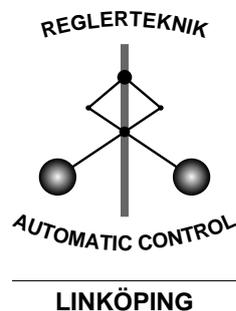# On Model Error Modeling in Set Membership Identification

Andrea Garulli* and Wolfgang Reinelt#

*Dipartimento di Ingegneria dell'Informazione
Università di Siena, 53100 Siena, Italy
WWW: `http://www-dii.ing.unisi.it/~garulli/`
Email: garulli@ing.unisi.it

#Department of Electrical Engineering
Linköping University, 581 83 Linköping, Sweden
WWW: `http://www.control.isy.liu.se/~wolle/`
Email: `wolle@isy.liu.se`

October 1999

REGLERTEKNIK

AUTOMATIC CONTROL

LINKÖPING

On model error modeling in set membership identification

Andrea Garulli[*]   and   Wolfgang Reinelt[#]

[*]Dipartimento di Ingegneria dell'Informazione, Università di Siena
Via Roma 56, 53100 Siena, Italy
E-mail: garulli@ing.unisi.it

[#]Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden
E-mail: wolle@isy.liu.se

Abstract

A recent perspective on model error modeling is applied to set membership identification techniques in order to highlight the separation between unmodeled dynamics and noise. Model validation issues are also easily addressed in the proposed framework. The computation of the minimum noise bound for which a nominal model is not falsified by i/o data, can be used as a rationale for selecting an appropriate model class. Uncertainty is evaluated in terms of the frequency response, so that it can be handled by $\mathcal{H}_\infty$ control techniques.

Keywords: identification for robust control, set membership, model error modeling, unmodeled dynamics, model validation.

1   Introduction

One of the main objectives of control-oriented identification is to estimate models that are suitable for robust control design techniques. To this purpose, the identification procedure must deliver not only a nominal model, but also a reliable estimate of the uncertainty associated to the model. Different paradigms for the description of uncertainty have been addressed in the literature (see e.g. the special issues [1, 2] and the recent book [3]). One of the most popular relies on the hypothesis that the identification error is unknown-but-bounded (UBB) and it is known as *bounded error* or *set membership* identification (see [4, 5] and the references therein).

In standard identification problems, the error originates from two different sources: a "variance" term, due to noise affecting the data, and a "bias" term, due to system dynamics which is not captured by the estimated nominal model (often addressed also as the *model error*). Clearly, the nature of these two error terms is quite different: the former is generally uncorrelated with the input signal, while the latter strongly depends on the estimated nominal model and on the input used in the identification experiment. The "size" of the model error is not negligible in most practical situations, especially those in which the order of the nominal model must be small (a typical requirement of robust design techniques). Moreover, while a priori information on measurement noise are often available, similar hypotheses on the unmodeled dynamics seem to be less realistic.

The first works on set membership identification [6, 7, 8] provided efficient algorithms for estimating the set of feasible models, compatible with the UBB error assumption, without separating

the contribution from unmodeled dynamics and noise (or equivalently, assuming that the plant generating the data belong to the considered model family). Model errors were explicitly accounted for in [9, 10]. Estimation of the size of the model error from residual data has been adopted in the set membership context in [11], for nominal models obtained via least-squares identification. Separation of unmodeled dynamics and noise in a setting including both stochastic and deterministic uncertainties has been studied in [12].

In this paper, it is shown how model error modeling concepts adopted by classical identification techniques [13, 14] can be easily extended to set membership identification, providing a general strategy for separating error contributions. A model error model identification procedure based on set membership techniques is formulated. In this context, error bounds can be used to extract information about the size of the unmodeled dynamics and the quality of the nominal model. Frequency domain uncertainty regions are considered in order to provide adequate models for $\mathcal{H}_\infty$ control design.

The paper is organized as follows. In Section 2, model error modeling ideas are briefly summarized. Section 3 introduces the main concepts of set membership identification and describes a procedure for identification of both nominal model and model error. The features of the proposed approach are discussed in Section 4, while Section 5 reports an example which clarifies the effectiveness of set membership model error modeling. Finally, Section 6 gives some concluding remarks.

