

# ERROR BOUNDS FOR FIR MODELS IN CONDITIONAL SET-MEMBERSHIP IDENTIFICATION

Marco Casini, Andrea Garulli, Antonio Vicino

*DII - Università di Siena  
via Roma 56, 53100 Siena, Italy  
Email: {casini,garulli,vicino}@ing.unisi.it*

Abstract: The present paper deals with the problem of reduced complexity model estimation in the framework of conditional set-membership identification. The measurement noise is assumed to be unknown but bounded, while the estimated model quality is evaluated according to a worst-case criterion. Since optimal conditional estimators are generally hard to compute, projection estimators are often used in view of their better tractability from a complexity viewpoint. Tight bounds on the suboptimality level of central projection estimators as compared to optimal ones are derived for the case when FIR models are employed for approximation. These bounds improve over known bounds holding for the general class of linearly parameterized models. *Copyright ©2005 IFAC*

Keywords: Set-membership identification, bounded error, FIR models.

## 1. INTRODUCTION

The objective of robust identification is to provide a nominal model of an unknown plant, together with an estimate of the uncertainty associated to the model (see e.g. (Mäkilä *et al.*, 1995; Garulli *et al.*, 1999a) and references therein). Nominal models must belong to a pre-specified class of limited complexity models, while uncertainty is usually expressed in terms of norm bounds on the unmodeled dynamics.

Conditional Set-Membership Identification (CSMI) is a research line that falls into the above framework (Giarrè *et al.*, 1997; Kacewicz, 1999; Garulli *et al.*, 2000). Its main features are: i) the true plant belongs to the set of *feasible* systems, i.e. all systems compatible with input/output data and a priori knowledge; ii) the model lies in a linearly parameterized subspace; iii) the size of the unmodeled dynamics is measured in terms of the worst-case modelling error (in some norm) with respect to all feasible systems. Since optimal identification algorithms generally require the solution of non-convex min-max optimization problems, a typical challenge in CSMI is to devise suboptimal algorithms and to assess their performance by finding guaranteed bounds on their identification error. A popular class of suboptimal algorithms is that of *projection estimators*, in which the nominal model is selected by projecting a suitable system onto the model class. The worst-case suboptimality degree of projection estimators has been studied in (Garulli *et al.*, 1999b; Garulli, 1999): tight bounds

on the ratio between the identification error of projection estimators and the error provided by optimal algorithms have been derived, in the case of general feasible system sets and linearly parameterized model classes.

In this paper, it is shown how these bounds can be significantly improved when the class of FIR models is considered. The suboptimality degree turns out to be further improved when the feasible system set is balanced, i.e. it admits a symmetry center. All  $\ell_p$  norms,  $1 \leq p \leq \infty$ , are considered. Such results motivate the use of projection algorithms in CSMI, in place of highly computationally demanding optimal algorithms.

The paper is organized as follows. The CSMI framework, including optimal and suboptimal identification algorithms, is introduced in Section 2. Tight bounds on the suboptimality degree of central projection estimators for FIR model classes, are given in Section 3 in the case of balanced feasible system sets, and in Section 4 for general feasible system sets. Concluding remarks are provided in Section 5.

## 2. CONDITIONAL SET-MEMBERSHIP IDENTIFICATION

In this paper, LTI discrete-time SISO systems are considered. The impulse response of a system is denoted by  $h = \{h_i\}_{i=0}^{\infty}$ , and belongs to a linear normed space  $\mathcal{H}$ , equipped with the norm  $\|\cdot\|_{\mathcal{H}}$ . A priori knowledge on the system is expressed as  $h \in \mathcal{S}$ , where  $\mathcal{S}$  is a set contained in  $\mathcal{H}$ . Let

$\bar{h} \in \mathcal{S}$  be the true system generating the data. Data consists of a set of  $n$  input/output pairs  $z = \{(u_k, y_k), k = 1, \dots, n\}$ , related by

$$y_k = \sum_{i=1}^k \bar{h}_i u_{k-i} + e_k \quad (1)$$

where  $e_k$  is the disturbance affecting the measurement  $y_k$ . Equation (1) can be written in compact form as

$$y = T(u)\bar{h}^n + e \quad (2)$$

where  $y = [y_1 \dots y_n]'$ ,  $u = [u_1, \dots, u_n]'$ ,  $\bar{h}^n = [\bar{h}_1 \dots \bar{h}_n]'$ ,  $e = [e_1 \dots e_n]'$  and  $T(u)$  is the lower triangular Toeplitz matrix of inputs, i.e.

$$T(u) = \begin{bmatrix} u_1 & 0 & \dots & 0 \\ u_2 & u_1 & \dots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ u_n & \dots & u_2 & u_1 \end{bmatrix}. \quad (3)$$

It is assumed that  $e$  is unknown-but-bounded, i.e.

$$\|e\|_Y \leq \varepsilon \quad (4)$$

where  $\|\cdot\|_Y$  denotes a suitable norm in  $\mathbb{R}^n$  and  $\varepsilon$  is a known positive scalar. According to (2) and (4), the *feasible system set* is defined as

$$\mathcal{F} = \{h \in \mathcal{S} : \|y - T(u)h^n\|_Y \leq \varepsilon\}. \quad (5)$$

where  $h^n = [h_1, \dots, h_n]'$ .

