G}Q}(r(t) - v(t)) \tag{16}$$

where $T^* = GQ$ is the ideal closed loop transfer function and the second equality follows after some algebraic manipulation. Assuming *r(t)* and *v(t)* to be independent, and using Parseval's relation, the equation (16) can be expressed in frequency domain as:

$$\Phi_{\tilde{y}}(\omega) = \frac{\left|T^*\right|^2 \left|Q\right|^2}{\left|1+\tilde{G}Q\right|^2} \left|\tilde{G}\right|^2 \left\{\Phi_r(\omega) + \Phi_v(\omega)\right\} \tag{17}$$

The expected value of $\left|\tilde{G}\right|^2$ may be expressed as:

$$E\left\{\left|\tilde{G}\right|^2\right\} = Cov(\hat{G}) = \frac{n}{N} \frac{\left|1+\tilde{G}Q\right|^2}{\left|Q\right|^2} \frac{\Phi_v(\omega)}{\Phi_r(\omega)} \tag{18}$$

for large N and n. The second equality is from (Forsell and Ljung, 1998) and the fact that in IMC parameterization we have:

$$u(t) = \frac{Q}{1+\tilde{G}Q}\left(r(t) - v(t)\right) \tag{19}$$

Therefore, using the criterion J to be minimized in the identification experiment can be written as:

$$\begin{aligned} J &= E\left\{(\tilde{y}^2)\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{\tilde{y}}(\omega) d\omega \\ &= \frac{1}{2\pi}\frac{n}{N}\int_{-\pi}^{\pi} \left|T^*\right|^2 \frac{\Phi_v(\omega)}{\Phi_r(\omega)} \left\{\Phi_r(\omega) + \Phi_v(\omega)\right\} d\omega \end{aligned} \tag{20}$$

If the input power is constrained as:

$$\int_{-\pi}^{\pi} \Phi_r(\omega) d\omega \le \alpha \tag{21}$$

Then, the optimal input to the system is given by (Ljung, 1987):

$$\Phi_r^{opt}(\omega) = \mu \left|T^*\right| \Phi_v(\omega) \tag{22}$$

where μ is a constant so that (21) becomes an equality. The expression (22) contain unknown terms such as G and $\Phi_v(\omega)$, so input design may be done iteratively as:

$$\Phi_r^i(\omega) = \mu_i \left|T_i^*\right| \Phi^i{}_v(\omega) \tag{23}$$

If the experiments are to be done in open loop, substituting (22) to (19), the following expression can obtained:

$$\Phi_u^{opt}(\omega) = \delta \frac{|Q|^2}{|1+\tilde{G}Q|^2} \left\{ |T^*| + 1 \right\} \Phi_v(\omega)$$
$$= \delta \frac{|T|^2}{|G|^2} \left\{ |T^*| + 1 \right\} \Phi_v(\omega) \qquad (24)$$

δ is a constant, which is adjusted so that input power constraints are satisfied.

4. AN EXAMPLE

Consider the following system:

$$G(s) = \frac{2}{s+1} \frac{229}{s^2 + 30s + 229}$$

Sampling with T=0.04, the discrete model is:

$$G(q) = \frac{0.0036q^{-1} + 0.0107q^{-2} + 0.0019q^{-3}}{1 - 2.0549q^{-1} + 1.3524q^{-2} - 0.2894q^{-3}}$$

The system was simulated using two different inputs and additive white noise of variance 0.01. The first input covers the lower frequencies and its power is concentrated between frequencies 0 to 0.05(π/T). Second input covers the frequency range between 0.2(π/T) and π/T. The estimated spectra of the inputs are shown in Figure 2. Amplitude of the inputs is so adjusted that they have the same total power.

Two tests with these inputs are done with 2000 data points. 500 Monte Carlo simulations with different noise realizations are done. The presented results are the averages of these simulations. Two 10th order ARX models were estimated, one based on the low frequency, the other on the high frequency input. The estimated transfer functions, along with the true system are shown in Figure 3. For G_1, model is good at lower frequencies. On the other hand, G_2 gives better estimate of the system at the higher frequencies, but is poorer at lower frequencies. This is obviously related to the concentration of the input power to the lower (u_1) and higher (u_2) frequencies. The estimated variances of these models, obtained using expression (12) are given in Figure 4. The combined transfer function estimate, obtained using expression (4) and (5) is shown, together with the frequency response of the true system, in Figure 5. We can see that there is a good agreement, despite the fact that individual estimates of G_1 and G_2 are not that good. A simple frequency domain parameter estimation algorithm, which is a combination of Sanathanan-Koerner iteration and weighted least squares in frequency domain, gives a parametric transfer function estimate, whose frequency response is shown along with the true system in Figure 5. Clearly, there is good agreement.

