Frequency domain uncertainty sets with guaranteed probability level in Prediction Error Identification †

X. Bombois[(1)], B.D.O. Anderson[(2)], M. Gevers[(3)]

[(1)] Delft Center for Systems and Control, Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
Tel: +31-15-2785150, Fax: +31-15-2786679, Email: X.J.A.Bombois@dcsc.tudelft.nl
[(2)] Research School of Information Sciences and Engineering
Australian National University, and National ICT Australia
[(3)] CESAME, Université Catholique de Louvain, Bâtiment Euler
B1348 Louvain la Neuve, Belgium

Extended Abstract

A model $G(z,\hat{\theta})$ identified in the Prediction Error (PE) framework [6] is always an approximation of the real-life system $G(z,\theta_0)$ we want to identify. A model is therefore useless if it is not accompanied with information about the achieved error $G(z,\hat{\theta}) - G(z,\theta_0)$. A classical domain where it is convenient to express the features of this error is the frequency domain (i.e. the Nyquist plane). In this presentation, we characterize this error in the frequency domain by determining a set of parametrized transfer functions $G(z,\theta)$ whose frequency response is at each frequency constrained to lie in an ellipse in the Nyquist plane and that has the property to contain the true system $G(z,\theta_0)$ at a probability level of at least α (say, 95%). We shall call such a set a frequency domain uncertainty region. This uncertainty region can be subsequently used e.g. for the design of a robust controller for $G(z,\theta_0)$ [10].

We consider the identification of model structures $G(z,\theta)$ that are linear in the parameter vector θ: $G(z,\theta) = \Lambda(z)\theta$, where $\Lambda(z)$ is a set of independent rational basis functions. Our results also hold for more general model structures, modulo a linearization of these structures around the identified model. We further assume that the true system can be parametrized exactly within the chosen model structure for some parameter vector θ_0, i.e. $G_0(z) = G(z,\theta_0)$, and that it can be represented by $y(t) = G_0(z)u(t) + v(t)$ where $v(t)$ is additive (colored) Gaussian noise.

A Prediction Error identification experiment consists of the collection of a finite number of input and output data, from which an estimate $\hat{\theta}$ of the parameter vector θ_0 is estimated using a prediction error criterion: see [6] for details. Due to the stochastic assumptions on the noise $v(t)$ and the assumptions on the model structure, the parameter vector $\hat{\theta}$ is a Gaussian random variable with mean θ_0 and a covariance matrix P_θ that can be estimated from the data[1]:

$$\hat{\theta} \sim \mathcal{N}(\theta_0, P_\theta). \tag{1}$$

†The European Commission is herewith acknowledged for its financial support in part to the research reported on in this presentation. The support is provided via the Program Training and Mobility of Researchers (TMR) and Project System Identification (ERB FMRX CT98 0206) to the European Research Network System Identification (ERNSI). The authors also acknowledge the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. National ICT Australia is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through *Backing Australia's Ability* and the ICT Centre of Excellence program. The scientific responsibility rests with its authors.

[1]The Gaussian assumption on $v(t)$ is not essential for the prediction error framework; it can be replaced by an assumption of quasistationarity on $v(t)$, in which case the estimated parameter vector $\hat{\theta}$ is asymptotically Gaussian.

By the linearity of the model structure, the frequency response of the estimated model $G(z,\hat{\theta})$ is also Gaussian. If we represent the frequency response of the model $G(z,\theta)$ by a vector $g(e^{j\omega},\theta)$ containing its real and imaginary parts

$$g(e^{j\omega},\theta) \triangleq \begin{pmatrix} Re(G(e^{j\omega},\theta)) \\ Im(G(e^{j\omega},\theta)) \end{pmatrix} = \overbrace{\begin{pmatrix} Re(\Lambda(e^{j\omega})) \\ Im(\Lambda(e^{j\omega})) \end{pmatrix}}^{T(e^{j\omega})} \theta, \tag{2}$$

we then have

$$\hat{g}(e^{j\omega}) \triangleq g(e^{j\omega},\hat{\theta}) \sim \mathcal{N}(g(e^{j\omega},\theta_0), P(\omega)) \tag{3}$$

$$\text{with} \quad P(\omega) \triangleq T(e^{j\omega}) P_\theta T(e^{j\omega})^T \in \mathbf{R}^{2\times 2}. \tag{4}$$