Following the terminology of the Information-Based Complexity (IBC) theory (see (Milanese and Vicino, 1991)), an *identification algorithm*  $\phi$  is a mapping from the output set to the space of models. When the latter coincides with the space of systems  $\mathcal{H}$ , one has:  $\phi : \mathbb{R}^n \rightarrow \mathcal{H}$ . The *worst-case identification error* associated to the identification algorithm turns out to be

$$E[\phi] = \sup_{h \in \mathcal{F}} \|h - \phi(y)\|_{\mathcal{H}}.$$

In robust identification, it is customary to identify reduced-complexity models that are suitable for robust control design techniques. A typical choice is to select a linearly parameterized model class such as  $\mathcal{M} = \{g \in \mathcal{H} : g = M\theta, \theta \in \mathbb{R}^m\}$ , where  $M$  is a linear operator,  $M : \mathbb{R}^m \rightarrow \mathcal{H}$  and  $\theta$  is the  $m$ -dimensional parameter vector to be identified,  $m < n$ . A *conditional identification algorithm* is defined as a mapping  $\phi : \mathbb{R}^n \rightarrow \mathcal{M}$ . The model class  $\mathcal{M}$  can be chosen as a collection of impulse responses of linear filters, such as FIRs, Laguerre or Kautz functions (Wahlberg, 1991; Wahlberg, 1994), or Generalized Orthonormal Basis Functions (Van den Hof *et al.*, 1995).

The identification of a model within the class  $\mathcal{M}$  has been addressed in the literature as *conditional set-membership identification* (Giarrè *et al.*, 1997; Kacewicz, 1999; Garulli *et al.*, 2000). In particular, the optimal model is given by the *conditional central algorithm*, which minimizes the worst-case error among the elements of  $\mathcal{M}$ , i.e.

$$\phi_{cc}(y) = \arg \inf_{g \in \mathcal{M}} \sup_{h \in \mathcal{F}} \|h - g\|_{\mathcal{H}}. \quad (6)$$

The min-max optimization problem in (6) is generally very difficult to solve. A procedure for computing  $\phi_{cc}(y)$  has been given in (Garulli *et al.*, 2000) for the case  $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_Y = \|\cdot\|_2$ , while suboptimal identification algorithms have been considered in (Garulli, 1999).

A typical suboptimal algorithm is the *central projection algorithm* given by

$$\phi_{cp}(y) = \arg \inf_{g \in \mathcal{M}} \|g - c\|_{\mathcal{H}}, \quad (7)$$

where  $c$  denotes the Chebyshev center of  $\mathcal{F}$ ,

$$c = \inf_{x \in \mathcal{H}} \sup_{h \in \mathcal{F}} \|x - h\|_{\mathcal{H}}. \quad (8)$$

For easiness of presentation, in the following we set  $\mathcal{H} = \mathbb{R}^n$  and consider  $\mathcal{S}$  and  $\mathcal{M}$  as subsets of  $\mathbb{R}^n$ . This corresponds to assuming that the data set length  $n$  is such that the impulse response samples  $\bar{h}_{n+1}, \bar{h}_{n+2}, \dots$  can be neglected. This assumption can usually be relaxed by making suitable hypotheses on the tail of the system impulse response and/or on the structure of the model class  $\mathcal{M}$  (see e.g. (Garulli *et al.*, 2000)).

In this paper we consider FIR models, i.e.  $g \in \mathcal{H} : g_i = 0, i > m$  (this corresponds to choosing  $M = [I_{m \times m} \ 0_{m \times n}]'$ ). For this case, tight bounds on the suboptimality degree of  $\phi_{cp}$  with respect to  $\phi_{cc}$  are derived for both balanced and non-balanced feasible sets, and for different  $\mathcal{H}$ -norms.

### 3. SUBOPTIMALITY BOUNDS FOR BALANCED FEASIBLE SYSTEM SETS

The following assumptions are made.

*Assumption 1.* Let  $\mathcal{M}$  be an  $m$ -dimensional linear manifold in  $\mathbb{R}^n$ , with  $m < n$ , such that  $\mathcal{M} = \{g \in \mathbb{R}^n : g = [g_1, \dots, g_m, 0, \dots, 0]', g_i \in \mathbb{R}\}$ .

*Assumption 2.* Let the feasible set  $\mathcal{F}$  be *balanced*, i.e. there exists  $c : c + v \in \mathcal{F}$  implies  $c - v \in \mathcal{F}$ , where  $c$  represents the symmetry center of  $\mathcal{F}$ .

*Remark 1.* Notice that the set  $\mathcal{V} = \{h \in \mathcal{H} : \|y - T(u)h^n\|_Y \leq \varepsilon\}$  is balanced for any norm  $Y$ . Moreover, the a priori set  $\mathcal{S}$  is usually assumed to be balanced. Being  $\mathcal{F} = \mathcal{S} \cap \mathcal{V}$ , Assumption 2 is verified for example if  $\mathcal{V} \subseteq \mathcal{S}$  (the so-called *residual prior*, i.e. no a priori constraint is given on the first  $n$  samples of the system impulse response).