2   Model error modeling

In this section, the basic ideas of model error modeling are briefly resumed (see [14] for a thorough treatment). Let (u_m, y_m) be a collection of i/o measurements and assume that a nominal model G_n of the system that generated the data has been estimated, according to some identification procedure. Then, the model error modeling strategy can be summarized as follows.

1. Compute the residual $\epsilon = y_m - G_n u_m$.

2. Consider the "error" system, with input u_m and output ϵ, and identify a model G_e for this system. This is an estimate of the error due to undermodeling, the so-called *model error model*. Standard statistical properties of the estimated model lead to the definition of the *uncertainty region of the model error* (e.g., given by the 99% confidence region of the estimated model error).

3. The *uncertainty region of the nominal model* is obtained by adding up the nominal model and the uncertainty region of the model error.

4. Model validation: the nominal model is not falsified if and only if it lies inside its own uncertainty region or, equivalently, if and only if zero is an element of the model error uncertainty region. This can be easily checked by looking at the corresponding uncertainty bands in the Bode or Nyquist plots.

Identification of the model error from residual data provides a separation between noise and unmodeled dynamics. In fact, G_e can be seen as an estimate of the dynamic system $\Delta(\cdot)$ (possibly nonlinear, time-varying, etc.), such that

$$\epsilon = \Delta(u) + v \tag{1}$$

where v is the noise term, which is assumed to be uncorrelated with the input u.

A few observations can be made.

- The model error G_e must be selected among a pre-specified class of models. Unfortunately, there is not a standard procedure for selecting the structure of the model error. This is rather based on a priori knowledge on the system generating the data, and on the purposes for which the model uncertainty is estimated. For example, a typical requirement is that unmodeled dynamics is captured at those frequencies that are significant for control design purposes.

- The size of the uncertainty region clearly depends on the selected nominal model. When a severe undermodeling occurs (due for example to the fact that we are required to deliver a low order nominal model to the robust control designer), one cannot expect the model error uncertainty region to be small at all frequencies, i.e. the nominal model will be falsified, whatever the structure of the model error.

- A falsified nominal model may still be accepted, provided that the related uncertainty band is small at frequencies of interest. However, it would be useful to evaluate the "minimum" model error uncertainty such that the nominal model is not falsified. As we will see, this can be easily done in the framework of set membership identification.

3 Set membership identification

The model error modeling procedure described in the previous section can be applied to any nominal model G_n. In this paper, nominal models provided by set membership identification techniques will be considered. In the following, we summarize the main features of these techniques.

Assume that the measured i/o data are generated by a system S_0, according to

$$y_m = S_0(u_m) + v$$

where S_0 belongs to a set $\mathcal{K}$ (the a priori information on the system) and the noise v is bounded in some norm $\mathcal{Y}$, i.e. $\|v\|_{\mathcal{Y}} \leq \delta$, for given $\delta > 0$. Then it is possible to define the *feasible system set*

$$FSS = \{S \in \mathcal{K} : \|y_m - S(u_m)\|_{\mathcal{Y}} \leq \delta\} \tag{2}$$

which is the set of systems that are compatible with the measured data and the a priori assumptions. The set FSS can be very complicated, depending on the structure of $\mathcal{K}$ and the norm $\|\cdot\|_{\mathcal{Y}}$. For example, if $\mathcal{K}$ is the class of LTI systems whose impulse response has an assigned exponential decay and noise is ℓ_∞-bounded, then FSS is an infinite dimensional polytope. Obviously, for real-world systems, including nonlinearities or time-varying drifts, FSS may be much more complex.

Since the set FSS contains all the information provided by data and a priori assumptions, it is natural to evaluate the quality of a nominal model G_n according to its worst-case error with respect to elements of FSS. In other words, the *identification error* associated to G_n is given by

$$E(G_n) = \sup_{S \in FSS} \|S - G_n\|_{\mathcal{S}} \tag{3}$$

where $\|\cdot\|_{\mathcal{S}}$ is a suitable norm in the system space.