Using the normal distribution of $g(e^{j\omega},\hat{\theta})$, we can easily build, at each frequency in the Nyquist plane, elliptic confidence regions $U(\omega,\chi_2)$ that are guaranteed to contain the frequency response of $g(e^{j\omega},\theta_0)$ at a prescribed probability level α:

$$U(\omega,\chi_2) = \{g \in \mathbf{R}^2 \mid (g-\hat{g}(e^{j\omega}))^T P(\omega)^{-1} (g-\hat{g}(e^{j\omega})) \leq \chi_2\} \tag{5}$$

The size χ_2 of the confidence region $U(\omega,\chi_2)$ is determined by the probability level α via the χ^2 distribution with 2 degrees of freedom: it is chosen such that $Pr(\chi^2(2) < \chi_2) = \alpha$. Examples of such confidence regions abound in the literature (see e.g. [4, 1, 8, 5, 9]) and they can be computed with the System Identification Toolbox of Matlab[2]).

Using the above results, constructing a frequency domain uncertainty set containing $G_0(z)$ at the probability level α could seem simple: we just glue together the ellipses $U(\omega,\chi_2)$ at each frequency in order to obtain the set $\mathcal{L}(\chi_2) \triangleq \{G(z,\theta) \mid g(e^{j\omega},\theta) \in U(\omega,\chi_2)\ \forall\omega\ \}$. Such an approach is nevertheless not correct. Indeed, as we will show, the probability that $G_0(z) \in \mathcal{L}(\chi_2)$ (or, equivalently, $Pr(g(e^{j\omega},\theta_0) \in U(\omega,\chi_2)\ \ \forall\omega)$) is always smaller (and can be much smaller) than α.

Note that the difficulty of extrapolating the confidence level α obtained for the individual uncertainty sets $U(\omega,\chi_2)$ to a confidence level for the uncertainty region $\mathcal{L}(\chi_2)$ obtained by connecting these sets together has been noticed by Tjärnström in [7]. In that paper, in order to obtain an overall confidence region for the true $G_0(z)$ with a predefined confidence level, Tjärnström proposes a method based on the bootstrap technique developed in statistics in the seventies. In a nutshell, it consists of estimating the probability distribution of the prediction errors, and then generating a large number of simulated input-output data sets from the known inputs and residuals drawn from the estimated distribution function. For each of these data sets, a model is identified. An uncertainty set with prespecified confidence level can then be computed experimentally from the large number of estimated models. The procedure is interesting, but very heavy on computer time.

The problem considered in this presentation is therefore similar to the one addressed by Tjärnström, namely how to construct an uncertainty region for the whole transfer function $G_0(e^{j\omega})$ with a prespecified confidence level, when this transfer function is obtained by Prediction Error Identification. However, we propose a much simpler procedure based on the commonly used elliptic uncertainty sets $U(\omega,\chi)$. More precisely, we solve the following problem: "*How should one choose the size χ of the ellipses $U(\omega,\chi)$ in such a way that the probability that $G_0(z) \in \mathcal{L}(\chi)$ is at least equal to some prespecified level α?*".

To address this problem, we first analyze why the choice $\chi = \chi_2$ leads to a set for which $Pr(G_0(z) \in \mathcal{L}(\chi_2)) < \alpha$. For this purpose, we study the parameter set $C_\theta(U(\omega,\chi_2))$ defined at each frequency as the set of parameters θ such that the frequency response of $G(z,\theta)$, evaluated at $z = e^{j\omega}$, lies in $U(\omega,\chi_2)$. We show that $C_\theta(U(\omega,\chi_2))$ is a different parameter set at each frequency, and that these sets have infinite size.

[2]However, there, the uncertainty regions are represented in a bode plot as uncertainty bands around the magnitude and the phase.

The probability that the entire frequency response $g(e^{j\omega}, \theta_0)$ of the true system lies in **every** $U(\omega, \chi_2)$ (i.e. $Pr(G_0(z) \in \mathcal{L}(\chi_2))$) is thus equal to the probability that θ_0 lies in the intersection of the sets $C_\theta(U(\omega, \chi_2))$ over all frequencies, and is consequently smaller than $Pr(g(e^{j\omega}, \theta_0) \in U(\omega, \chi_2)) = Pr(\theta_0 \in C_\theta(U(\omega, \chi_2))) = \alpha$. We will present examples showing that the former probability can indeed be much smaller than the latter, particularly as the size of the parameter vector θ increases.