In the following we denote by  $\|\cdot\|_p$  any  $\ell_p$  norm,  $1 \leq p \leq \infty$ . Let us define the following quantities:

$$\hat{c} \triangleq \arg \inf_{g \in \mathcal{M}} \|g - c\|_p, \quad (9)$$

$$\hat{h} \triangleq \arg \sup_{h \in \mathcal{F}} \|\hat{c} - h\|_p, \quad (10)$$

$$\tilde{g} \triangleq \phi_{cc}(y) = \arg \inf_{g \in \mathcal{M}} \sup_{h \in \mathcal{F}} \|g - h\|_p, \quad (11)$$

$$\hat{E}_p(\mathcal{M}, \mathcal{F}) \triangleq E[\phi_{cp}(y)] = \|\hat{c} - \hat{h}\|_p, \quad (12)$$

$$\tilde{E}_p(\mathcal{M}, \mathcal{F}) \triangleq E[\phi_{cc}(y)] = \sup_{h \in \mathcal{F}} \|\tilde{g} - h\|_p. \quad (13)$$

Note that  $\hat{c}$  is the projection of  $c$  on  $\mathcal{M}$ ,  $\hat{h}$  is the farthest point in  $\mathcal{F}$  from  $\hat{c}$ ,  $\tilde{g}$  is the conditional center, and  $\hat{E}_p(\mathcal{M}, \mathcal{F})$  and  $\tilde{E}_p(\mathcal{M}, \mathcal{F})$  are

the central projection error and the conditional error, respectively. Moreover, let  $\hat{k} = 2c - \hat{h}$  (the symmetric of  $\hat{h}$  w.r.t.  $c$ ). The aim of this section is to evaluate the maximum ratio between  $\hat{E}_p$  and  $\tilde{E}_p$ , for every linear subspace  $\mathcal{M}$  and  $\mathcal{F}$  as in Assumptions 1-2, i.e.

$$\mu_p = \max_{\mathcal{M}, \mathcal{F}} \frac{\hat{E}_p(\mathcal{M}, \mathcal{F})}{\tilde{E}_p(\mathcal{M}, \mathcal{F})}. \quad (14)$$

Let us now reformulate the problem in a simpler way by choosing an appropriate coordinate frame. Let  $c = [c_1, \dots, c_n]'$  be the center of symmetry of  $\mathcal{F}$ . Without loss of generality, let us choose the coordinate system so that  $c_i = 0$ ,  $1 \leq i \leq m$ . It follows that, for any norm  $p$ ,  $\hat{c} = [0, \dots, 0]'$  and

$$\hat{E}_p = \|\hat{h} - \hat{c}\|_p = \|\hat{h}\|_p. \quad (15)$$

Moreover,  $\hat{k} = [-\hat{h}_1, \dots, -\hat{h}_m, \hat{k}_{m+1}, \dots, \hat{k}_n]'$ , where  $\hat{k}_i = 2c_i - \hat{h}_i$ ,  $m < i \leq n$ . Let us first analyze the case  $p = \infty$ .

*Proposition 1.* Let  $\mathcal{F}$  be any set. Then, for every  $\mathcal{M}$  as in Assumption 1, one has

$$\hat{E}_\infty(\mathcal{M}, \mathcal{F}) = \tilde{E}_\infty(\mathcal{M}, \mathcal{F}).$$

*Proof:* Let  $a$  be any point. It is well known that  $\sup_{h \in \mathcal{F}} \|h - a\|_\infty = \sup_{h \in \text{BOX}(\mathcal{F})} \|h - a\|_\infty$  where  $\text{BOX}(\mathcal{F})$  denotes the minimum box containing  $\mathcal{F}$ . Let us choose the center of  $\mathcal{F}$  as the symmetry center of  $\text{BOX}(\mathcal{F})$ ; notice that  $\text{BOX}(\mathcal{F})$  is a convex and symmetric set and thus satisfies Assumption 2. By (15),  $\hat{E}_\infty = \|\hat{h}\|_\infty = \max_{i=1}^n |\hat{h}_i|$ . Let us consider the case that  $\max_{i=1}^n |\hat{h}_i| = |\hat{h}_s|$ ,  $s > m$ . Since  $\tilde{g} \in \mathcal{M}$ ,  $\tilde{g} = [\tilde{g}_1, \dots, \tilde{g}_m, 0, \dots, 0]'$ , and  $\tilde{E}_\infty \geq \|\hat{h} - \tilde{g}\|_\infty = \max\{\max_{i \leq m} |\hat{h}_i - \tilde{g}_i|, |\hat{h}_s|\} \geq |\hat{h}_s|$ . Since  $\tilde{E}_\infty$  can not be greater than  $\hat{E}_\infty$ ,  $\tilde{E}_\infty = \hat{E}_\infty$ .

Let us now assume  $\max_{i=1}^n |\hat{h}_i| = |\hat{h}_d|$ ,  $1 \leq d \leq m$ . Since  $\hat{k} = [-\hat{h}_1, \dots, -\hat{h}_m, \hat{k}_{m+1}, \dots, \hat{k}_n]'$ , it follows that  $\hat{E}_\infty = \|\hat{h}\|_\infty = \|\hat{k}\|_\infty = |\hat{h}_d|$  and

$$\begin{aligned} \tilde{E}_\infty &\geq \max\{\|\tilde{g} - \hat{h}\|_\infty, \|\tilde{g} - \hat{k}\|_\infty\} \\ &\geq \max\{|\tilde{g}_d - \hat{h}_d|, |\tilde{g}_d - \hat{k}_d|\} \\ &= \max\{|\tilde{g}_d - \hat{h}_d|, |\tilde{g}_d + \hat{h}_d|\} \geq |\hat{h}_d|. \end{aligned}$$

Again  $\tilde{E}_\infty$  can not be greater than  $\hat{E}_\infty$ , and then  $\tilde{E}_\infty = \hat{E}_\infty$ .  $\square$

*Remark 2.* Proposition 1 states that, when the  $\ell_\infty$ -norm is used, the conditional center coincides with the central projection in  $\ell_2$ -norm. Moreover, this holds also when  $\mathcal{F}$  is a generic set, not necessarily balanced. Since the center of  $\mathcal{F}$  may not be unique, we choose the center of  $\mathcal{F}$  as the symmetry center of  $\text{BOX}(\mathcal{F})$ .