As another numerical experiment, the second input was generated under feedback, using the IMC controller and the reference input was generated using expression (23). IMC filter Q was designed by inverting the minimum phase part of the estimate G_1 and by multiplying it by a filter F so that Q is proper (Morari and Zafiriou, 1989). The designed closed loop system has an approximate bandwidth of 15 rad/s. Input spectra, estimated transfer functions, their variances and the combined estimate are given in Figures 6 to 9 for this case. It can be seen that (Figure 6) the second input u_2 generated under feedback has its maximum power around the bandwidth of the closed loop system. This is reflected in the estimated variance shown in Figure 8. Although the combined estimate in Figure 9 does not seem as good as the one in Figure 5, the closed loop responses are indistinguishable.

5. DISCUSSION AND CONCLUSIONS

In this paper, the well-known inverse variance weighting has been employed to combine different estimates of a transfer function. A simple to obtain, but important result is that, the weighting scheme is proposed is valid under quite general conditions, i.e., when the model set does not contain the true system. Therefore, the models that have considerable bias error in the process model may be combined, to obtain a final model that has lower variance. However, in general, since both bias and variance errors are proportional to the inverse of the input spectrum at a given frequency, combining models may be expected to reduce bias error as well.

A large class of models may be combined using the inverse variance weighting. Of particular importance is the combination of models for control-oriented identification. The extra, but useful knowledge obtained from a relay test (Aström and Hagglund, 1988), which gives a reliable estimate of the ultimate gain and period of the system, may be combined with other tests, to obtain models useful for control. This might especially be useful for PID based controller designs. Combining the data from the main experiment with probing sine wave tests at certain important frequencies, is another possibility. The recently developed methods of ‘identification for control’ can also benefit from combining models. This fact was pointed out in (Ljung, 1997).

One problem with inverse variance weighting in combining models is that, the estimated, instead of the true variance (which is unknown) has to be used in weighting. If the estimates of the variance, obtained from expressions such as equations (11) or (12) are not representative of the actual variance, results obtained may not make sense. The accuracy of the variance expressions depends on the disturbance models.

Therefore, we need to estimate good disturbance models, in addition to the process model.

REFERENCES

Aström K.J. (1993) Matching Criteria for Control and Identification, Proceedings of 2nd ECC pp. 248-251.

Aström K.J., Hagglund T., (1988) *Automatic Tuning of PID Controllers*, ISA Press.

Forsell U., Ljung L. (1998) Identification for control: Some results on optimal experiment design. Proceedings of the 37th CDC pp 3384-3389

Gevers M. (1993) Towards a Joint Design of Identification and Control? In '*Essays on Control' Trentelman* H.L., Willems J.C. Editors, Birkhauser. Pp. 111-151.

Hjalmarsson H. (1993) A Model Variance Estimator. Proceedings of IFAC Triennial World Congress, Sydney, pp.335-340

Hjalmarsson H., Gunnarsson S., Gevers M. (1994) A Convergent Iterative Restricted complexity Control Design Scheme Proceedings of 33rd CDC pp. 1735-1740.

Lee W.S., Anderson B.D.O., Mareels M.Y., Kosut R.L. (1995) On Some Key Issues in the Windsurfer Approach to Adaptive Robust Control. Automatica v.31, pp. 1619-1636.

Ljung L. (1987) *System Identification*. Prentice Hall.

Ljung L. (1997) Identification, Model Validation and Control, Slides of the CDC-97 Plenary Lecture.

Morari M., Zafiriou E. (1989) *Robust Process Control*, Prentice Hall.

Ninnes B., Goodwin G. (1994) Estimation of Model Quality. Proceedings of IFAC Symposium on System Identification. Copenhagen, pp. 25-44.

Schoukens J., Pintelon R. (1991) Identification *of Linear Systems*. Pergamon Press.