In order to identify a nominal model, a model class must be chosen. A common requirement is that the nominal model must be simple (low dimensional, linearly parameterized, etc.). Hence, a typical structure for G_n is

$$G_n(q; \theta) = \sum_{i=1}^{n} \theta_i B_i(q) \tag{4}$$

where the $B_i(q)$ are user defined basis functions, such as FIR filters, Laguerre or Kautz filters [15], orthonormal basis functions [16], etc. If we denote by $\mathcal{M}$ the set of nominal models parameterized as in (4), the problem of selecting a model in $\mathcal{M}$ according to the criterion (3) is usually addressed as *conditional* set membership identification [17]. The optimal nominal model is given by the conditional central estimate

$$G_n^* = \arg \inf_{G \in \mathcal{M}} \sup_{S \in FSS} \|S - G\|_{\mathcal{S}}. \tag{5}$$

In most practical situation, finding an exact solution of the min-max optimization problem (5) is a prohibitive task and computationally tractable alternatives must be looked for. These may include recursive approximations of the set FSS (employing different approximating regions like ellipsoids, orthotopes, parallelotopes, see e.g. [6, 8, 18]), and/or suboptimal pointwise estimators like projection algorithms, interpolatory algorithms, etc. The evaluation of bounds on the identification error of these algorithms has been tackled in different contexts and several results are now available (see e.g. [19, 20]). On the other hand, these bounds may be conservative in many situations of interest for several reasons. First, the identification error (3) depends on FSS, whose computation is often untractable. Moreover, the tightness of these bounds is related to that of the noise bound δ: if the latter is overestimated, the former will also be excessively large.

Whatever nominal model has been identified using the above procedures, the actual size of the unmodeled dynamics from available data can be evaluated via the error modeling technique described in Section 2. The identification of the error system (1) can be performed using set membership identification algorithms, exploiting the noise bound on v. If the structure of the model error model is chosen as

$$G_e(q; \bar{\theta}) = \sum_{i=1}^{\bar{n}} \bar{\theta}_i \bar{B}_i(q) \tag{6}$$

the feasible set for the model error parameters is given by

$$FES = \{\bar{\theta} \in \mathbb{R}^{\bar{n}} : \|\epsilon - \sum_{i=1}^{\bar{n}} \bar{\theta}_i \bar{B}_i(q) u_m\|_{\mathcal{Y}} \le \delta\}. \tag{7}$$

Then, one may select the worst-case optimal error model by computing the Chebishev center of FES

$$\bar{\theta}^* = \arg \inf_{\bar{\theta} \in \mathbb{R}^{\bar{n}}} \sup_{\tilde{\theta} \in FES} \|\tilde{\theta} - \bar{\theta}\| \tag{8}$$

where $\| \cdot \|$ denotes the Euclidean norm (of course, other norms can be chosen).

As for (5), problem (8) may be computationally unfeasible, and suboptimal solutions are sought. For example, let us consider the case of ℓ_∞-bounded noise. Since FES is a polytope in $\mathbb{R}^{\bar{n}}$, it can be recursively approximated by ellipsoids or parallelotopes, and the center of these approximating sets may be chosen as an estimate of $\bar{\theta}^*$.

If robust identification is oriented to $\mathcal{H}_\infty$ control, an uncertainty band associated to the nominal model frequency response must be delivered to the control designer. In this framework, the frequency plot of the error model uncertainty region can be obtained by mapping the set FES onto the complex plane for each frequency of interest. This leads to the frequency domain uncertainty set

$$\mathcal{V}(FES) = \bigcup_{\omega} \mathcal{V}_\omega(FES)$$

where

$$\mathcal{V}_\omega(FES) = \{z \in \mathbb{C} : \ z = A_\omega \bar{\theta}, \ \bar{\theta} \in FES\} \tag{9}$$

and $A_\omega = [\bar{B}_1(e^{j\omega}) \dots \bar{B}_{\bar{n}}(e^{j\omega})] \in \mathbb{C}^{1 \times \bar{n}}$. Once again, the computation of $\mathcal{V}(FES)$ may be a formidable task if the exact FES defined by (7) is considered. Hence, set approximations are useful also in this respect.

The overall set membership identification strategy can be summarized as follows.

1. Identify the nominal model $G_n(q; \theta)$, using conditional estimators based on the feasible system set (2).

2. Compute the residual $\epsilon = y_m - G_n(q; \theta)u_m$.

3. Select a model error structure and compute (or approximate) the feasible model error set (7).

4. Identify a nominal model error $G_e(q; \bar{\theta})$, using optimal or suboptimal estimators based on FES.

5. Map the nominal model plus the model error and its uncertainty region onto the frequency domain.

Remark. The use of residual data in set membership identification for evaluating the worst-case $\mathcal{H}_\infty$ norm of the unmodeled dynamics has been introduced in [11], for standard least-squares nominal models. The above set membership model error modeling strategy can be seen as a general framework in which the structures of the nominal and error models, and the corresponding identification algorithms must be chosen by the user according to the specific problem (a priori knowledge, noise bound, error norm, etc.) In the next section, it will be shown how it is possible to exploit this framework to obtain a useful measure of the size of the unmodeled dynamics. This will also provide useful information on the selection of the nominal model structure.

