

Identification of Probabilistic System Uncertainty Regions by Explicit Evaluation of Bias and Variance Errors

Richard G. Hakvoort and Paul M. J. Van den Hof, *Senior Member, IEEE*

Abstract—A procedure is developed for identification of probabilistic system uncertainty regions for a linear time-invariant system with unknown dynamics, on the basis of time sequences of input and output data. The classical framework is handled in which the system output is contaminated by a realization of a stationary stochastic process. Given minor and verifiable prior information on the system and the noise process, frequency response, pulse response, and step response confidence regions are constructed by explicitly evaluating the bias and variance errors of a linear regression estimate. In the model parameterizations, use is made of general forms of basis functions. Conservatism of the uncertainty regions is limited by focusing on direct computational solutions rather than on closed-form expressions. Using an instrumental variable method for identification, the procedure is suitable also for input-output data obtained from closed-loop experiments.

Index Terms—Robust control, system identification, uncertainty modeling.

I. INTRODUCTION

IN MANY situations where identified models are used to reflect the dynamic properties of an underlying dynamical system, there is interest not only in the specification of one single nominal model, but simultaneously in a quantification of the corresponding model uncertainty. In other words, the question “how reliable is my identified model?” is a question which is very relevant in many model applications. One of the most apparent examples of this is in model-based robust control design. Here, robust controller analysis and synthesis requires the availability of an upper bound on the model-plant mismatch, specified in some norm [7].

In “classical” prediction error identification [26], [38] it is possible to quantify asymptotic confidence intervals for estimated parameters and frequency responses, however only in the situation that there is no unmodeled dynamics. This situation, referring to the case that we can exactly model the plant under study by a (low-order) finite-dimensional linear time-invariant model is considered to be overly optimistic.

Motivated by applications in model-based robust control theory, many authors have considered the problem of deriving

system uncertainty regions on the basis of measurement data and prior assumptions. Next to the problem of dealing with unmodeled dynamics, a second goal here has been to arrive at hard-bounded (worst case deterministic) uncertainty regions, providing the user with 100% certainty about the dynamic properties of the system at hand. This is in contrast with a probabilistic formulation of system uncertainty, as present in the classical methods.

It has been clearly motivated [20], [27] that the worst case deterministic versus probabilistic issue is completely determined by the prior assumptions concerning the noise disturbance on the data.

The worst case deterministic type of problem has been addressed in terms of frequency response data in [5], [10], [17], [18], [25], [33], and [34]. In the case of time-domain data, a deterministic/worst case approach is often referred to as set-membership identification or as parameter bounding identification. Accounts are given in [8], [30], and [32]. Several norms are used to outerbound the obtained parametric uncertainty sets, as e.g., in terms of frequency response magnitude [42], \mathcal{H}_∞ -norms [24], [44], and ℓ_1 -norms [23], [28], [39]. Direct outerbounds on the frequency response of the system are considered in [11] and [12]. Contributions based on model (in)validation are provided in [35] and [36]. An important drawback is that due to the worst case character of the assumed disturbances, the obtained system uncertainty regions generally are very conservative, if this worst case disturbance does not actually occur. This is shown in [20] and [31].

Approaches that consider disturbance signals to be stochastic and that also account for undermodeling are given in [1], [6], [9], and [45]. In this approach the noise is assumed to behave noisy, i.e., random and uncorrelated to the input signal, and prior assumptions about the system are made, varying from deterministic [1], [6] to stochastic [9]. Typically these procedures yield frequency response uncertainty regions which are correct with a certain specified probability, provided the prior assumptions that are made are correct. The current limitations of these methods are that they are restricted to periodic input signals [1], [6] or that a probabilistic embedding of the unmodeled dynamics has to be specified *a priori* [9].

In this paper we present a new procedure for the identification of probabilistic system uncertainty regions that properly handles effects of bias and is not limited to periodic input excitation. Instead of deriving closed form expressions for model error bounds, our procedure is designed for arriving

Manuscript received May 19, 1995; revised January 20, 1997 and May 20, 1997. Recommended by Associate Editor, J. C. Spall.

R. G. Hakvoort was with the Mechanical Engineering Systems and Control Group, Delft University of Technology, 2628 CD Delft, The Netherlands. He is now with Royal Dutch/Shell Group, KSEPL, Rijswijk, The Netherlands.

P. M. J. Van den Hof is with the Mechanical Engineering Systems and Control Group, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: p.m.j.vandenhof@wbmt.tudelft.nl).

Publisher Item Identifier S 0018-9286(97)08158-0.

at computational solutions by explicitly calculating bias and variance errors of a linear regression [least squares (LS) or instrumental variable (IV)] estimate. In this way we are able to arrive at expressions for system uncertainty in terms of frequency response, pulse response, and step response, for any user-chosen sequence of frequency points and time indexes, respectively. As a result, conservatism induced by upper-bounding the errors is limited. We will adopt the classical prediction error framework with stochastic noise disturbance on the data; the required prior information on the system is limited and is shown to be partly constructible from the data.

We will employ a linear model parameterization in terms of general basis functions. Recent work on Laguerre, Kautz, and generalized functions has shown to have great potentials in identification problems; see [19], [31], [40], [41], and [43]. In this way approximate knowledge about pole locations of the unknown system can be incorporated by the choice of appropriate basis functions, leading to a model parameterization in terms of a series expansion that exhibits an increased rate of convergence.

The outline of the paper is as follows. In the next section the identification setting is described. Section III presents the least squares/instrumental variable estimate. In Section IV, the frequency response error of the estimated model is evaluated, which leads to probabilistic frequency response system uncertainty regions. In the subsequent sections pulse response and step response uncertainty regions are derived. In Section VII, it is shown how the prior information that is required for our procedure can be identified from the data itself. Next, in Section VIII, a simulation example is shown. The present paper is an extended version of [15]. A worst case deterministic counterpart of the procedure is presented in [14].

Concerning notation, \mathcal{H}_2 will denote the Hilbert space of complex functions analytic in $|z| > 1$ and squared integrable on the unit circle; \mathcal{H}_∞ is the Hardy space of complex functions that are analytic and bounded in $|z| > 1$; \mathbb{R} is the set of real-valued scalars; ℓ_∞ is the set of bounded signals, i.e., $\sup_t |u(t)| \leq c \in \mathbb{R}$.

II. IDENTIFICATION SETTING

We consider a data generating system $G_0(z) \in \mathcal{A}$, the class of proper, bounded-input/bounded-output (ℓ_∞ -stable) linear time-invariant systems, generating input–output data according to

$$y(t) = G_0(q)u(t) + v_0(t) \quad (1)$$

with q the forward shift: $qu(t) = u(t+1)$, u , and y scalar input and output signals and v_0 a noise process satisfying the following assumption which is standard in prediction error identification (cf. [26]).

Assumption 2.1: The noise process v_0 is a stationary stochastic process with auto-covariance function $R_{v_0}(\tau) = Ev_0(t+\tau)v_0(t)$ and it satisfies $v_0(t) = H_0(q)e_0(t)$ for some ℓ_2 -stable $H_0(q)$, where $\{e_0(t)\}$ is a sequence of independent random variables with zero mean values, variances σ_e , and bounded fourth moments.

The input signal is assumed to be a quasi-stationary signal [26] and may be dependent on the noise process e_0 so as to allow for experimental data that is obtained from closed-loop experiments.

A data set taken from the system G_0 will typically consist of data sequences $\{u(t)\}_{t=1, \dots, N}$, $\{y(t)\}_{t=1, \dots, N}$, where N denotes the length of the observation interval.

In order to cope with unknown initial conditions the input signal in the past is assumed to be bounded by

$$|u(t)| \leq \bar{u}, \quad \forall t \leq 0 \quad (2)$$

for some given \bar{u} .

When characterizing uncertainty sets concerning the system dynamics on the basis of the available data sequences, use will be made of representations of the system $G_0 \in \mathcal{A}$ in the form of a general series expansion.

As $\mathcal{A} \subset \mathcal{H}_\infty \subset \mathcal{H}_2$, it follows that for $\{F_k(z)\}_{k=0, \dots, \infty}$ being a prechosen sequence of basis functions in \mathcal{H}_2 , any system $G_0 \in \mathcal{A}$ can be uniquely written as

$$G_0(z) = \sum_{k=0}^{\infty} g_0(k)F_k(z). \quad (3)$$

In the function space \mathcal{H}_2 , many choices for (orthonormal) basis functions exist, as e.g., the standard pulse basis $F_k(z) = z^{-k}$, the Laguerre basis $F_k(z) = \sqrt{1-a^2}z(1-az)^k/(z-a)^{k+1}$ for any $a \in \mathbb{R}$, $|a| < 1$, and the two-parameter Kautz basis which is a second-order generalization of the Laguerre basis. Recently, the latter two bases are exploited in terms of system identification of expansion coefficients; see [41] and [43]. More general constructions of orthonormal basis functions have been introduced in [19] and used for identification purposes in [40]. The most important phenomenon to notice here is that an appropriate choice of basis functions can essentially improve the rate of convergence in a series expansion (3). To this end, the basis functions introduced in [19] can incorporate dynamics of any complexity (order) and thus extend the Laguerre case where one single pole can be positioned in the functions. In this paper we will consider the functions $F_k(z)$ to be chosen *a priori*.

As *a priori* information on the data generating system G_0 , the coefficients $g_0(k)$ are assumed to be bounded by

$$|g_0(k)| \leq \bar{g}(k), \quad k = 0, \dots, \infty \quad (4)$$

for given $\bar{g}(k) \in \mathbb{R}$ that satisfies an exponential decay, i.e., there exist *a priori* given numbers $M, \rho, k^* \in \mathbb{R}$ with $M, k^* \geq 0$, $\rho < 1$, such that

$$\bar{g}(k) \leq M\rho^k, \quad \forall k > k^*. \quad (5)$$

Methods to verify these prior bounds on the basis of measurement data will be discussed later on.

As we will also be dealing with bounding the uncertainty in the pulse and step response of the system G_0 , we denote $G_0(z) = \sum_{k=0}^{\infty} p_0(k)z^{-k}$, with $\{p_0(k)\}_{k=0, \dots, \infty}$ the pulse response sequence of G_0 , while $\{s_0(k) := \sum_{i=0}^k p_0(i)\}_{k=0, \dots, \infty}$ refers to the step response sequence of G_0 .