Zang Z., Bitmead R., Gevers M. (1995) Iterative Weighted Least Squares Identification and Weighted LQG Control Design. Automatica v.31, pp. 1577-1594.

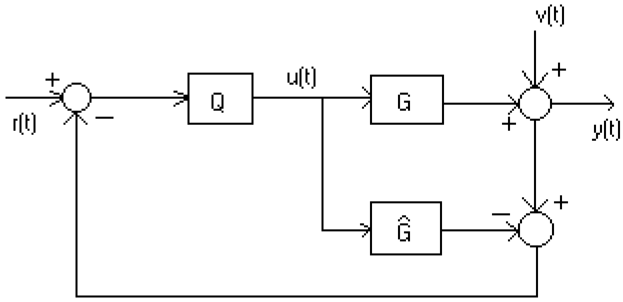

Figure 1. IMC Parameterization.

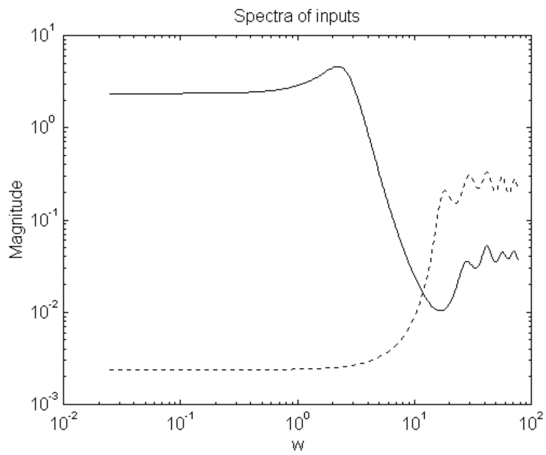

Figure 2. Spectra of inputs: Solid line $\Phi_{u1}(\omega)$, dashed line $\Phi_{u2}(\omega)$.

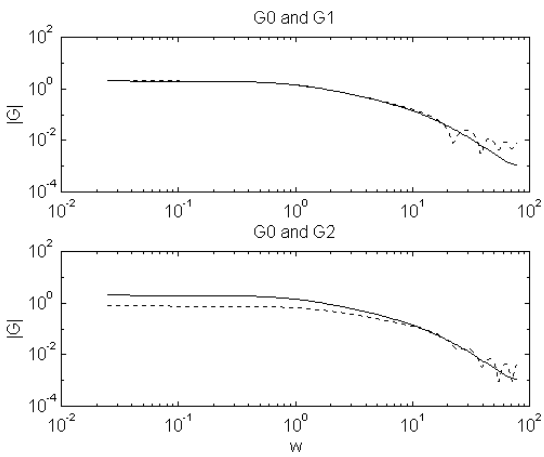

Figure 3. True (solid) and estimated (dashed) transfer functions.

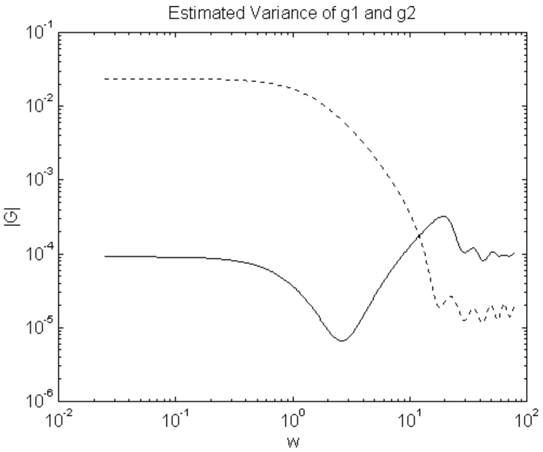

Figure 4. Estimated Variance of transfer functions: variance of G1(solid), and variance of G2 (dashed) from expression (12).

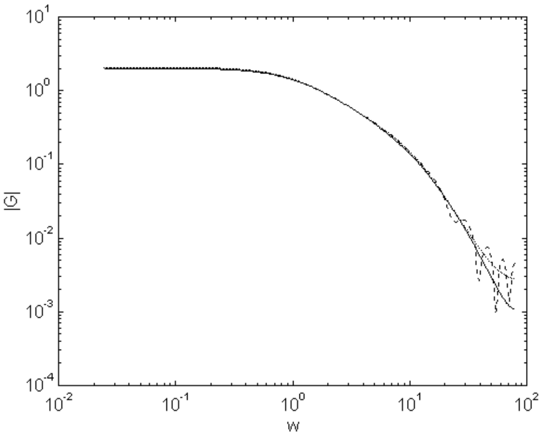

Figure 5. True transfer function (solid), combined estimate (dashed) and parametric estimate (dotted).