4   Model validation and model class selection

When identification of the model error model is performed via standard prediction error methods, one can choose the 99% confidence region of the model error as the uncertainty set to be added to the nominal model [14]. However, it is very likely that the nominal model does not lie inside the resulting uncertainty region, whenever a remarkable undermodeling occurs.

For this reason, it is useful to evaluate the "distance from validation" of the nominal model. In the following, this is done in the set membership setting of Section 3. Let us consider the set $\mathcal{V}_\omega(FES)$ in (9) and define

$$d(\omega) = \min_{z \in \mathcal{V}_\omega(FES)} |z|. \tag{10}$$

If we set

$$d = \sup_\omega d(\omega) \tag{11}$$

then it is clear that the nominal model can be deemed to be unfalsified if $d = 0$. If ellipsoidal approximations of FES are considered, the sets $\mathcal{V}_\omega(FES)$ are ellipses in the complex plane and the optimization problem (10) can be easily solved. Then, d in (11) can be approximated by taking the maximum over a finite number of frequencies in the range of interest.

It is worth observing that d is a function of δ, since FES (and its frequency image $\mathcal{V}_\omega(FES)$) depend on the noise bound. Then, one can look for the minimum value of δ for which the nominal model is still not falsified by data, i.e.

$$\delta^* = \min_{\{\delta:\ d=0\}} \delta. \tag{12}$$

Since d is an increasing function of δ, the computation of δ^* can be easily performed within the desired precision, using a standard bisection on δ.

It is believed that the quantity δ^* is important in several respects.

- It allows one to deliver the smallest frequency domain uncertainty region associated to the model error model, which does not falsify the nominal model. This meets a typical requirement of the control designer: the "tightest" uncertainty band around the nominal model, without contradicting the information given by data.

- It can be used to "validate" (or invalidate) the a priori noise bound. In fact, if δ^* turns out to be much smaller than the a priori noise bound δ, one may conclude that this bound was overestimated. Conversely, underestimation of the noise bound can be detected by high values of δ^*, provided that a sufficiently rich model class has been chosen for the identification of the model error model.

- It gives a rationale for selecting the structure of the nominal model. Let us denote by p a vector of parameters that define a model class for G_n (for example, the number and/or the pole locations of the family of basis functions $B_i(q)$). For each p, one can perform set membership identification of nominal model and model error model as suggested in Section 3, and compute $\delta^*(p)$. Then, at least in principle, one can choose the optimal model structure by solving

$$\inf_p \delta^*(p). \tag{13}$$

In practice, the optimization in (13) can be performed over a finite set $\mathcal{P} = \{p_1 \ldots p_m\}$ (see the example in the next section, where p is the pole of a Laguerre expansion of fixed order).

5 Example

Let us consider the same plant as in [21], fed with an harmonic input signal. The output is corrupted by additive white noise, normally distributed with variance 0.0625. Notice that this is an unfavorable situation in the set membership framework, since the ℓ_∞ norm of the noise v is about 1, but only 10 samples out of 4000 used in the identification experiment are larger than 0.7. This means that, except for a few "outliers" the noise bound $\delta = 1$ is overestimated, while it is well known that set membership identification algorithms enjoy nice properties when the noise is boundary visiting.

Laguerre expansions of order $n = 4$ and $\bar{n} = 10$ are selected as nominal model class $\mathcal{M}$ and model error model class, respectively. The pole of both Laguerre bases is chosen as $p = 0.96$ (this choice will be discussed later). As a nominal model, we take $G_n(q, \theta^r)$, where

$$\theta^r = \arg \inf_{\theta \in \mathbb{R}^n} \|y_m - \sum_{i=1}^n \theta_i B_i(q) u_m\|_\infty.$$

This is known as restricted projection estimate and can be computed by linear programming (its properties in the worst-case setting are discussed in [20]). The model error model $G_e(q; \bar{\theta})$ is estimated by recursively approximating the set FES in (7) (with $\| \cdot \|_{\mathcal{Y}} = \ell_\infty$ and $\delta = 1$) by ellipsoids, and then taking $\bar{\theta}$ as the center of the ellipsoid.