In order to state the main result concerning  $\ell_p$  norms,  $1 \leq p < \infty$ , some lemmas are needed.

*Lemma 1.* Let  $\mathcal{M}$  and  $\mathcal{F}$  be given, and define  $\hat{\mathcal{F}} = \{\hat{h}, \hat{k}\}$ . Then,

$$\frac{\hat{E}_p(\mathcal{M}, \mathcal{F})}{\tilde{E}_p(\mathcal{M}, \mathcal{F})} \leq \frac{\hat{E}_p(\mathcal{M}, \hat{\mathcal{F}})}{\tilde{E}_p(\mathcal{M}, \hat{\mathcal{F}})}.$$

*Proof:* Note that  $\hat{\mathcal{F}} \subseteq \mathcal{F}$ , due to Assumption 2, and hence  $\hat{E}_p(\mathcal{M}, \hat{\mathcal{F}}) \leq \hat{E}_p(\mathcal{M}, \mathcal{F})$ . Moreover, being  $\hat{h} \in \hat{\mathcal{F}}$ ,  $\tilde{E}_p(\mathcal{M}, \hat{\mathcal{F}}) = \tilde{E}_p(\mathcal{M}, \mathcal{F})$ .  $\square$

An immediate consequence of Lemma 1 is that we can replace  $\mathcal{F}$  by the set  $\hat{\mathcal{F}} = \{\hat{h}, \hat{k}\}$ .

*Lemma 2.* Let  $\mathcal{F} = \hat{\mathcal{F}}$ . Under Assumption 1,  $\tilde{g}_i = \beta_i \hat{h}_i$ ,  $\forall i \leq m$ ,  $|\beta_i| \leq 1$ .

*Proof:* By contradiction, let us assume  $\tilde{g} : \tilde{g}_s = \gamma \hat{h}_s$ ,  $s \leq m$  and  $|\gamma| > 1$ . Then,

$$\begin{aligned} \|\tilde{g} - \hat{h}\|_p^p &= \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p + |\gamma \hat{h}_s - \hat{h}_s|^p \\ &= \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p + |\gamma - 1|^p |\hat{h}_s|^p, \end{aligned}$$

$$\begin{aligned} \|\tilde{g} - \hat{k}\|_p^p &= \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p + |\gamma \hat{h}_s + \hat{h}_s|^p \\ &= \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p + |\gamma + 1|^p |\hat{h}_s|^p. \end{aligned}$$

Let us define  $\bar{g} : \bar{g}_i = \tilde{g}_i$ ,  $i \neq s$ ,  $\bar{g}_s = \delta \hat{h}_s$ ,  $\delta = \text{sgn}(\gamma)$ . Then,

$$\|\bar{g} - \hat{h}\|_p^p = \begin{cases} \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p & \text{if } \gamma > 0 \\ \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p + 2^p |\hat{h}_s|^p & \text{if } \gamma < 0 \end{cases},$$

$$\|\bar{g} - \hat{k}\|_p^p = \begin{cases} \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p + 2^p |\hat{h}_s|^p & \text{if } \gamma > 0 \\ \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p & \text{if } \gamma < 0 \end{cases}.$$

Therefore,  $\tilde{g}$  cannot be the conditional center.  $\square$

*Lemma 3.* Let  $\mathcal{F} = \hat{\mathcal{F}}$ . Under Assumption 1, if  $\|\tilde{g} - \hat{h}\|_p > \|\tilde{g} - \hat{k}\|_p$  then  $\tilde{g}_i = \hat{h}_i$ ,  $i \leq m$ .

*Proof:* Let  $\|\tilde{g} - \hat{h}\|_p^p - \|\tilde{g} - \hat{k}\|_p^p = \nu$ ,  $\nu > 0$ . By contradiction let us assume  $\exists s$ ,  $s \leq m$  :  $\tilde{g}_s \neq \hat{h}_s$ . Then,

$$\begin{aligned} \tilde{E}_p^p &= |\tilde{g}_s - \hat{h}_s|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p \\ &= |\hat{h}_s + \tilde{g}_s|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p + \nu. \end{aligned} \quad (16)$$

Let us assume  $\hat{h}_s > 0$ . By Lemma 2, one has  $\hat{h}_s \geq \tilde{g}_s$ . Moreover, notice that  $\tilde{g}_s \geq 0$ , otherwise  $\tilde{g}$  is no more the conditional center. Indeed, if  $\tilde{g}_s < 0$  it is possible to define  $\tilde{t} : \tilde{t}_i = \tilde{g}_i$ ,  $i \neq s$ ,  $\tilde{t}_s = \tilde{g}_s + \varepsilon$ ,  $\varepsilon > 0$  sufficiently small, such that

$$\|\tilde{t} - \hat{k}\|_p \leq \|\tilde{t} - \hat{h}\|_p < \|\tilde{g} - \hat{h}\|_p.$$