The problem that will be considered is formulated as follows.

Given are the prior information (4), (5) on G_0 , the stochastic nature of the noise process as formulated in Assumption 2.1, and the bound on the past input signal (2). Also given is a sequence of input and output measurement data of length N .

- For any user-chosen frequency $\omega \in [0, \pi]$, determine a region in the complex plane that will contain $G_0(e^{i\omega})$ asymptotically with a prescribed probability.
- For any user-chosen time index $k \geq 0$, determine a real-valued interval that will contain $p_0(k)$ [respectively, $s_0(k)$] asymptotically with a prescribed probability.

This problem will be handled along the following lines.

The system G_0 is split into two parts: $G_0(z) = \tilde{G}_0(z) + \bar{G}_0(z)$, with

$$\tilde{G}_0(z) = \sum_{k=0}^n g_0(k)F_k(z) \quad (6)$$

$$\bar{G}_0(z) = \sum_{k=n+1}^{\infty} g_0(k)F_k(z) \quad (7)$$

for some user-defined truncation value n .

For solving the first part of the problem, we are going to identify the coefficients of \tilde{G}_0 from measurement data. As a result, probabilistic uncertainty bounds will be derived for $\tilde{G}_0(e^{i\omega})$ using asymptotic variance expressions of a linear regression-type estimate. These variance expressions are based on the stochastic noise Assumptions 2.1. In the variance expressions the influence of the undermodeling part $\bar{G}_0(z)$ is properly taken into account. Additionally, deterministic uncertainty bounds will be determined for the tail $\bar{G}_0(e^{i\omega})$, using the deterministic prior bounds $\bar{g}(k)$ given in (4). The sum of the deterministic uncertainty bounds for $\bar{G}_0(e^{i\omega})$ and the probabilistic uncertainty bounds for $\tilde{G}_0(e^{i\omega})$ provides probabilistic uncertainty regions for the system $G_0(e^{i\omega})$. A similar procedure will be followed to solve the second part of the problem, related to the pulse (step) response coefficients $p_0(k)$ ($s_0(k)$).

The truncation order n is chosen so as to minimize the uncertainty bounds. If n is chosen too small, the resulting bounds will be heavily determined by the prior information (4), which is generally conservative. If n is chosen too large, the confidence regions for $\tilde{G}_0(z)$ will be large as the variance increases with the number of parameters to be estimated. An optimal choice of the truncation order will be commented upon later.

III. AN INSTRUMENTAL VARIABLE ESTIMATE

In this section we will formulate the linear regression/instrumental variable estimate that will be used for that part of the system (\tilde{G}_0) that will be parametrically identified.

We will consider the model structure

$$\varepsilon(t, \theta) = y(t) - G(q, \theta)u(t) \quad (8)$$

with

$$G(z, \theta) = \sum_{k=0}^n g(k)F_k(z) \quad (9)$$

where $\varepsilon(t, \theta)$ is an (parameterized) output error signal, $G(z, \theta)$ is the parameterized model, while the model parameters $\{g(k)\}_{k=0, \dots, n}$ are collected in the parameter vector $\theta = [g(0) \ g(1) \ \dots \ g(n)]^T \in \mathbb{R}^{n+1}$. In calculating the output error signal, we will adopt the convention that $u^+(t) := u(t)$ for $t > 0$, and $u^+(t) = 0$ for $t \leq 0$.

The following signals are defined:

$$x_k(t) := F_k(q)u^+(t) \quad (10)$$

$$z_k(t) := F_k(q)r(t) \quad (11)$$

with r an instrumental variable signal, correlated with the input u but not correlated with the noise process e_0 . Now denoting the column-vectors

$$\phi(t) := \begin{bmatrix} x_0(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \zeta(t) := \begin{bmatrix} z_0(t) \\ \vdots \\ z_n(t) \end{bmatrix} \quad (12)$$

the instrumental variable estimate $\hat{\theta}_N$ is obtained by [26], [37], and [38]

$$\hat{\theta}_N := \text{sol}_{\theta \in \Theta} \left\{ \frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\varepsilon(t, \theta) = 0 \right\} \quad (13)$$

with $\Theta \subset \mathbb{R}^{n+1}$ an appropriate parameter space, $t_s > 0$, representing the starting sample used in the IV estimate, and $\tilde{N} = N - t_s + 1$. By algebraic manipulation the standard expression for a linear regression model estimate follows:

$$\hat{\theta}_N = \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\phi^T(t) \right]^{-1} \frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)y(t) \quad (14)$$

provided that the inverted matrix is nonsingular. This is achieved by appropriate choice of the instruments $z_k(t)$ in $\zeta(t)$ and thus by appropriate choice of r .

In open-loop experimental situations the input signal u will satisfy the requirements for a proper instrument, and thus the choice $r = u$ in that case leads to a simple linear LS estimate of the expansion coefficients. Under closed-loop experimental conditions, the instrument r can be chosen so as to match an external reference signal that is correlated with the input signal u , but independent of the noise process e_0 .

By adopting the additional notation $u^-(t) = u(t)$ for $t \leq 0$, and $u^-(t) = 0$ for $t > 0$, the system output $y(t)$ can be written as

$$y(t) = \tilde{G}_0(q)u^+(t) + \bar{G}_0(q)u^+(t) + G_0(q)u^-(t) + v_0(t) \quad (15)$$

with $\tilde{G}_0(q)u^+(t) = \phi^T(t)\theta_0$ and $\theta_0 = [g_0(0) \ \dots \ g_0(n)]$. Substituting these expressions in (14) shows that

$$\hat{\theta}_N = \theta_0 + \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\phi^T(t) \right]^{-1} \frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\gamma(t) \quad (16)$$

with

$$\gamma(t) = \bar{G}_0(q)u^+(t) + G_0(q)u^-(t) + v_0(t). \quad (17)$$

The second term in the right-hand side of (16) that is dependent on $\gamma(t)$ clearly reflects the three sources of modeling errors.

The first term in $\gamma(t)$ reflects the effect of the neglected tail of the series expansion on the estimated parameters and thus points to an effect of undermodeling; the second term characterizes the effect of unknown initial conditions, and the third term reflects the effect of noise disturbance.

IV. FREQUENCY RESPONSE UNCERTAINTY REGIONS

A. Introduction

In this section we will analyze the effect of the uncertainty in the estimated parameters on the frequency response of the estimated model.

Denote $\hat{\theta}_N = [\hat{g}(0) \cdots \hat{g}(n)]^T$. Then

$$\hat{G}(z) = \sum_{k=0}^n \hat{g}(k)F_k(z) \quad (18)$$

and with $\tilde{G}_0(z)$ as defined in (6) it follows from (16) that

$$\begin{aligned} \hat{G}(z) &= \tilde{G}_0(z) + [F_0(z) \cdots F_n(z)] \\ &\cdot \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\phi^T(t) \right]^{-1} \frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\gamma(t) \end{aligned}$$

and thus $\hat{G}(z) - G_0(z) = -\bar{G}_0(z) +$

$$[F_0(z) \cdots F_n(z)] \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\phi^T(t) \right]^{-1} \frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\gamma(t).$$

Next we evaluate this difference term for a specifically chosen frequency ω , leading to the expression

$$\hat{G}(e^{i\omega}) - G_0(e^{i\omega}) = -\bar{G}_0(e^{i\omega}) + \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_\omega(t)\gamma(t) \quad (19)$$

where for fixed ω , $\kappa_\omega(t)$ is a complex scalar-valued signal defined by

$$\begin{aligned} \kappa_\omega(t) &:= [F_0(e^{i\omega}) \cdots F_n(e^{i\omega})] \\ &\cdot \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t)\phi^T(t) \right]^{-1} \zeta(t). \quad (20) \end{aligned}$$

The signal $\kappa_\omega(t)$ is a filtered version of the signal $r(t)$ and can be computed, as it only depends on known quantities once the frequency ω has been specified. It will play an essential role throughout the following derivation of model error bounds. Combining (17) and (19) we can write:

$$\begin{aligned} \hat{G}(e^{i\omega}) - G_0(e^{i\omega}) &= -\bar{G}_0(e^{i\omega}) + \eta + \beta \\ &+ \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_\omega(t)v_0(t) \quad (21) \end{aligned}$$

with

$$\eta := \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_\omega(t)\bar{G}_0(q)u^+(t) \quad (22)$$

$$\beta := \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_\omega(t)G_0(q)u^-(t). \quad (23)$$

In the next subsection we will formulate upper bounds for the several terms in this expression.

B. Bounding the Model Error

Analyzing (21) we have to find upper bounds for the several terms on the right-hand side of the equation. Given the fact that the expressions are complex-valued, we will formulate separate bounds for Real and Imaginary parts. To this end we denote

$$\operatorname{Re}\{\kappa_\omega(t)\} =: \kappa_{R,\omega}(t) \quad (24)$$

$$\operatorname{Im}\{\kappa_\omega(t)\} =: \kappa_{I,\omega}(t). \quad (25)$$

When there is no confusion possible, we will drop the ω subscript, in order to simplify notation. The quantities η and β can simply be bounded by using the available data and prior information, as follows.