Figure 1 shows the resulting frequency uncertainty region for both the nominal model and the model error, as suggested by the model error modeling approach outlined in Section 2. The true plant is also reported (dashed line in the upper plot) in order to show that a real undermodeling occurs at high frequencies. Nevertheless, the uncertainty is pretty narrow at frequencies up to $\omega = 5 \ rad/s$, and from the lower plot it is evident that the nominal model is not falsified by the data.
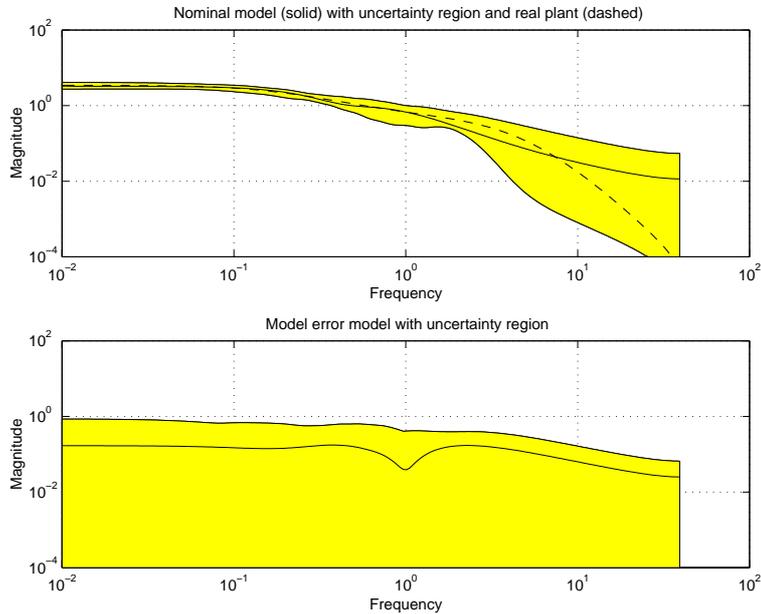

Figure 1: Nominal model and model error model for a Laguerre expansion with pole $p = 0.96$ ($\delta = 1$). The true plant is shown in the upper plot (dashed line).

Let us consider δ^* in (12). From a numerical optimization over different values of δ, one gets $\delta^* \approx 0.82$. Loosely speaking, this "confirms" the a priori assumption on the noise bound. The nominal model and uncertainty band for $\delta = \delta^*$ are shown in Figure 2a, while the corresponding plot for an overestimated noise bound ($\delta = 2$) is reported in Figure 2b. In the latter case, the actual value of δ^* allows one to deem the obtained uncertainty region unreasonably large.

An obvious criticism is that a rigorous way for selecting the model classes for both the nominal and the model error model is lacking. As said above, the structure of the model error model should rely on the purpose for which the model is used. For example, if one wants to estimate more carefully the unmodeled dynamics at high frequencies, a possible choice is a Kautz model with two real poles in 0.96 and 0.7. The result is shown in Figure 3 and it presents a wider uncertainty band for $\omega > 5 \ rad/s$, as one might expect.