Then, in order to determine a size χ for which $Pr(G_0(z) \in \mathcal{L}(\chi))$ is at least α, we proceed in two steps:

1. first, we provide an estimate for the intersection of the sets $C_\theta(U(\omega, \chi))$ over all frequencies, by observing that the parameter set $U_\theta(\chi) \triangleq \{\theta \mid (\theta - \hat{\theta})^T P_\theta^{-1} (\theta - \hat{\theta}) \leq \chi\}$ is a (strict) subset of this intersection. As a result, $Pr(g(e^{j\omega}, \theta_0) \in U(\omega, \chi) \ \ \forall \omega) > Pr(\theta_0 \in U_\theta(\chi))$.

2. the solution to the problem is then to select a χ such that $Pr(\theta_0 \in U_\theta(\chi))$ is equal to the desired level α. Due to the distribution (1), this is trivial. Let us stress that, as soon as as $\dim(\theta) > 2$, such χ is larger than χ_2.

To summarize, the main message and contribution of this presentation is to pinpoint the fact that the intuitive and commonly used procedure that consists of "gluing together" in the Nyquist plane ellipses that contain the frequency response of the true system at every frequency with probability α, say 0.95, does not produce an uncertainty set that contains the true system with probability α at all frequencies. The actual confidence level for this whole uncertainty set is smaller, and can be much smaller for moderate size parameter vectors. However, as we show in the presentation, it is actually easy to construct a global uncertainty set that contains the true system with a probability of at least α. To achieve this, one must construct ellipses of larger size χ than the size that produces a probability level of α at any given frequency.

Finally, note that we published most of the technical results of this presentation in [2, 3]. However, the novelty of this presentation lies in the interpretation we give to these technical results. Note also that an extended version of this presentation has been submitted to Systems and Control Letters in December 2003.

References

[1] P. Andersen, S. Toffner-Clausen, and T.S. Pedersen. Estimation of frequency domain model uncertainties with application to robust control design. In *Proc. IFAC Symposium on System Identification*, pages 603–608, Copenhagen, 1994.

[2] X. Bombois, B.D.O. Anderson, and M. Gevers. Frequency domain image of a set of linearly parametrized transfer functions. 6th European Control Conference, paper WA07-1, pp. 1416-1421, Porto, 2001.

[3] X. Bombois, B.D.O. Anderson, and M. Gevers. Mapping parametric confidence ellipsoids to Nyquist plane for linearly parametrized transfer functions. In G.C. Goodwin, editor, *Model Identification and Adaptative Control*, pages 53–71. Springer Verlag, 2000.

[4] G.C. Goodwin, M. Gevers, and B. Ninness. Quantifying the error in estimated transfer functions with application to model order selection. *IEEE Trans. Automatic Control*, 37:913–928, 1992.

[5] R.G. Hakvoort and P.M.J. Van den Hof. Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors. *IEEE Trans. Automatic Control*, 42(11):1516–1528, 1997.

[6] L. Ljung. *System Identification: Theory for the User, 2nd Edition.* Prentice-Hall, Englewood Cliffs, NJ, 1999.

[7] F. Tjarnstrom. Computing uncertainty regions with simultaneous confidence degree using bootstrap. In *Proc. IFAC Symposium on System Identification*, pages 1133–1138, Santa Barbara, 2000.

[8] S. Toffner-Clausen, P. Andersen, J. Stoustrup, and H.H. Niemann. Estimated frequency domain model uncertainties used in robust controller design - a μ-approach. In *Proc. IEEE Conference on Control Applications*, pages 1585–1590, Glasgow, 1994.

[9] L. Wang and G.C. Goodwin. Integrating identification with robust control: a mixed H_2/H_∞ approach. In *CD-ROM Proc. 39th Conference on Decision and Control*, Sydney, Australia, 2000.

[10] K. Zhou and J. Doyle. *Essentials of Robust Control*. Prentice Hall, Upper Saddle River, New Jersey, 1998.