Lemma 4.1:

$$\begin{aligned} |\operatorname{Re}\{\eta\}| &\leq \bar{\eta}_R := \sum_{k=n+1}^{\infty} \bar{g}(k) \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_{R,\omega}(t)x_k(t) \right| \\ |\operatorname{Im}\{\eta\}| &\leq \bar{\eta}_I := \sum_{k=n+1}^{\infty} \bar{g}(k) \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_{I,\omega}(t)x_k(t) \right|. \end{aligned}$$

Proof: The expressions follow by substitution of (7) and (10) into (22) of η , and by utilizing (4). \square

Lemma 4.2: Denote the Laurent series expansion of $F_k(z)$ by $F_k(z) = \sum_{j=0}^{\infty} f_k(j)z^{-j}$, then

$$\begin{aligned} |\operatorname{Re}\{\beta\}| &\leq \bar{\beta}_R \\ &:= \sum_{k=0}^{\infty} \bar{g}(k) \sum_{\ell=0}^{\infty} \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_{R,\omega}(t)f_k(t+\ell) \right| \bar{u}, \\ |\operatorname{Im}\{\beta\}| &\leq \bar{\beta}_I \\ &:= \sum_{k=0}^{\infty} \bar{g}(k) \sum_{\ell=0}^{\infty} \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_{I,\omega}(t)f_k(t+\ell) \right| \bar{u}. \end{aligned}$$

Proof: Substituting (3) into (23) and utilizing the Laurent expansion of $F_k(z)$ shows that $\beta = 1/\tilde{N} \sum_{t=t_s}^N \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \kappa_\omega(t)g_0(k)f_k(j)u^-(t-j)$. By variable substitution $\ell = j - t$, the above expressions directly result. \square

Lemma 4.3:

$$\begin{aligned} |\operatorname{Re}\{\bar{G}_0(e^{i\omega})\}| &\leq \bar{\delta}_R := \sum_{k=n+1}^{\infty} \bar{g}(k) |\operatorname{Re}\{F_k(e^{i\omega})\}| \\ |\operatorname{Im}\{\bar{G}_0(e^{i\omega})\}| &\leq \bar{\delta}_I := \sum_{k=n+1}^{\infty} \bar{g}(k) |\operatorname{Im}\{F_k(e^{i\omega})\}|. \end{aligned}$$

Proof: The results follow directly by combining (4) and (7). \square

The above three Lemmas represent computable bounds for the contribution of the first three terms in (21). The actual computation of the expressions involves the evaluation of infinite sums. However, due to the fact that $\bar{g}(k)$ shows exponential decay rate in k and $f_k(t)$ shows exponential decay rate in t , the infinite sums will converge. In [13] it is shown

that all sums can be calculated to within arbitrary accuracy. This is done by truncating the infinite sums to a finite sum and providing a bound for the truncation error which will vanish with increasing length of the truncated sum. For further details we refer to [13].

Clearly $\bar{\beta}_R, \bar{\beta}_I$ will be small if t_s is chosen large; $\bar{\eta}_R, \bar{\eta}_I, \bar{\delta}_R, \bar{\delta}_I$ will be small if the effect of the neglected tail $\{g_0(k)\}_{k=n+1, \dots, \infty}$ is small. This tail effect will be small if n is chosen large and/or the basis is chosen appropriately so as to have a series expansion (3) with a high rate of convergence. This latter property can be achieved by choosing basis functions that are adapted to the system dynamics, as is possible with the generalized basis functions introduced in [19]. For the remaining term $1/\tilde{N} \sum_{t=t_s}^N \kappa_\omega(t)v_0(t)$, we establish the following key lemma which is in fact a multivariate extension of [16, Proposition 2.5].

Lemma 4.4: Let v_0 satisfy Assumptions 2.1 and be independent of the bounded deterministic quasi-stationary sequence $\{r(t)\}$. Let r_1 and r_2 be two signals determined by $r_1(t) = F_1(q)r(t)$, $r_2(t) = F_2(q)r(t)$ for two ℓ_∞ -stable linear filters F_1 and F_2 . Denote

$$\Lambda_{r_1 r_2}^N := E \frac{1}{\tilde{N}} \begin{bmatrix} \sum_{t=t_s}^N r_1(t)v_0(t) \\ \sum_{t=t_s}^N r_2(t)v_0(t) \end{bmatrix} \cdot \begin{bmatrix} \sum_{t=t_s}^N r_1(t)v_0(t) & \sum_{t=t_s}^N r_2(t)v_0(t) \end{bmatrix}$$

and $\Lambda_{r_1 r_2} := \lim_{N \rightarrow \infty} \Lambda_{r_1 r_2}^N$. Also, denote for $i, j = 1, 2$

$$R_{r_i r_j}^N(\tau) := \frac{1}{\tilde{N} + \tau} \sum_{t=t_s}^{N+\tau} r_i(t)r_j(t-\tau),$$

$$\tau = -N + t_s, \dots, 0,$$

$$R_{r_i r_j}^N(\tau) := \frac{1}{\tilde{N} - \tau} \sum_{t=t_s}^{N-\tau} r_i(t+\tau)r_j(t),$$

$$\tau = 1, \dots, N - t_s$$

and $R_{r_i}^N(\tau) := R_{r_i r_i}^N(\tau)$, $i = 1, 2$. Then

$$1) \quad \Lambda_{r_1 r_2}^N(i, j) = \sum_{\tau=-N+t_s}^{N-t_s} \frac{\tilde{N} - |\tau|}{\tilde{N}} R_{r_i r_j}^N(\tau) R_{v_0}(\tau),$$

$$i, j = 1, 2$$

$$2) \quad \Lambda_{r_1 r_2}(i, j) = \sum_{\tau=-\infty}^{\infty} R_{r_i r_j}(\tau) R_{v_0}(\tau),$$

$$i, j = 1, 2$$

$$3) \quad \frac{1}{\sqrt{\tilde{N}}} \begin{bmatrix} \sum_{t=t_s}^N r_1(t)v_0(t) \\ \sum_{t=t_s}^N r_2(t)v_0(t) \end{bmatrix} \xrightarrow{N \rightarrow \infty} \mathcal{N}(0, \Lambda_{r_1 r_2})$$

where $\mathcal{N}(0, \Lambda_{r_1 r_2})$ denotes the Multivariate Normal distribution with mean zero and covariance matrix $\Lambda_{r_1 r_2}$. Moreover, if $\Lambda_{r_1 r_2}^N$ is invertible

$$4) \quad \frac{1}{\tilde{N}} \begin{bmatrix} \sum_{t=t_s}^N r_1(t)v_0(t) & \sum_{t=t_s}^N r_2(t)v_0(t) \end{bmatrix} (\Lambda_{r_1 r_2}^N)^{-1} \cdot \begin{bmatrix} \sum_{t=t_s}^N r_1(t)v_0(t) \\ \sum_{t=t_s}^N r_2(t)v_0(t) \end{bmatrix} \xrightarrow{N \rightarrow \infty} \chi^2(2)$$

where $\chi^2(2)$ denotes the Chi-square distribution with 2 degrees of freedom.

Proof: The proof is given in [13] and [16]. In the proof fruitful use is made of results established in [26, Th. 9.1] and [20, ch. 2]. Existence of the limit $\lim_{N \rightarrow \infty} \Lambda_{r_1 r_2}^N$ is guaranteed by the fact that r is a bounded deterministic quasi-stationary sequence, and v_0 satisfies Assumption 2.1. \square

The result of this Lemma can be applied to our error term $1/\tilde{N} \sum_{t=t_s}^N \kappa_\omega(t)v_0(t)$ by realizing that $\kappa_R(t)$ and $\kappa_I(t)$ satisfy the conditions that are formulated for r_1 and r_2 in this Lemma. In this sense, the Lemma establishes an asymptotic result for the probability density function of the stochastic term in our model error. Use of a central limit theorem provides an asymptotic normal distribution. Consequently, the results given in 3) and 4) are asymptotic results. For finite N the given distributions are approximations of the true ones. In [13, Appendix 5.B] extensive Monte Carlo simulations illustrate the relevance of this approximation also for finite N already. Note that the expression for the covariance matrix in Part 1) is a nonasymptotic result; it is correct for any N .

The asymptotic covariance matrix $\Lambda_{r_1 r_2}$ in the Lemma is dependent on the covariance function $R_{v_0}(\tau)$ of the unknown noise disturbance process. For the moment we will assume that it is known. In Section VII, it will be shown that we can replace this covariance function by an estimate without conflicting with the original result stated above.

C. Frequency Response Uncertainty Regions

Using the results of the previous subsection, a computable bound for the model error $\hat{G}(e^{i\omega}) - G_0(e^{i\omega})$ is straightforwardly obtained in the form of a confidence region for the system's frequency response $G_0(e^{i\omega})$. The bound is given in the following main theorem.

Theorem 4.5: Consider the IV estimate (14) with corresponding frequency response $\hat{G}(e^{i\omega})$. Let v_0 satisfy Assumptions 2.1 and be independent of the bounded deterministic quasi-stationary sequence $\{r(t)\}$. Let $\bar{\eta}_R, \bar{\eta}_I, \bar{\beta}_R, \bar{\beta}_I, \bar{\delta}_R,$ and $\bar{\delta}_I$ be given by Lemmas 4.1–4.3. Let $\kappa_R(t), \kappa_I(t)$ be as defined in (24) and (25), for a fixed choice $\omega = \omega_j \in [0, \pi]$, and denote

$$\Lambda_{\kappa_R \kappa_I}^N = \begin{bmatrix} \lambda_{\kappa_R}^N & \lambda_{\kappa_R \kappa_I}^N \\ \lambda_{\kappa_I \kappa_R}^N & \lambda_{\kappa_I}^N \end{bmatrix} \quad (26)$$

conformable to Part 1) of Lemma 4.4. Then, asymptotically¹ in N

$$\begin{aligned} 1) \quad & |\operatorname{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j}))| \\ & \leq c_{N,\alpha} \sqrt{\frac{\lambda_{\kappa_R}^N}{\tilde{N}}} + \bar{\eta}_R + \bar{\beta}_R + \bar{\delta}_R, \text{ w.p. } \geq \alpha \\ 2) \quad & |\operatorname{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j}))| \\ & \leq c_{N,\alpha} \sqrt{\frac{\lambda_{\kappa_I}^N}{\tilde{N}}} + \bar{\eta}_I + \bar{\beta}_I + \bar{\delta}_I, \text{ w.p. } \geq \alpha \end{aligned}$$

where $c_{N,\alpha}$ corresponds to a probability α in the standard Normal distribution, such that $x \in \mathcal{N}(0, 1)$ implies $\operatorname{prob}(|x| \leq c_{N,\alpha}) = \alpha$.