Let  $\bar{g} : \bar{g}_i = \tilde{g}_i$ ,  $i \neq s$ ,  $\bar{g}_s = \tilde{g}_s + \gamma$ ,  $\gamma > 0$ , and  $\bar{E}_p = \sup_{h \in \mathcal{F}} \|\bar{g} - h\|_p = \max\{\bar{E}_{p1}, \bar{E}_{p2}\}$ , where,

$$\bar{E}_{p1}^p = |\hat{h}_s - \tilde{g}_s - \gamma|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p, \quad (17)$$

$$\bar{E}_{p2}^p = |\hat{h}_s + \tilde{g}_s + \gamma|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p. \quad (18)$$

Let  $\bar{\gamma} > 0$  such that  $|\hat{h}_s + \tilde{g}_s + \bar{\gamma}|^p = |\hat{h}_s + \tilde{g}_s|^p + \nu$ , and choose  $\gamma < \min\{(\hat{h}_s - \tilde{g}_s), \bar{\gamma}\}$ . By comparing (17)-(18) with (16) one has

$$\tilde{E}_p^p > \bar{E}_p^p = \max\{\bar{E}_{p1}^p, \bar{E}_{p2}^p\},$$

and then  $\tilde{g}$  is not the conditional center, which contradicts the hypothesis.

It is straightforward to repeat the previous reasoning for  $\hat{h}_s < 0$ .  $\square$

*Lemma 4.* Let  $\mathcal{F} = \hat{\mathcal{F}}$ . Under Assumption 1,  

$$\tilde{E}_p = \|\tilde{g} - \hat{h}\|_p. \quad (19)$$

*Proof:* Since  $\tilde{E}_p = \max\{\|\tilde{g} - \hat{h}\|_p, \|\tilde{g} - \hat{k}\|_p\}$ , by contradiction let us suppose that

$\tilde{E}_p^p = \|\tilde{g} - \hat{k}\|_p^p = \|\tilde{g} - \hat{h}\|_p^p + \nu$ ,  $\nu > 0$ , that is

$$\begin{aligned} \tilde{E}_p^p &= \sum_{i=1}^m |\tilde{g}_i + \hat{h}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p \\ &= \sum_{i=1}^m |\tilde{g}_i - \hat{h}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p + \nu. \end{aligned}$$

Since by (10) and (15),  $\sum_{i=m+1}^n |\hat{h}_i|^p \geq \sum_{i=m+1}^n |\hat{k}_i|^p$ ,

this means that  $\exists s, 1 \leq s \leq m : |\tilde{g}_s + \hat{h}_s| > |\tilde{g}_s - \hat{h}_s|$ . Then,

$$\begin{aligned} \tilde{E}_p^p &= |\tilde{g}_s + \hat{h}_s|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p \\ &= |\hat{h}_s - \tilde{g}_s|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p + \nu. \quad (20) \end{aligned}$$

So,  $\tilde{g}_s$  and  $\tilde{h}_s$  have the same sign. To simplify the proof, let us assume  $\tilde{g}_s > 0$ . By Lemma 2, one has  $\hat{h}_s \geq \tilde{g}_s > 0$ .

Let us define  $\bar{g} : \bar{g}_i = \tilde{g}_i, i \neq s, \bar{g}_s = \tilde{g}_s - \gamma, \gamma > 0$ , and  $\bar{E}_p = \sup_{h \in \mathcal{F}} \|\bar{g} - h\|_p = \max\{\bar{E}_{p1}, \bar{E}_{p2}\}$ , where

$$\bar{E}_{p1}^p = |\bar{g}_s + \hat{h}_s - \gamma|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{k}_i|^p, \quad (21)$$

$$\bar{E}_{p2}^p = |\hat{h}_s - \tilde{g}_s + \gamma|^p + \sum_{i \neq s} |\tilde{g}_i - \hat{h}_i|^p. \quad (22)$$

Let  $\bar{\gamma} > 0$  such that  $|\hat{h}_s - \tilde{g}_s + \bar{\gamma}|^p = |\hat{h}_s - \tilde{g}_s|^p + \nu$ , and choose  $\gamma < \min\{(\tilde{g}_s + \hat{h}_s), \bar{\gamma}\}$ . By comparing (21)-(22) with (20) one has

$$\tilde{E}_p^p > \bar{E}_p^p = \max\{\bar{E}_{p1}^p, \bar{E}_{p2}^p\},$$

and then  $\tilde{g}$  is not the conditional center, which contradicts the hypothesis.

It is straightforward to repeat the previous reasoning for  $\tilde{g}_s < 0$ .  $\square$

Notice that the previous lemma implies

$$\tilde{E}_p = \|\hat{h} - \tilde{g}\|_p^p \geq \|\hat{k} - \tilde{g}\|_p^p.$$

*Lemma 5.* Let  $\mathcal{F} = \hat{\mathcal{F}}$ . Under Assumption 1, there always exists an optimal solution of (11) such that  $\tilde{g}_i = \beta \hat{h}_i, \forall i \leq m$ , and  $0 \leq \beta \leq 1$ .