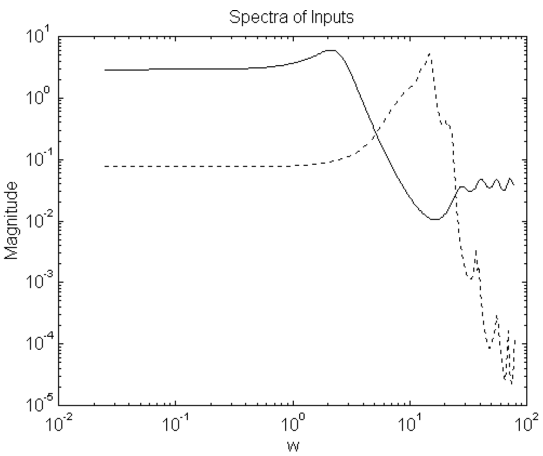

Figure 6. Spectra of inputs for closed loop experiment: Solid line $\Phi_{u1}(\omega)$, dashed line $\Phi_{u2}(\omega)$.

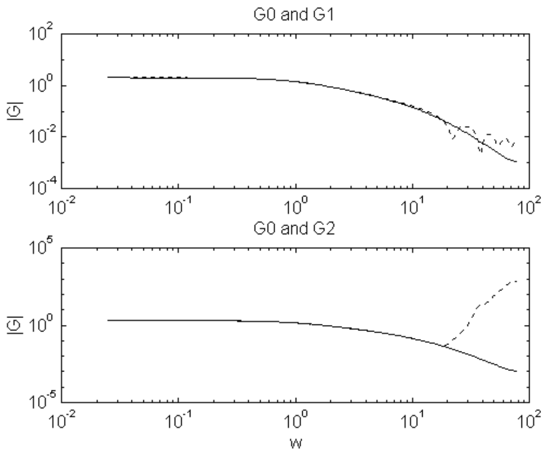

Figure 7. True (solid) and estimated (dashed) transfer functions for closed loop experiment.

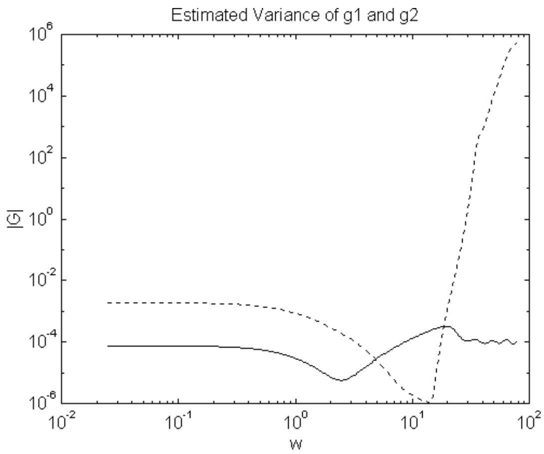

Figure 8. Estimated Variance of transfer functions: variance of G1(solid), and variance of G2 (dashed).

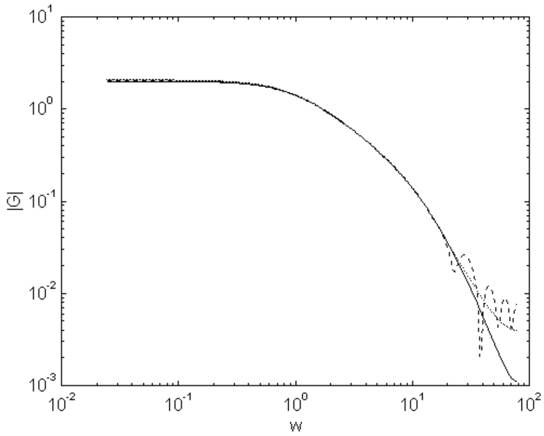

Figure 9. True transfer function (solid), combined estimate (dashed) and parametric estimate (dotted) for the closed loop experiment.