Moreover, if $\Lambda_{\kappa_R \kappa_I}^N$ is invertible, introduce $\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$ as the square-root of the inverse of $\Lambda_{\kappa_R \kappa_I}^N$, i.e., $\Gamma^T \Gamma = (\Lambda_{\kappa_R \kappa_I}^N)^{-1}$ and $\Gamma^T = \Gamma$. Then, asymptotically in N

$$\begin{aligned} 3) \quad & \begin{bmatrix} \operatorname{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \\ \operatorname{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \end{bmatrix}^T \Gamma^T \Gamma \\ & \cdot \begin{bmatrix} \operatorname{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \\ \operatorname{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \end{bmatrix} \\ & \leq \left(\sqrt{\frac{c_{\chi,\alpha}}{\tilde{N}}} + \sqrt{\gamma_{11}^2 + \gamma_{21}^2}(\bar{\eta}_R + \bar{\beta}_R + \bar{\delta}_R) \right. \\ & \left. + \sqrt{\gamma_{12}^2 + \gamma_{22}^2}(\bar{\eta}_I + \bar{\beta}_I + \bar{\delta}_I) \right)^2, \quad \text{w.p. } \geq \alpha \end{aligned}$$

where $c_{\chi,\alpha}$ corresponds to a probability α in the Chi-square distribution with 2 degrees of freedom, such that $x \in \chi^2(2)$ implies $\operatorname{prob}(x \leq c_{\chi,\alpha}) = \alpha$.

Proof: See the Appendix. \square

Parts 1) and 2) of this theorem provide probabilistic bounds for the real and imaginary parts of the model error for one specific choice of ω_j , and as such pointwise for the frequency response of the system $G_0(z)$. The uncertainty region for $G_0(e^{i\omega_j})$ can be constructed as rectangular system confidence regions in the complex plane using Bonferroni's inequality [29]. In particular, if any complex-valued random variable x has the property that $\operatorname{Re}(x) \leq a_1$, w.p. $\geq \alpha_1$, and $\operatorname{Im}(x) \leq a_2$, w.p. $\geq \alpha_2$, then $\{\operatorname{Re}(x) \leq a_1 \wedge \operatorname{Im}(x) \leq a_2\}$, w.p. $\geq 1 - (1 - \alpha_1) - (1 - \alpha_2)$.

Ellipsoidal system confidence regions are obtained with Part 3) of the above theorem, provided the matrix $\Lambda_{\kappa_R \kappa_I}^N$ is invertible. Note that this is generally the case, except for frequencies $\omega_j = 0, \pi$. For these frequencies the signal $\{\kappa_I(t)\}$ is identically zero, as $\operatorname{Im}(F_k(e^{i\omega_j}))$ appearing in (20) is zero. This very naturally means that for frequencies zero and π there is no imaginary system uncertainty.

As before, the several sources of uncertainty can clearly be distinguished; the λ -dependent terms reflect a variance effect, being caused by the noise term v_0 . The terms with $\bar{\eta}_R, \bar{\eta}_I$

are due to the effect of the neglected tail on the accuracy of the estimated coefficients of $\tilde{G}_0(q)$ and represent a bias contribution. The third contribution with $\bar{\beta}_R, \bar{\beta}_I$ is due to the unknown initial conditions in the data, i.e., the effect of data outside the interval that is measured. Finally, the contribution of $\bar{\delta}_R, \bar{\delta}_I$ corresponds to the neglected dynamics of the tail $\tilde{G}_0(q)$ itself and also represents a bias contribution.

The different error sources can be a tradeoff. In particular the truncation value n can be used to make a tradeoff between bias and variance. A larger value n means a smaller bias, but a larger variance. An optimal value can be determined by varying n so as to minimize the size of the uncertainty. Similarly, the integer t_s offers the possibility to tradeoff the influence of initial conditions to the variance. A larger value t_s means a decrease of the error contributions $\bar{\beta}_R, \bar{\beta}_I$, but an increase of the variance, due to a decreasing $\tilde{N} = N - t_s + 1$.

It is emphasized that the identification of the IV model is not a goal as such, but serves only as a basis for the construction of system uncertainty regions. The design variables in the IV identification, such as the IV model order n , should not be used to obtain a tractable (low-order) nominal model, but should be tuned in such a way that the uncertainty regions are as small as possible. The identification of an appropriate (low-order) nominal model, suited for use in control design, is not the issue here.

The result of Theorem 4.5 is that for any user-chosen frequency ω_j an uncertainty region can be determined that is valid with prescribed probability. Typically an uncertainty region will be determined for a user-chosen frequency-grid, requiring a recalculation of all the terms that determine the uncertainty region for each separate ω . This mechanism of recalculation for each separate ω requires considerable computational effort. However, with contemporary computing power this is not a severe problem. On the other hand, by replacing the usual closed form expressions by these computational solutions, a considerable reduction in conservatism is achieved.

Remark 4.6: In literature variance expressions are given for IV and FIR estimates, assuming that the system is in the model set and neglecting the influence of the initial conditions; see e.g., [26], [37], and [38]. The first condition implies that the available results are asymptotic in the model order n also. Here we clearly have presented results that are valid for finite n . Some progress has also been made in [21] and [22], where for a different identification setting a procedure is presented to incorporate the influence of the bias when computing the variance. Another important aspect is that variance expressions found in literature are often derived for the asymptotic case $N \rightarrow \infty$, whereas here tractable nonasymptotic expressions have been derived. In particular, the expression for the covariance matrix, as given in Part 1) of Lemma 4.4, is correct in the nonasymptotic case. Only the distributions in Theorem 4.5 are asymptotic.

As Theorem 4.5 leads to the construction of system uncertainty regions for frequencies in a user-chosen frequency-grid, it does not provide statements about the system uncertainty for frequencies outside the grid. However, the system's prior information (4) concerns a smoothness condition on the frequency response, leading to the following Lemma.

¹The probabilistic expressions are based on the asymptotic distribution of the corresponding stochastic variable, and so for finite N they will not be exact, but approximative.

Lemma 4.7: Let $G_0 \in \mathcal{A}$ be finite-dimensional and satisfy (4) and (5). Then

$$\left| \frac{d G_0(e^{i\omega})}{d\omega} \right| \leq \sum_{k=0}^{\infty} \bar{g}(k) \sup_{\omega} \left| \frac{d F_k(e^{i\omega})}{d\omega} \right|$$

for all $\omega \in [0, \pi]$.

Proof: For $G_0 \in \mathcal{A}$ and finite-dimensional, G_0 will be differentiable on the unit circle, and so

$$\left| \frac{d G_0(e^{i\omega})}{d\omega} \right| = \left| \sum_{k=0}^{\infty} g_0(k) \frac{d F_k(e^{i\omega})}{d\omega} \right|$$

$$\leq \sum_{k=0}^{\infty} \bar{g}(k) \left| \frac{d F_k(e^{i\omega})}{d\omega} \right|.$$

□

As the basis functions are given *a priori*, the obtained upper bound expression can be computed, providing upper bounds on the system uncertainty for frequencies that are not in the chosen frequency-grid as meant before.

V. PULSE RESPONSE UNCERTAINTY REGIONS

In this section an analysis is made of the uncertainty region for the estimated pulse response, leading to a pulse response uncertainty region for the system $G_0(z)$.

Additional to the notation $G_0(z) = \sum_{k=0}^{\infty} p_0(k)z^{-k}$, we introduce $\tilde{G}_0(z) = \sum_{k=0}^{\infty} \tilde{p}_0(k)z^{-k}$ and $\hat{G}_0(z) = \sum_{k=0}^{\infty} \hat{p}(k)z^{-k}$.

Expanding the basis functions $F_k(z)$ according to $F_k(z) = \sum_{\ell=0}^{\infty} f_k(\ell)z^{-\ell}$, and using expressions (3), (6), (7), and (18), it follows that

$$\hat{p}(k) = \sum_{j=0}^n \hat{g}(j) f_j(k)$$

$$\tilde{p}_0(k) = \sum_{j=0}^n g_0(j) f_j(k).$$

Now, multiplication of the left- and right-hand side of (16) with the row vector $[f_0(k) \ f_1(k) \ \cdots \ f_n(k)]$ leads to

$$\hat{p}(k) = \tilde{p}_0(k) + \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_k(t) \gamma(t) \quad (27)$$

with

$$\kappa_k(t) := [f_0(k) \ \cdots \ f_n(k)] \left[\frac{1}{\tilde{N}} \sum_{t=t_s}^N \zeta(t) \phi^T(t) \right]^{-1} \zeta(t). \quad (28)$$

Note that for each $k = 0, 1, \dots, n$, $\kappa_k(t)$ is a scalar real-valued signal that replaces the complex valued signal $\kappa_{\omega}(t)$ in the analysis of the previous section. Similarly, the signals $\kappa_k(t)$ are computable on the basis of prior knowledge and measurement data.

Expression (27) now essentially matches the similar expression (19), and so the analysis and bounding of the model error follows along very similar lines. This leads to the following theorem, which provides probabilistic error bounds for the parameters $\hat{p}(k)$, and as such confidence intervals for the system parameters $p_0(k)$.