*Proof:* Let us suppose  $\|\tilde{g} - \hat{h}\|_p > \|\tilde{g} - \hat{k}\|_p$ . By Lemma 3,  $\tilde{g}_i = \hat{h}_i, i \leq m$ , and thus  $\beta = 1$ . Thanks to Lemma 4, it remains to consider only the case  $\|\tilde{g} - \hat{h}\|_p = \|\tilde{g} - \hat{k}\|_p$ . By Lemma 2, one gets

$$\begin{aligned} \sum_{i=m+1}^n |\hat{h}_i|^p &\leq \|\tilde{g} - \hat{k}\|_p^p = \sum_{i=1}^m |\tilde{g}_i + \hat{h}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p \\ &= \sum_{i=1}^m |(\beta_i + 1)\hat{h}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p \end{aligned}$$

$$\leq \sum_{i=1}^m 2|\hat{h}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p,$$

that is,

$$\sum_{i=m+1}^n |\hat{h}_i|^p - \sum_{i=m+1}^n |\hat{k}_i|^p \leq 2^p \sum_{i=1}^m |\hat{h}_i|^p. \quad (23)$$

One has,

$$\tilde{E}_p^p = \sum_{i=1}^m |\tilde{g}_i - \hat{h}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p = \sum_{i=1}^m |\tilde{g}_i + \hat{h}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p$$

that is,

$$\begin{aligned} &\sum_{i=1}^m |\tilde{g}_i + \hat{h}_i|^p - \sum_{i=1}^m |\tilde{g}_i - \hat{h}_i|^p = \\ &= \sum_{i=m+1}^n |\hat{h}_i|^p - \sum_{i=m+1}^n |\hat{k}_i|^p \triangleq \gamma. \quad (24) \end{aligned}$$

From (10) and (15), it is easy to see that  $\gamma \geq 0$ .

Let  $\tilde{g}_i = \beta \hat{h}_i, 0 \leq \beta \leq 1$ . By substituting in (24),

$$\begin{aligned} &\sum_{i=1}^m |\beta \hat{h}_i + \hat{h}_i|^p - \sum_{i=1}^m |\beta \hat{h}_i - \hat{h}_i|^p = \\ &= (\beta + 1)^p \sum_{i=1}^m |\hat{h}_i|^p - (1 - \beta)^p \sum_{i=1}^m |\hat{h}_i|^p = \gamma, \end{aligned}$$

i.e., by (23),

$$\begin{aligned} 0 &\leq (\beta + 1)^p - (1 - \beta)^p = \gamma \left( \sum_{i=1}^m |\hat{h}_i|^p \right)^{-1} \\ &= \left( \sum_{i=m+1}^n |\hat{h}_i|^p - \sum_{i=m+1}^n |\hat{k}_i|^p \right) \left( \sum_{i=1}^m |\hat{h}_i|^p \right)^{-1} \leq 2^p. \quad (25) \end{aligned}$$

It remains to prove that there always exists  $0 \leq \beta \leq 1$  such that (25) is achieved. Let us define  $f(\beta) \triangleq (\beta + 1)^p - (1 - \beta)^p$ , one has  $f(0) = 0$ ,  $f(1) = 2^p$ , and since  $f(\beta)$  is a continuous function,  $\exists \beta, 0 \leq \beta \leq 1 : f(\beta) = \delta, 0 \leq \delta \leq 2^p$ , and the lemma is proved.  $\square$

*Remark 3.* Note that when  $\beta = 0$ , the conditional center coincides with the central projection, giving  $\hat{E}_p/\tilde{E}_p = 1$ . Since we want to maximize this ratio, in the next we will consider  $0 < \beta \leq 1$ .

It is now possible to state the main result.

*Theorem 1.* Let Assumptions 1-2 hold. Let  $1 \leq p < \infty$ . Then

$$\mu_p \leq \begin{cases} 3/2 & \text{if } p = 1, \\ \sqrt[p]{1 + \frac{1}{(2^{p-\sqrt{2}} - 1)^{p-1}}} & \text{if } 1 < p < \infty. \end{cases}$$

Moreover, the bounds are tight.

*Proof:* Due to Lemma 1, we can consider the feasible set  $\hat{\mathcal{F}} = \{\hat{h}, \hat{k}\}$ . From Lemma 4, one has

$$\sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p \geq \sum_{i=1}^m |\hat{h}_i + \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p. \quad (26)$$

By (15) and (19), it follows that

$$\begin{aligned}
\frac{\widehat{E}_p^p}{\widetilde{E}_p^p} &= \frac{\sum_{i=1}^m |\hat{h}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p}{\sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p} \\
&= 1 + \frac{\sum_{i=1}^m |\hat{h}_i|^p - \sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p}{\sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p} \\
&\leq 1 + \frac{\sum_{i=1}^m |\hat{h}_i|^p - \sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p}{\sum_{i=1}^m |\hat{h}_i + \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{k}_i|^p} \quad (27) \\
&\leq 1 + \frac{\sum_{i=1}^m |\hat{h}_i|^p - \sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p}{\sum_{i=1}^m |\hat{h}_i + \tilde{g}_i|^p} \quad (28)
\end{aligned}$$

where the first inequality follows from (26). By Lemma 5,  $\hat{h}_i = \alpha \tilde{g}_i$ ,  $\forall i \leq m$ ,  $\alpha \geq 1$ , and since  $\|\tilde{g}\|_p^p > 0$ , one has

$$\begin{aligned}
\mu_p^p &\leq 1 + \frac{\sum_{i=1}^m |\alpha \tilde{g}_i|^p - \sum_{i=1}^m |\alpha \tilde{g}_i - \tilde{g}_i|^p}{\sum_{i=1}^m |\alpha \tilde{g}_i + \tilde{g}_i|^p} \\
&= 1 + \frac{\alpha^p \sum_{i=1}^m |\tilde{g}_i|^p - (\alpha - 1)^p \sum_{i=1}^m |\tilde{g}_i|^p}{(\alpha + 1)^p \sum_{i=1}^m |\tilde{g}_i|^p} \\
&= 1 + \frac{\alpha^p - (\alpha - 1)^p}{(\alpha + 1)^p}. \quad (29)
\end{aligned}$$