As long as the nominal model structure is concerned, the choice can be based on the value of δ^*, as explained in Section 4. In Figure 4, δ^* is plotted for different values of the Laguerre pole p, in the case when a Laguerre expansion is used for both nominal and model error model ($n = 4$, $\bar{n} = 10$). It can be seen that $p = 0.96$ guarantees model validation with the lowest noise bound ($\delta \approx 0.82$). This corresponds to the tightest uncertainty band around the nominal model, among
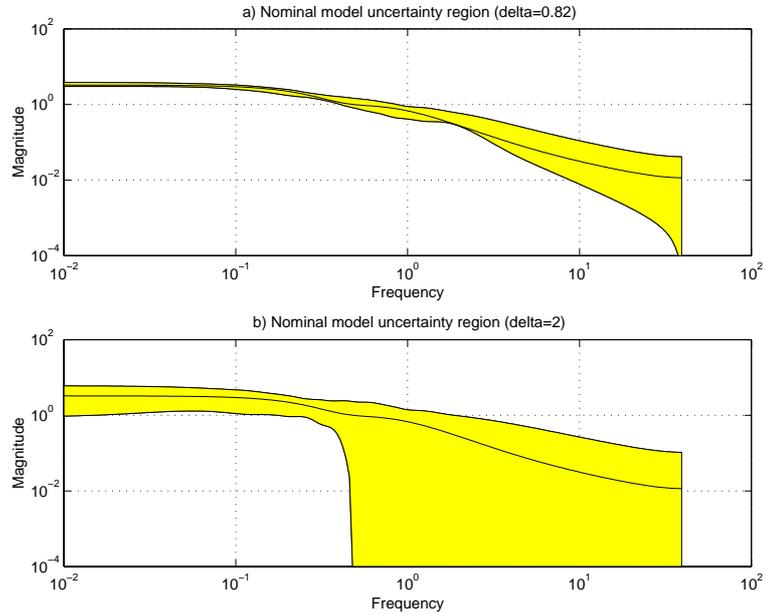
Figure 2: Nominal model uncertainty bands for a Laguerre expansion with pole $p = 0.96$. (a) $\delta = \delta^*$, (b) $\delta = 2$.

the considered nominal model structures (see Figs. 1 and 2).

6   Conclusions and Future Work

General model error modeling ideas and their application in the context of set membership identification have been analyzed. In particular, it has been shown that the separation of noise and unmodeled dynamics is quite natural in this framework, and that the minimum noise bound for which a nominal model is not falsified by the data can be easily computed.

A lot of work is still to be done, in several directions. For example, more complicated model error model structures are needed, to cope with nonlinearities or time-varying drifts. Obviously, this requires more sophisticated identification algorithms and smarter approximations in the computation of the uncertainty regions. Moreover, accurate criteria for the a priori selection of the structure of the model error model, depending on the specific control-oriented identification problem, must still be investigated.

References

[1] R. L. Kosut, G. C. Goodwin, and M. P. Polis (Eds.). Special issue on system identification for robust control design. *IEEE Transactions on Automatic Control*, 37:899–1008, 1992.

[2] T. Söderström and K. J. Aström (Eds.). Special issue on trends in system identification. *Automatica*, 31(12), 1995.

[3] A. Garulli, A. Tesi, and A. Vicino (Eds.). *Robustness in Identification and Control*. Lecture Notes in Control and Information Sciences. Springer, London, 1999.
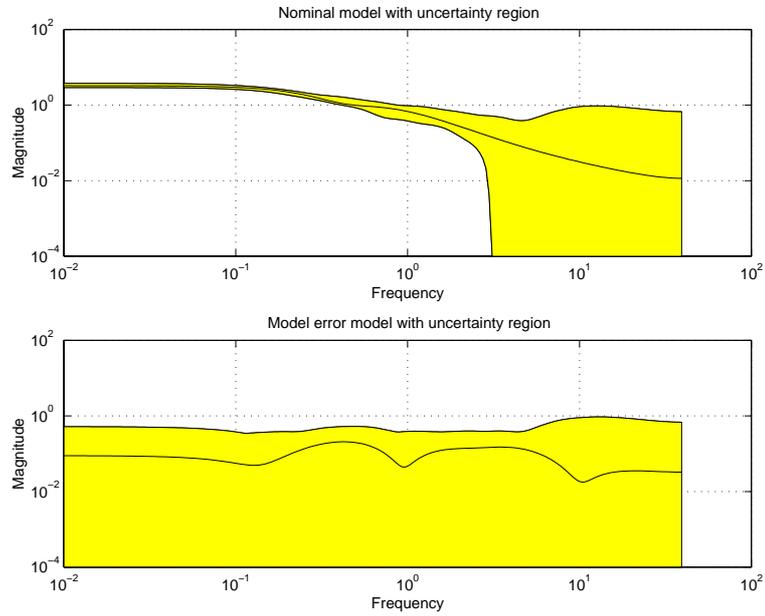
Figure 3: Nominal model: Laguerre expansion with pole $p = 0.96$; model error model: Kautz expansion with poles $p = 0.96$, 0.7 ($\delta = 1$).
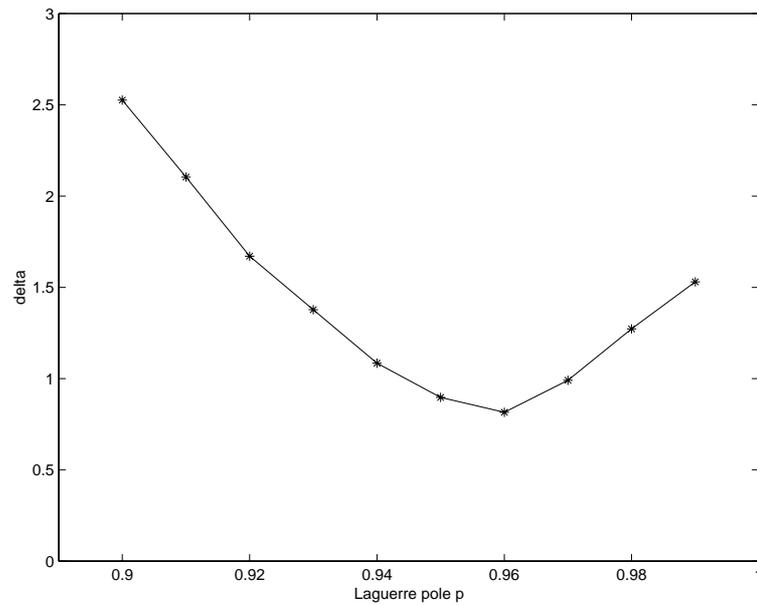