Theorem 5.1: Consider the IV estimate (14) with corresponding pulse response $\{\hat{p}(k)\}$. Let v_0 satisfy Assumptions 2.1 and be independent of the bounded deterministic quasi-stationary sequence $\{r(t)\}$. Let $c_{\mathcal{N}, \alpha}$ correspond to a probability α in the standard Normal distribution, such that, $x \in \mathcal{N}(0, 1)$ implies $\text{prob}(|x| \leq c_{\mathcal{N}, \alpha}) = \alpha$. Denote

$$\lambda_{\kappa_k}^N := \sum_{\tau=-N+t_s}^{N-t_s} \frac{\tilde{N} - |\tau|}{\tilde{N}} R_{\kappa_k}^N(\tau) R_{v_0}(\tau)$$

with $\kappa_k(t)$ given by (28). Then, for $N \rightarrow \infty$

$$|\hat{p}(k) - p_0(k)| \leq c_{\mathcal{N}, \alpha} \sqrt{\frac{\lambda_{\kappa_k}^N}{\tilde{N}}} + \bar{\eta}(k) + \bar{\beta}(k) + \bar{\delta}(k) \text{ w.p. } \geq \alpha$$

where

$$\bar{\eta}(k) = \sum_{j=n+1}^{\infty} \bar{g}(j) \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_k(t) x_j(t) \right| \quad (29)$$

$$\bar{\beta}(k) = \sum_{j=0}^{\infty} \bar{g}(j) \sum_{\ell=0}^{\infty} \left| \frac{1}{\tilde{N}} \sum_{t=t_s}^N \kappa_k(t) f_j(t+\ell) \right| \bar{u} \quad (30)$$

$$\bar{\delta}(k) = \sum_{j=n+1}^{\infty} \bar{g}(j) |f_j(k)|. \quad (31)$$

Proof: The proof is similar to the proof of Part 1) of Theorem 4.5, replacing $\kappa_{\omega}(t)$ by $\kappa_k(t)$, and using the fact that $p_0(k) = \tilde{p}_0(k) + \bar{p}_0(k)$ with $\tilde{G}_0(z) = \sum_{k=0}^{\infty} \tilde{p}_0(k)z^{-k}$. □

The confidence intervals can be calculated for $k = 0, \dots, m$, for some user-defined integer m , leading to parameter intervals

$$p_l(k) \leq p_0(k) \leq p_u(k), \quad \text{w.p. } \geq \alpha.$$

An upper bound can be specified for the remaining parameters, corresponding to $k = m+1, \dots, \infty$, using the prior bounds in (4)

$$|p_0(k)| \leq \sum_{j=0}^{\infty} \bar{g}(j) |\kappa_j(k)|, \quad k = m+1, \dots, \infty.$$

Note that again, as in the previous frequency response analysis, the bounds refer to computable expressions of known quantities, however, containing infinite sums. They can be calculated to within each desired accuracy.

Remark 5.2: An analysis similar to the one presented here for the pulse response coefficients of G_0 can be given for the

pulse response coefficients of WG_0 with W any stable and stably invertible weighting function; see [13].

VI. STEP RESPONSE UNCERTAINTY REGIONS

In many industrial application situations where estimated models and their uncertainty region are of interest, a step response uncertainty region will be of higher relevance than a pulse response uncertainty region.

It appears that the analysis given above can straightforwardly be extended to cover the situation of step response coefficients.

By denoting $s_0(k) := \sum_{j=0}^k p_0(j)$, $\tilde{s}_0(k) := \sum_{j=0}^k \tilde{p}_0(j)$, and $\hat{s}(k) = \sum_{j=0}^k \hat{p}(j)$, we can repeat the analysis as in the situation of pulse responses, however, now by multiplication of the left- and right-hand side of (16) with the row vector $\sum_{j=0}^k [f_0(j) \ f_1(j) \ \cdots \ f_n(j)]$. This leads to

$$\hat{s}(k) = \tilde{s}_0(k) + \frac{1}{N} \sum_{t=t_s}^N \kappa_k(t) \gamma(t)$$

with

$$\kappa_k(t) := \sum_{j=0}^k [f_0(j) \ \cdots \ f_n(j)] \left[\frac{1}{N} \sum_{t=t_s}^N \zeta(t) \phi^T(t) \right]^{-1} \zeta(t),$$

As a result the full analysis of the pulse response situation goes through, however, now with the row vector $[f_0(k) \ \cdots \ f_n(k)]$ in $\kappa_k(t)$ replaced by the row vector $\sum_{j=0}^k [f_0(j) \ \cdots \ f_n(j)]$.

VII. ESTIMATION OF THE PRIOR INFORMATION FROM DATA

A. Introduction

The uncertainty intervals derived in the previous sections can be calculated, provided some prior information on the system and that its disturbance signals is available. This specifically concerns knowledge of the autocovariance function of the noise, $\{R_{v_0}(\tau)\}$, as present in the expression for the matrix $\Lambda_{r_1 r_2}^N$ occurring in Lemma 4.4-1). In specific forms this matrix is employed in the formulation of Theorems 4.5 and 5.1.

A second source of prior information that is used in our formulated error bounds is the exponential bound on the decay rate of the expansion coefficients $\bar{g}(k)$. We will first show how we can estimate the noise covariance information from the data, and next we will discuss a mechanism how to adjust the coefficient bounds in accordance with the data.

B. Estimation of Noise Auto-Covariance Function

The next Proposition shows that we can safely replace the prior information on the noise process in the results of the several Theorems by an appropriate estimate.

Proposition 7.1:

- 1) If in Theorem 4.5 the unknown matrix $\Lambda_{\kappa_R \kappa_I}^N$ is replaced by an estimate $\hat{\Lambda}_{\kappa_R \kappa_I}^N$ satisfying $\hat{\Lambda}_{\kappa_R \kappa_I}^N - \Lambda_{\kappa_R \kappa_I}^N \geq 0$, then the results 1)–3) of this Theorem still hold true.
- 2) If in Theorem 5.1 the unknown variable $\lambda_{\kappa_k}^N$ is replaced by an estimate $\hat{\lambda}_{\kappa_k}^N$, satisfying $\hat{\lambda}_{\kappa_k}^N - \lambda_{\kappa_k}^N \geq 0$, then the results of this Theorem still hold true. \square

Proof: It follows directly that an overestimate substituted in Parts 1) and 2) of Theorem 4.5 will lead to right-hand sides in the corresponding expressions that are larger than necessary. The expressions thus hold true, but the resulting uncertainty regions (for a given probability) get larger and thus more conservative. This same holds for Part 3) of Theorem 4.5, where the ellipsoid on the left-hand side of the expression will get smaller.

Part 2) of the Proposition follows similarly as Parts 1) and 2) of Theorem 4.5. \square

The Proposition shows that we may use an “over”-estimate of the corresponding matrices that contain the unknown noise autocovariance function. The result is that the obtained uncertainty regions are somewhat bigger, introducing conservatism. However, the probabilistic expressions in the Theorems remain correct.

In line with the expressions used in Lemma 4.4, we will now show how we can construct, for two general filtered reference signals r_1 and r_2 , an estimate $\hat{\Lambda}_{r_1 r_2}^N$ that asymptotically satisfies the conditions as mentioned in the above Proposition and thus overestimates the exact expression $\Lambda_{r_1 r_2}$ as defined in Lemma 4.4-2). As this situation also covers the estimate of the scalar entry as meant on Part 2) of the Proposition above, we will only consider this matrix case. The procedure that we will sketch is similar in spirit to a related procedure that is introduced and analyzed for a different problem in [16].

To this end we consider having available an ℓ_∞ -stable nominal model $\hat{G}_N(q)$ that has been obtained independently of the considered data set, for example by identification based on another data set. We consider the output error

$$\hat{e}(t) := y(t) - \hat{G}_N(q)u(t) = \psi(t) + v_0(t) \quad (32)$$

with $\psi(t) := (G_0(q) - \hat{G}_N(q))u(t)$, and the idea is to use this output error for estimating the second-order statistics of the noise process v_0 . Next denote

$$\hat{R}_{\hat{e}}^N(\tau) := \frac{1}{\tilde{N} - |\tau|} \sum_{t=t_s}^{N-|\tau|} \hat{e}(t) \hat{e}(t + |\tau|) \quad (33)$$

and consider the following estimate for $\Lambda_{r_1 r_2}^N$:

$$\hat{\Lambda}_{r_1 r_2}^N = \sum_{\tau=-w(N)}^{w(N)} c_w(\tau) \frac{\tilde{N} - |\tau|}{N} \hat{R}_{\hat{e}}^N(\tau) \cdot \begin{bmatrix} R_{r_1}^N(\tau) & R_{r_1 r_2}^N(\tau) \\ R_{r_1 r_2}^N(\tau) & R_{r_2}^N(\tau) \end{bmatrix} \quad (34)$$

where $c_w(\tau)$ is a (positive real-valued) window function, similar to the ones used in spectral analysis [26]. Now the following Theorem can be established for the situation that we are dealing with open loop measurements.

Theorem 7.2: Consider the situation of Lemma 4.4, and consider the expressions $\Lambda_{r_1 r_2}^N$ and its estimate $\hat{\Lambda}_{r_1 r_2}^N$ as defined in (33) and (34), where the window function $c_w(\tau)$ satisfies some technical assumptions (see proof), and the model $\hat{G}_N(q)$ has been obtained independent of the noise process $\{v_0(t)\}_{t=t_s, \dots, N}$.

Suppose that the input signal $\{u(t)\}$ is independent of the noise process e_0 and has an autocovariance function $R_u(\tau)$

that satisfies $|R_u(\tau)| \leq M_u \rho_u^\tau$ for some finite $M_u, \rho_u \in \mathbb{R}$, $\rho_u < 1$. Then under weak conditions on the input signal (see proof)

- 1) $\lim_{N \rightarrow \infty} \hat{\Lambda}_{r_1 r_2}^N = \Lambda_{r_1 r_2} + \Lambda_\Psi$ w.p. 1 with $\Lambda_\Psi =$

$$\begin{bmatrix} \sum_{\tau=-\infty}^{\infty} R_{r_1}(\tau)R_\psi(\tau) & \sum_{\tau=-\infty}^{\infty} R_{r_1 r_2}(\tau)R_\psi(\tau) \\ \sum_{\tau=-\infty}^{\infty} R_{r_1 r_2}(\tau)R_\psi(\tau) & \sum_{\tau=-\infty}^{\infty} R_{r_2}(\tau)R_\psi(\tau) \end{bmatrix}$$
- 2) $\Lambda_\Psi \geq 0$.

Proof: See the Appendix. \square

The consequence of this result is that the given estimate $\hat{\Lambda}_{r_1 r_2}^N$ can be used to replace the true value $\Lambda_{r_1 r_2}^N$ in Theorem 4.5 in order to determine frequency response uncertainty regions. Asymptotically correct results are obtained with respect to both the ellipsoidal and rectangular confidence regions. The result shows that classical spectral estimation techniques can fruitfully be applied to estimate the variance of an IV or FIR LS estimate. This has also been shown in [20].

In the Theorem it is assumed that the input signal has an exponentially decaying auto-covariance function. This is necessary in order for the auto-covariance function of the output error $\hat{e}(t)$ to be exponentially decaying. Consequently, the input u is not allowed to contain undecaying deterministic components such as sinusoids. If they are present in the input, they should be detected and removed from both input and output signals.