Let us define  $f(\alpha) \triangleq 1 + \frac{\alpha^p - (\alpha - 1)^p}{(\alpha + 1)^p}$ . We want to maximize  $f(\alpha)$  for  $\alpha \geq 1$ . Denoting by  $\alpha_{max}$  the value for which  $f(\alpha)$  is maximum, one gets

$$\alpha_{max} = \begin{cases} 1 & \text{if } p = 1, \\ \frac{p^{-1}\sqrt{2}}{p^{-1}\sqrt{2} - 1} & \text{if } 1 < p < \infty. \end{cases}$$

Substituting  $\alpha_{max}$  in (29), after some algebra,

$$\mu_p \leq \begin{cases} 3/2 & \text{if } p = 1, \\ \sqrt[p]{1 + \frac{1}{(2^{p-1}\sqrt{2} - 1)^{p-1}}} & \text{if } 1 < p < \infty. \end{cases}$$

It remains to prove that such bounds are tight. This can be done by choosing  $\hat{h}$  and  $\hat{k}$  such that the inequalities in (27)-(28) become equalities, that is  $\hat{k}_i = 0$ ,  $m < i \leq n$ , and

$$\sum_{i=1}^m |\hat{h}_i - \tilde{g}_i|^p + \sum_{i=m+1}^n |\hat{h}_i|^p = \sum_{i=1}^m |\hat{h}_i + \tilde{g}_i|^p. \quad (30)$$

The upper bound is achieved when  $\hat{h}_i = \alpha_{max} \tilde{g}_i$ ,  $1 \leq i \leq m$ . After some algebra, (30) becomes

$$\sum_{i=m+1}^n |\hat{h}_i|^p = \frac{(\alpha_{max} + 1)^p - (\alpha_{max} - 1)^p}{\alpha_{max}^p} \sum_{i=1}^m |\hat{h}_i|^p. \quad (31)$$

Table 1. Upper bound (approximated) on  $\mu_p$  for balanced sets.

$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$\dots$	$p = \infty$
$\mu_p$	1.5	$\simeq 1.15$	$\simeq 1.09$	$\simeq 1.06$	$\simeq 1.05$	$\dots$ 1

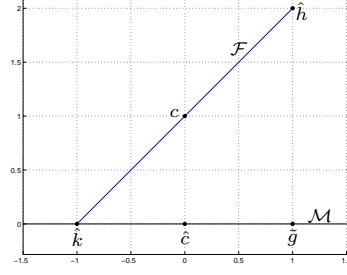


Fig. 1. Example 1:  $\mu_1 = 3/2$ .

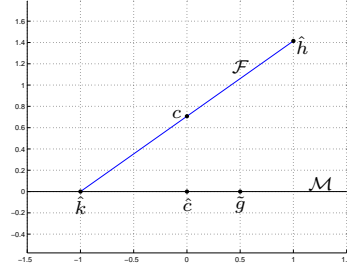


Fig. 2. Example 2:  $\mu_2 = \sqrt{4/3}$ .

Since  $\hat{h}_i$ ,  $i > m$ , can be arbitrarily chosen, they can be selected so that (31) is satisfied.  $\square$

Table 1 reports upper bounds on  $\mu_p$  for different norms.

In the following, two examples showing cases where the upper bound is achieved are reported for  $p = 1$  and  $p = 2$ .

*Example 1.* Let  $p = 1$ ,  $n = 2$ ,  $m = 1$ ,  $c = [0, 1]'$ ,  $\hat{h} = [1, 2]'$ . It follows that  $\hat{k} = [-1, 0]'$  and  $\tilde{g} = [1, 0]'$ . Moreover, one has  $\widehat{E} = 3$  and  $\widetilde{E} = 2$  and thus  $\mu_1 = 3/2$  (see Fig. 1).

*Example 2.* Let  $p = 2$ ,  $n = 2$ ,  $m = 1$ ,  $c = [0, \sqrt{2}/2]'$ ,  $\hat{h} = [1, \sqrt{2}]'$ . It follows that  $\hat{k} = [-1, 0]'$  and  $\tilde{g} = [1/2, 0]'$ . One has  $\widehat{E} = \sqrt{3}$  and  $\widetilde{E} = 3/2$  and thus  $\mu_2 = \sqrt{4/3}$  (see Fig. 2).

#### 4. SUBOPTIMALITY BOUNDS FOR GENERAL FEASIBLE SYSTEM SETS

In this section, bounds similar to those previously found are given for a generic set, not necessarily convex and symmetric. In particular, we focus on the  $\ell_1, \ell_2, \ell_\infty$ -norms.

In (Garulli *et al.*, 1999b) it has been shown that, for a generic set and for any linear  $\mathcal{M}$  (not necessarily obtained by FIRs),  $\mu_2 = \sqrt{4/3}$ . Since this result coincides with that reported in Table 1 regarding balanced sets and  $\mathcal{M}$  FIR, one can conclude that such a bound also holds for generic sets. Since Proposition 1 holds for a generic  $\mathcal{F}$ , one has  $\mu_\infty = 1$ . In the following we analyze the case  $p = 1$ .