Figure 4: Different values of δ^* as a function of the Laguerre pole p.

[4] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*, 27(6):997–1009, 1991.

[5] M. Milanese, J. P. Norton, H. Piet-Lahanier, and E. Walter (eds.). *Bounding Approaches to System Identification*. Plenum Press, New York, 1996.

[6] E. Fogel and F. Huang. On the value of information in system identification - bounded noise case. *Automatica*, 18:229–238, 1982.

[7] J. P. Norton. Identification and application of bounded-parameter models. *Automatica*, 23:497–507, 1987.

[8] E. Walter and H. Piet-Lahanier. Estimation of parameter bounds from bounded-error data: a survey. *Mathematics in Computer and Simulation*, 32:449–468, 1990.

[9] B. Wahlberg and L. Ljung. Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Transactions on Automatic Control*, 37(7):900–912, 1992.

[10] R. L. Kosut, M. K. Lau, and S. P. Boyd. Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Transactions on Automatic Control*, 37(7):929–941, 1992.

[11] L. Giarrè, M. Milanese, and M. Taragna. $\mathcal{H}_\infty$ identification and model quality evaluation. *IEEE Transactions on Automatic Control*, 42(2):188–199, 1997.

[12] S. R. Venkatesh and M. A. Dahleh. Identification in the presence of classes of unmodeled dynamics and noise. *IEEE Transactions on Automatic Control*, 42(12):1620–1635, 1997.

[13] L. Ljung. Comments on model validation as set membership identification. In A. Garulli, A. Tesi, and A. Vicino, editors, *Robustness in Identification and Control*, pages 7–16. Springer, 1999.

[14] L. Ljung. Model validation and model error modeling. In *Proc. of the Åström Symposium on Control*, pages 15–42, Lund, Sweden, August 1999.

[15] B. Wahlberg. System identification using Laguerre models. *IEEE Transactions on Automatic Control*, 39:1276–1282, 1991.

[16] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor. System identification with generalized orthonormal basis functions. *Automatica*, 31(12):1821–1834, 1995.

[17] A. Garulli, A. Vicino, and G. Zappa. Conditional central algorithms for worst-case set membership identification and filtering. *IEEE Transactions on Automatic Control*, 1999. To appear.

[18] L. Chisci, A. Garulli, A. Vicino, and G. Zappa. Block recursive parallelotopic bounding in set membership identification. *Automatica*, 34(1):15–22, 1998.

[19] M. Milanese and M. Taragna. SM identification of model sets for robust control design from data. In A. Garulli, A. Tesi, and A. Vicino, editors, *Robustness in Identification and Control*, pages 17–34. Springer, 1999.

[20] A. Garulli, B. Z. Kacewicz, A. Vicino, and G. Zappa. Error bounds for conditional algorithms in restricted complexity set membership identification. *IEEE Transactions on Automatic Control*, 2000. To appear.

[21] G. C. Goodwin, J. H. Braslavsky, and M. M. Seron. Non-stationary stochastic embedding for transfer function estimation. In *Proc. of the 14th IFAC World Congress*, Bejing, China, July 1999.