C. Estimation of Parameter Bounds

In this subsection it is indicated how data may be used to derive an appropriate prior bound $\bar{g}(k)$ on the expansion coefficients of the system to be modeled. Analogously to the derivation of pulse response uncertainty regions in Section V, it is possible to derive parameter confidence intervals. It follows directly from (16) that for $k \leq n$

$$\hat{g}(k) = \hat{g}_0(k) + \frac{1}{N} \sum_{t=t_s}^N \kappa_k(t) \gamma(t)$$

with

$$\kappa_k(t) := e_k^T \left[\frac{1}{N} \sum_{t=t_s}^N \zeta(t) \phi^T(t) \right]^{-1} \zeta(t) \quad (35)$$

where e_k is the k th Euclidean unit vector. Note that in comparison with (28), $\kappa_k(t)$ is now obtained by replacing the row vector $[f_0(k) \ f_1(k) \ \cdots \ f_n(k)]$ in (28) by the unit vector e_k^T . Again, the analysis presented before fully applies, as formulated in the following Corollary.

Corollary 7.3: Consider the situation of Theorem 5.1. Denote $\lambda_{\kappa_k}^N$ as in Theorem 5.1 with $\kappa_k(t)$ as defined in (35). Then, for $N \rightarrow \infty$

$$|\hat{g}(k) - g_0(k)| \leq c_{N,\alpha} \sqrt{\frac{\lambda_{\kappa_k}^N}{N}} + \bar{\eta}(k) + \bar{\beta}(k) \text{ w.p. } \geq \alpha$$

where $\bar{\eta}(k)$ and $\bar{\beta}(k)$ are defined in (29) and (30). \square

Proof: The proof is the same as the proof of Theorem 5.1. Note that in this case $\bar{\delta}(k) = 0$ because of the fact that $\hat{g}_0(k) = g_0(k)$, $0 \leq k \leq n$. \square

The procedure of the previous subsection can be used to estimate the variance $\lambda_{\kappa_k}^N$, in order to actually compute the parameter confidence intervals.

The corollary establishes parameter confidence intervals, although for the computation of these intervals a prior bound $\bar{g}(k)$ has to be chosen. This result seems to end up in a vicious circle. However, it is still useful if properly combined with an iterative procedure. The posterior uncertainty interval, being the result of the above Corollary, is a means to validate the choice of the prior uncertainty interval induced by $\bar{g}(k)$. As a tighter choice of $\bar{g}(k)$ will lead to smaller uncertainty intervals, we will iteratively reduce the prior bound so as to achieve posterior uncertainty intervals that are not invalidated by the prior bounds for a specifically chosen interval $[0, n^*]$. In [4] it is suggested to choose n^* such that for $n > n^*$ the value zero is contained in the posterior uncertainty interval for the corresponding parameters. Although there is no formal guarantee that correct results are achieved, this procedure has shown to yield good results in practice.

VIII. SIMULATION EXAMPLE

In this section the procedure for probabilistic uncertainty bounding identification is illustrated by means of an example. Consider the data generating system

$$G_0(z) = \frac{0.21z^4 + 0.35z^3 - 0.12z^2 - 0.11z + 0.23}{z^5 - 2.5z^4 + 3.3z^3 - 2.5z^2 + 1.2z - 0.3}$$

The output of the system is disturbed by low-pass noise

$$y(t) = G_0(q)u(t) + v_0(t), \quad v_0(t) = \frac{0.2q}{q - 0.5} e_0(t)$$

where $e_0(t)$ is Gaussian white noise with variance one.

First, 500 samples of the output have been measured while the input is zero. This is a so-called free-run experiment and is used to estimate $R_{v_0}(\tau)$ according to (33). This estimate is used later in (34) for the construction of an overbounded estimate $\hat{\Lambda}_{r_1 r_2}$ of $\Lambda_{r_1 r_2}$ to be used in the uncertainty bounding.

For the model parameterization, basis functions have been chosen to be generated by a third-order all-pass function with pole locations 0.4183, and $0.6858 \pm 0.697i$ (see [19]). The generalized pulse response parameter bounds (5) are *a priori* chosen according to $M = 10$ and $\rho = 0.9$. Later on these bounds are tightened using insight obtained from the data.

Next an identification experiment is performed, where $N = 1000$ samples of the output have been measured, while the input has been excited with a Random Binary Sequence, i.e., a signal with amplitude switching between -1 and 1 at random time instants. By design the initial conditions are zero, so $\bar{u} = 0$ and t_s is chosen equal to one. As the data are generated in open loop, the signal $\{r(t)\}$ is chosen to be equal to $\{u(t)\}$. For the variance estimates of (34) use is made of a Tukey-window [2] with window parameters $f = 0.6$ and $w = 100$.

With Corollary 7.3 stochastic bounds for $g_0(k)$ have been identified, with $n = 45$ and $c_{N,0.99} = 2.58$. The results are

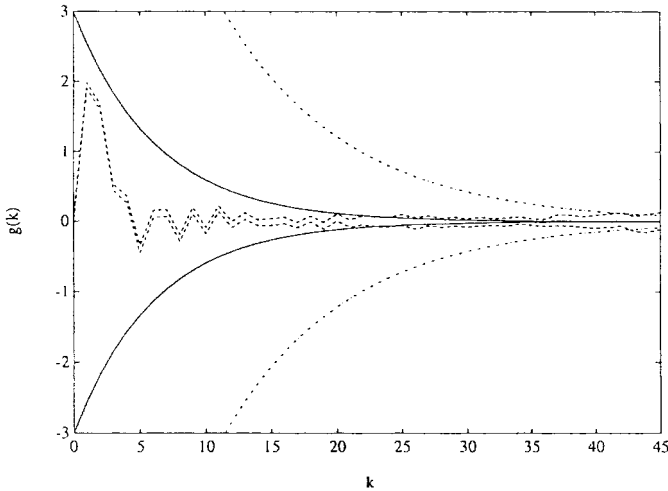


Fig. 1. Prior parameter bounds $\pm\bar{g}(k)$ (dash-dotted), identified bounds $g_l(k)$, $g_u(k)$ (dashed), and new prior bounds $\pm\bar{g}(k)$ (solid).

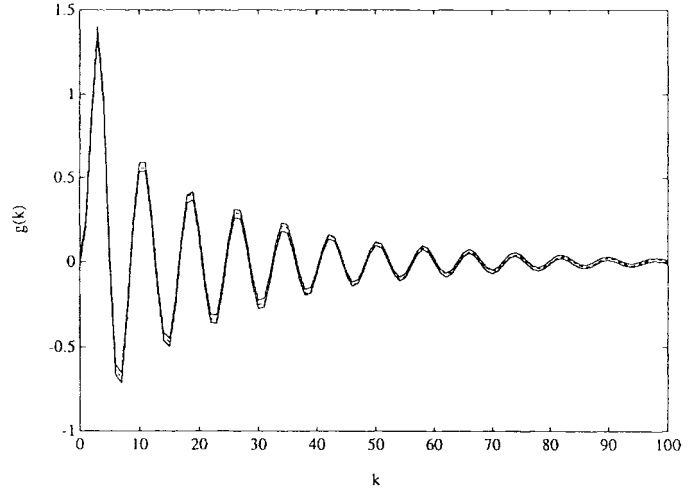


Fig. 3. Identified 99.9% confidence regions $[p_l(k), p_u(k)]$ for the system's pulse response (solid) and system's pulse response $p_0(k)$ (dashed).

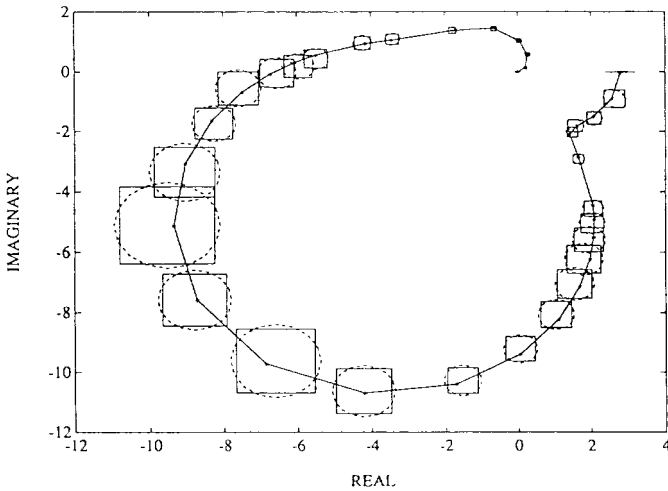


Fig. 2. Nyquist diagram identified 99.9%-confidence regions for the system's frequency response (rectangles, ellipsoids) and system's frequency response $G_0(e^{i\omega_j})$ (solid, *).

depicted in Fig. 1. The bounds $\pm\bar{g}(k)$ are shown, as well as the calculated confidence intervals.

Following the procedure proposed in Section VII-C, new and tighter prior bounds $\bar{g}(k)$ have been chosen such that they are consistent with the confidence intervals for the first 24 parameters. The new prior bounds are smaller than the confidence intervals of $g_0(k)$ for $k \geq 24$, where zero is in the confidence intervals. The resulting new bounds, corresponding to $M = 3$ and $\rho = 0.85$, are also depicted in Fig. 1. In fact, they appear to be satisfied by the system $G_0(z)$.

With the tighter chosen prior bounds $\bar{g}(k)$ and the covariance estimate $\hat{R}_\epsilon^N(\tau)$ obtained from the free-run experiment, probabilistic frequency response uncertainty regions are now identified on the basis of Theorem 4.5. Uncertainty regions are calculated for 32 frequencies between zero and π , while $n = 39$ parameters are identified. Both ellipsoidal confidence regions, corresponding to $c_{\chi, 0.999} = 13.815$, and rectangular confidence regions, corresponding to $c_{\mathcal{N}, 0.9995} = 3.5$, are calculated. Consequently, for each frequency both

the ellipsoidal and the rectangular confidence regions are correct, with probability larger than 0.999 (for the rectangles this follows by application of Bonferroni's inequality). The probability that the confidence regions are correct for all 32 frequencies uniformly is larger than $1 - 32 \cdot 0.001 = 0.97$. The results are depicted in Fig. 2, together with the system's frequency response $G_0(e^{i\omega_j})$. Moreover, with the procedure of Section V, 3.5σ -confidence regions are calculated for the first 100 pulse response parameters of the system $G_0(z)$. These pulse response uncertainty regions are depicted in Fig. 3, together with the system's pulse response. It appears that in this case the confidence regions are correct, i.e., the system is within the identified bounds, both in Figs. 2 and 3.

IX. CONCLUSIONS

In this paper an identification procedure has been developed which yields confidence regions for the frequency response, pulse response, and step response of some stable linear time-invariant system. The procedure involves the explicit calculation of bias and variance errors of an instrumental variable/least squares estimate. Relying on a computational solution, the procedure does not involve nonlinear optimization but restricts to convex optimization algorithms. Probabilistic uncertainty regions are obtained, while the separate sources of uncertainty (undermodeling, noise and initial conditions) can clearly be distinguished so as to indicate appropriate action for reduction of the uncertainty regions.

It is shown that the required prior information can be reliably estimated from the data itself. The procedure is applicable to closed-loop experiments, and extensions to the multivariable situation are available [13]. The use of generalized (orthonormal) basis functions enables us to incorporate prior knowledge about the system dynamics into the basis, thus leading to models with reduced bias and having a limited number of parameters.

The procedure makes use of results which are asymptotic in the number of data. As in applications there are always finite-data records, but the results might not be valid in

practice. On the other hand, Monte Carlo simulations [13] have shown that the error caused by the finiteness of the number of data can be very small, even for small values of N . A successful application of the presented identification method for identification and robust control of an industrial process has been reported in [3].

APPENDIX

A. Proof of Theorem 4.5

Combining (21) with Lemma 4.4 yields

$$\begin{aligned} & \sqrt{\tilde{N}} \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_R - \beta_R \\ \text{Im}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_I - \beta_I \end{bmatrix} \\ &= \frac{1}{\sqrt{\tilde{N}}} \begin{bmatrix} \sum_{t=t_s}^N \kappa_R(t) v_0(t) \\ \sum_{t=t_s}^N \kappa_I(t) v_0(t) \end{bmatrix} \xrightarrow{N \rightarrow \infty} \mathcal{N}(0, \Lambda_{\kappa_R \kappa_I}) \end{aligned}$$

as the conditions of Lemma 4.4 are satisfied. In particular $\{\kappa_R(t)\}$ and $\{\kappa_I(t)\}$ are filtered versions of the signal $\{r(t)\}$, which is independent of $\{v_0(t)\}$. Also, note that $\Lambda_{\kappa_R \kappa_I}^N$ equals $\Lambda_{\kappa_R \kappa_I}$, if $N \rightarrow \infty$. Using $c_{N, \alpha}$ taken from the Standard Normal distribution, gives, asymptotically in N

$$\begin{aligned} & \sqrt{\tilde{N}} |\text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) - \eta_R - \beta_R + \text{Re}(\bar{G}_0(e^{i\omega_j}))| \\ & \leq c_{N, \alpha} \sqrt{\lambda_{\kappa_R}^N}, \quad \text{w.p. } \alpha. \end{aligned}$$

Dividing the left- and right-hand side by $\sqrt{\tilde{N}}$, using the bounds derived in Lemma's 4.1–4.3, and applying the triangle inequality, gives that, w.p. $\geq \alpha$

$$\begin{aligned} c_{N, \alpha} \sqrt{\frac{\lambda_{\kappa_R}^N}{\tilde{N}}} & \geq |\text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j}))| - |\eta_R| \\ & \quad - |\beta_R| - |\text{Re}(\bar{G}_0(e^{i\omega_j}))| \\ & \geq |\text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j}))| - \bar{\eta}_R \\ & \quad - \bar{\beta}_R - \bar{\delta}_R \end{aligned}$$

which proves Part 1). Part 2) can be proven similar to this.

If $\Lambda_{\kappa_R \kappa_I}^N$ is invertible, Part 4) of Lemma 4.4 is applicable. Using $c_{\chi, \alpha}$ taken from the Chi-square distribution with 2 degrees of freedom, this gives asymptotically in N

$$\begin{aligned} & \tilde{N} \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_R - \beta_R \\ \text{Im}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_I - \beta_I \end{bmatrix}^T \Gamma^T \\ & \quad \cdot \Gamma \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_R - \beta_R \\ \text{Im}(\hat{G}(e^{i\omega_j}) - \tilde{G}_0(e^{i\omega_j})) - \eta_I - \beta_I \end{bmatrix} \\ & \leq c_{\chi, \alpha}, \quad \text{w.p. } \alpha. \end{aligned}$$

Hence, w.p. $\geq \alpha$ (36) holds true, as shown at the bottom of the page. Dividing the left- and right-hand side by $\sqrt{\tilde{N}}$ and

applying the triangle inequality, gives that, w.p. $\geq \alpha$

$$\begin{aligned} \sqrt{\frac{c_{\chi, \alpha}}{\tilde{N}}} & \geq \left| \Gamma \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \\ \text{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \end{bmatrix} \right| - \left| \Gamma \begin{bmatrix} \eta_R \\ \eta_I \end{bmatrix} \right| + \\ & \quad - \left| \Gamma \begin{bmatrix} \beta_R \\ \beta_I \end{bmatrix} \right| - \left| \Gamma \begin{bmatrix} \text{Re}(\bar{G}_0(e^{i\omega_j})) \\ \text{Im}(\bar{G}_0(e^{i\omega_j})) \end{bmatrix} \right| \\ & \geq \left| \Gamma \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \\ \text{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \end{bmatrix} \right| + \\ & \quad - \left| \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix} \right| (|\text{Re}(\bar{G}_0(e^{i\omega_j}))| + |\eta_R| + |\beta_R|) + \\ & \quad - \left| \begin{bmatrix} \gamma_{12} \\ \gamma_{22} \end{bmatrix} \right| (|\text{Im}(\bar{G}_0(e^{i\omega_j}))| + |\eta_I| + |\beta_I|) \\ & \geq \left| \Gamma \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \\ \text{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) \end{bmatrix} \right| + \\ & \quad - \sqrt{\gamma_{11}^2 + \gamma_{21}^2} (\bar{\delta}_R + \bar{\eta}_R + \bar{\beta}_R) + \\ & \quad - \sqrt{\gamma_{12}^2 + \gamma_{22}^2} (\bar{\delta}_I + \bar{\eta}_I + \bar{\beta}_I) \end{aligned}$$

from which result 3) follows by taking the second and third term to the left and squaring left- and right-hand side. \square

B. Proof of Theorem 7.2

With $\{\psi(t)\}$ being independent of $\{e_0(t)\}$ it follows that

$$R_{\hat{\epsilon}}^N(\tau) := E \hat{R}_{\hat{\epsilon}}^N(\tau) = R_{\psi}^N(\tau) + R_{v_0}(\tau), \quad |\tau| \leq \tilde{N} - 1.$$

Denote matrix-element (i, j) of $\hat{\Lambda}_{r_1 r_2}(\tau)$ defined in (34) by $\hat{\lambda}_{r_i r_j}^N$. Then for $i, j = 1, 2$

$$E \hat{\lambda}_{r_i r_j}^N = \sum_{\tau=-w}^w c_w(\tau) \frac{\tilde{N} - |\tau|}{\tilde{N}} R_{r_i r_j}^N(\tau) [R_{\psi}^N(\tau) + R_{v_0}(\tau)].$$

Since $\{r_1(t)\}$ and $\{r_2(t)\}$ are bounded deterministic and jointly quasi-stationary, and therefore $R_{r_i r_j}^N(\tau)$ is bounded, $i, j = 1, 2$. Moreover, the auto-covariance function of the output error

$$R_{\hat{\epsilon}}(\tau) = \lim_{N \rightarrow \infty} R_{\hat{\epsilon}}^N(\tau) = R_{\psi}(\tau) + R_{v_0}(\tau)$$

shows exponential decay rate in τ as both terms on the right-hand side show exponential decay. Also, for $R_{r_i r_j}^N(\tau)$ defined in Lemma 4.4

$$\lim_{N \rightarrow \infty} R_{r_i r_j}^N(\tau) = R_{r_i r_j}(\tau), \quad \forall |\tau| \leq w$$

where w is allowed to tend to infinity as long as w/N tends to zero, which is assumed to be the case.

In line with [20] the following assumptions are imposed on the window $c_w(\tau)$, stating that the window should converge to one if $N \rightarrow \infty$, but slow enough in comparison to the number of data N .

$$\sqrt{\tilde{N}} \left| \Gamma \begin{bmatrix} \text{Re}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) - \eta_R - \beta_R + \text{Re}(\bar{G}_0(e^{i\omega_j})) \\ \text{Im}(\hat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})) - \eta_I - \beta_I + \text{Im}(\bar{G}_0(e^{i\omega_j})) \end{bmatrix} \right| \leq \sqrt{c_{\chi, \alpha}} \quad (36)$$

Assumption 2.1: The sequence of integers $\{w(N)\}$ is a positive, monotonously increasing sequence such that for some $k > \frac{5}{2}$

$$\lim_{N \rightarrow \infty} \frac{w(N)}{\sqrt{N}/\log^k N} = C$$

for some finite C , and the real-valued window-function $c_w(\tau)$ is such that $|c_w(\tau)| < C, \forall w, \tau$, and $\lim_{w \rightarrow \infty} c_w(\tau) = 1, \forall \tau$. \square

Under these conditions it follows that

$$\lim_{N \rightarrow \infty} E\hat{\Lambda}_{r_i r_j}^N = \lambda_{r_i r_j} + \sum_{\tau=-\infty}^{\infty} R_{r_i r_j}(\tau) R_{\psi}(\tau)$$

where $\lambda_{r_i r_j}$ denotes element (i, j) of matrix $\Lambda_{r_i r_j}$. Under the conditions mentioned in the theorem, the above assumptions on the window, and the weak conditions on the input signal $u(t)$ formulated in [20, (2.11) and (3.14)], [20, Th. 3.1] can be applied which yields that with probability one

$$\lim_{N \rightarrow \infty} \hat{\Lambda}_{r_i r_j}^N = \lim_{N \rightarrow \infty} E\hat{\Lambda}_{r_i r_j}$$

Due to the fact that $R_{\psi}(\tau)$ is exponentially decaying, there exists a stationary stochastic process n such that $R_n(\tau) = R_{\psi}(\tau)$ for all τ . Without loss of generality it may be assumed that n is independent of $\{r_i(t)\}, i = 1, 2$.

Denote the signals $q_i(t) = r_i(t)n(t), i = 1, 2$, then

$$\begin{aligned} & \sum_{\tau=-\infty}^{\infty} R_{r_i r_j}(\tau) R_{\psi}(\tau) \\ &= \sum_{\tau=-\infty}^{\infty} R_{r_i r_j}(\tau) R_n(\tau) \\ &= \sum_{\tau=-\infty}^{\infty} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E n(t) r_i(t) n(t+\tau) r_j(t+\tau) \\ &= \sum_{\tau=-\infty}^{\infty} R_{q_j q_i}(\tau) = \Phi_{q_j q_i}(0), \quad i, j = 1, 2 \end{aligned}$$

where $\Phi_{q_j q_i}(\omega)$ refers to the cross-spectrum between q_j and q_i . As $\Phi_{q_j q_i}(0) \geq 0$ this proves that $\Lambda_{\Psi} \geq 0$. \square

REFERENCES

- [1] D. S. Bayard, "Statistical plant set estimation using Schroeder-phased multisinusoidal input design," in *Proc. Amer. Contr. Conf.*, Chicago, IL, 1992, pp. 2988–2995.
- [2] D. R. Brillinger, *Time Series: Data Analysis and Theory*, expanded ed. New York: McGraw-Hill, 1981.
- [3] D. de Roover and R. G. Hakvoort, "Identification and robust control design of an industrial glass tube manufacturing process," in *Proc. Amer. Contr. Conf.*, Seattle, WA, 1995, pp. 19–23.
- [4] D. K. de Vries, "Identification of model uncertainty for control design," Ph.D. dissertation, Mech. Eng. Syst. Contr. Group, Delft Univ. Technol., 1994.
- [5] D. K. de Vries and P. M. J. Van den Hof, "Quantification of model uncertainty from data," *Int. J. Robust Nonlinear Contr.*, vol. 4, pp. 301–319, 1994.
- [6] ———, "Quantification of uncertainty in transfer function estimation: A mixed probabilistic worst case approach," *Automatica*, vol. 31, pp. 543–557, 1995.
- [7] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. New York: MacMillan, 1992.
- [8] E. Fogel and Y. F. Huang, "On the value of information in system identification—Bounded noise case," *Automatica*, vol. 18, pp. 229–238, 1982.
- [9] G. C. Goodwin, M. Gevers, and B. M. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 913–928, July 1992.
- [10] G. Gu and P. P. Khargonekar, "A class of algorithms for identification in \mathcal{H}_{∞} ," *Automatica*, vol. 28, pp. 299–312, 1992.
- [11] R. G. Hakvoort, "Worst-case system identification in \mathcal{H}_{∞} : Error bounds and optimal models," in *Prepr. 12th IFAC World Congr.*, Sydney, Australia, 1993, vol. 8, pp. 161–164.
- [12] R. G. Hakvoort and P. M. J. Van den Hof, "Identification of model error bounds in ℓ_1 - and \mathcal{H}_{∞} -norm," in *Proc. Santa Barbara Wkshp. Modeling Uncertainty in Contr. Syst.*, 1992, Lecture Notes in Contr. and Inf. Sciences. London: Springer-Verlag, 1993, vol. 192, pp. 139–155.
- [13] R. G. Hakvoort, "System identification for robust process control: Nominal models and error bounds," Ph.D. dissertation, Mech. Eng. Syst. Contr. Group, Delft Univ. Technol., The Netherlands, 1994.
- [14] ———, "A linear programming approach to the identification of frequency domain error bounds," in *Syst. Identification (SYSID'94)*, M. Blanke and T. Söderström, Eds. Oxford, UK: Pergamon, IFAC Postprint 1995, vol. IPPV, pp. 497–502.
- [15] R. G. Hakvoort and P. M. J. Van den Hof, "An instrumental variable procedure for the identification of probabilistic frequency response uncertainty regions," in *Proc. 33rd IEEE Conf. Decision Contr.*, Lake Buena Vista, FL, 1994, pp. 3596–3601.
- [16] ———, "Consistent parameter bounding identification for linearly parametrized model sets," *Automatica*, vol. 31, no. 7, pp. 957–969, 1995.
- [17] A. J. Helmicki, C. A. Jacobson, and C. N. Nett, "Identification in \mathcal{H}_{∞} : A robustly convergent nonlinear algorithm," in *Proc. Amer. Contr. Conf.*, San Diego, CA, 1990, pp. 386–391.
- [18] ———, "Control-oriented system identification: A worst-case/deterministic approach in \mathcal{H}_{∞} ," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1163–1176, 1991.
- [19] P. S. C. Heuberger, P. M. J. Van den Hof, and O. H. Bosgra, "A generalized orthonormal basis for linear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 451–465, Mar. 1995.
- [20] H. Hjalmarsson, "Aspects on incomplete modeling in system identification," Ph.D. dissertation, Dept. Electr. Eng., Linköping Univ., Sweden, 1993.
- [21] ———, "A model variance estimator," in *Prepr. 12th IFAC World Congr.*, Sydney, Australia, 1993, vol. 9, pp. 5–10.
- [22] H. Hjalmarsson and L. Ljung, "Estimating model variance in the case of undermodeling," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 1004–1008, 1992.
- [23] C. A. Jacobson and C. N. Nett, "Worst case identification in ℓ_1 : Optimal algorithms and error bounds," in *Proc. Amer. Contr. Conf.*, Boston, MA, 1991, pp. 3152–3157.
- [24] R. L. Kosut, M. K. Lau, and S. P. Boyd, "Set-membership identification of systems with parametric and nonparametric uncertainty," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 929–941, 1992.
- [25] R. O. Lemaire, L. Valavani, M. Athans, and G. Stein, "A frequency domain estimator for use in adaptive control systems," *Automatica*, vol. 27, pp. 23–38, 1991.
- [26] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [27] L. Ljung, B. Wahlberg, and H. Hjalmarsson, "Model quality: The role of prior knowledge and data information," in *Proc. 30th IEEE Conf. Decision Contr.*, Brighton, U.K., 1991, pp. 273–278.
- [28] P. M. Mäkilä, "Robust identification and Galois sequences," *Int. J. Contr.*, vol. 54, pp. 1189–1200, 1991.
- [29] E. B. Manoukian, *Modern Concepts and Theorems of Mathematical Statistics*. New York: Springer-Verlag, 1986.
- [30] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, pp. 997–1009, 1991.
- [31] B. M. Ninness, "Stochastic and deterministic modeling," Ph.D. dissertation, Univ. Newcastle, Australia, 1993.
- [32] J. P. Norton, "Identification and application of bounded-parameter models," *Automatica*, vol. 23, pp. 497–507, 1987.
- [33] P. J. Parker and R. R. Bitmead, "Adaptive frequency response identification," in *Proc. 26th Conf. Decision Contr.*, Los Angeles, CA, 1987, pp. 348–353.
- [34] J. R. Partington, "Robust identification in \mathcal{H}_{∞} ," *J. Math. Anal. Appl.*, vol. 166, pp. 428–441, 1991.
- [35] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model invalidation," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 951–959, 1994.

- [36] R. S. Smith and J. C. Doyle, "Model invalidation: A connection between robust control and identification," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 942–952, 1992.
- [37] T. Söderström and P. Stoica, *Instrumental Variable Methods for System Identification*, Lect. Notes in Contr. and Inf. Sciences. Berlin: Springer-Verlag, 1983, vol. 57.
- [38] ———, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [39] D. N. C. Tse, M. A. Dahleh, and J. N. Tsitsiklis, "Optimal asymptotic identification under bounded disturbances," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1176–1190, 1993.
- [40] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor, "System identification with generalized orthonormal basis functions," *Automatica*, vol. 31, no. 12, pp. 1821–1834, 1995.
- [41] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Automat. Contr.* vol. 36, pp. 551–562, 1991.
- [42] B. Wahlberg and L. Ljung, "Hard frequency-domain model error bounds from least-squares like identification techniques," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 900–912, 1992.
- [43] B. Wahlberg, "System identification using Kautz models," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1276–1282, 1994.
- [44] R. C. Younce and C. E. Rohrs, "Identification with nonparametric uncertainty," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 715–728, 1992.
- [45] Y. C. Zhu, "Estimation of transfer functions: Asymptotic theory and a bound of model uncertainty," *Int. J. Contr.*, vol. 49, pp. 2241–2258, 1989.



Richard G. Hakvoort was born in Delft, The Netherlands, in 1968. He obtained the M.Sc. and Ph.D degrees in mechanical engineering from Delft University of Technology, The Netherlands, in 1990 and 1994, respectively.

Since 1995, he has been a Research Petrophysicist at Shell Research in Rijswijk, The Netherlands. His current research interests include modeling of electrical wireline tools and construction of efficient algorithms to process electrical logging data.

Dr. Hakvoort received for his M.Sc. thesis the 1990 Control Technology Award of the Dutch Royal Institute of Engineers (KIVI).



Paul M. J. Van den Hof (S'85–M'88–SM'95) was born in Maastricht, The Netherlands, in 1957. He obtained the M.Sc. and Ph.D. degrees both from the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands, in 1982 and 1989, respectively.

Since 1986, he has been working in the Mechanical Engineering Systems and Control Group at Delft University of Technology, The Netherlands, where since 1991, he holds a position as an Associate Professor. In 1992, he has held a short-term visiting position at the Centre for Industrial Control Science, The University of Newcastle, NSW, Australia. His research interests include issues of system identification, parameterization, and the interplay between identification and robust control design, with applications in mechanical servo systems and industrial process control systems.

Dr. Van den Hof is an Associate Editor of *Automatica*.