*Theorem 2.* Let  $\mathcal{F} \subset \mathbb{R}^n$  be any set and let  $\mathcal{M}$  satisfy Assumption 1. Then,  $\mu_1 \leq 2$ .

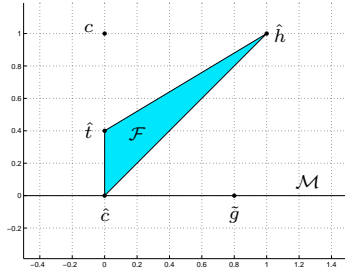


Fig. 3. Example 3:  $\mu_1 \rightarrow 2$  as  $\alpha \rightarrow 0$ .

*Proof:* Let us assume the same context considered in the previous section, except for the fact that the center may not be the symmetry center (when  $\mathcal{F}$  is balanced). Then,

$$\widehat{E}_1 = \|\hat{h}\|_1 = \sum_{i=1}^m |\hat{h}_i| + \sum_{i=m+1}^n |\hat{h}_i|. \quad (32)$$

Moreover,

$$\widetilde{E}_1 \geq \|\hat{h} - \tilde{g}\|_1 = \sum_{i=1}^m |\hat{h}_i - \tilde{g}_i| + \sum_{i=m+1}^n |\hat{h}_i| \geq \sum_{i=m+1}^n |\hat{h}_i|,$$

$$r \geq \|\hat{h} - c\|_1 = \sum_{i=1}^m |\hat{h}_i| + \sum_{i=m+1}^n |\hat{h}_i - c_i| \geq \sum_{i=1}^m |\hat{h}_i|.$$

Thus,

$$\widetilde{E}_1 + r \geq \sum_{i=1}^m |\hat{h}_i| + \sum_{i=m+1}^n |\hat{h}_i| = \widehat{E}_1,$$

i.e., since  $\widetilde{E}_1 \geq r$ ,

$$\mu_1 \leq \frac{\widetilde{E}_1 + r}{\widetilde{E}_1} = 1 + \frac{r}{\widetilde{E}_1} \leq 1 + 1 = 2. \quad \square$$

Table 2 summarizes the error bounds of central projection estimators for different norms, model classes and system feasible sets (see (Garulli *et al.*, 1999b; Garulli, 1999)). In the following, an example which shows a case when the upper bound of Theorem 2 is achieved is reported.

*Example 3.* Let  $p=1$ ,  $n=2$ ,  $m=1$  and  $\mathcal{F}$  be the triangle of vertices  $\hat{c}$ ,  $\hat{h}$ ,  $\hat{t}$ , where  $\hat{c} = [0, 0]'$ ,  $\hat{h} = [1, 1]'$  and  $\hat{t} = [0, \alpha]'$ ,  $0 \leq \alpha \leq 1$  (see Fig. 3). Notice that  $\mathcal{F}$  is not balanced and there are different possible choices of its Chebyshev center, according to definition (8). Let us choose  $c = [0, 1]'$ . Then,  $\widehat{E} = 2$ . Moreover, it is easy to show that the conditional center is  $\tilde{g} = [1 - \frac{\alpha}{2}, 0]'$  and the conditional error  $\widetilde{E} = \|\tilde{g} - \hat{h}\|_1 = \|\tilde{g} - \hat{t}\|_1 = 1 + \frac{\alpha}{2}$ . Thus,  $\mu_1 = \frac{4}{2+\alpha}$ , and  $\mu_1 \rightarrow 2$  when  $\alpha \rightarrow 0$ .

*Remark 4.* Note that the bound of Theorem 2 differs from that derived in Theorem 1 because

Table 2. Upper bound on  $\mu_p$  for  $\ell_1, \ell_2$  and  $\ell_\infty$ -norms.

	$\ell_1$	$\ell_2$	$\ell_\infty$
Generic $\mathcal{M}, \mathcal{F}$	3	$\sqrt{4/3}$	2
$\mathcal{M}$ FIR, generic $\mathcal{F}$	2	$\sqrt{4/3}$	1
$\mathcal{M}$ FIR, $\mathcal{F}$ balanced	3/2	$\sqrt{4/3}$	1

in the case of unbalanced sets the choice of the Chebyshev center is no more trivial. Since the center may not be unique, Theorem 2 provides the upper bound for the worst set and for the worst choice of the center.

## 5. CONCLUSIONS

In this paper tight bounds on the suboptimality level of central projection estimators with respect to optimal estimators are derived, for the case when the class of FIR models is used in conditional set-membership identification. These bounds improve over known bounds holding for the larger class of linearly parameterized models. Some numerical examples showing tightness of the derived bounds are also reported. The study of more strict bounds for other classes of linearly parameterized models, such as Laguerre or Kautz filters, will be the object of future work.

## REFERENCES

- Garulli, A. (1999). Tight error bounds for projection algorithms in conditional set membership estimation. *Sys. Contr.Lett.* **37**, 293–300.
- Garulli, A., A. Tesi and A. Vicino (Eds.) (1999a). *Robustness in Identification and Control*. Lecture Notes in Control and Information Sciences. Springer. London.
- Garulli, A., A. Vicino and G. Zappa (2000). Conditional central algorithms for worst-case set membership identification and filtering. *IEEE Trans. Automat. Contr.* **45**(1), 14–23.
- Garulli, A., B. Z. Kacewicz, A. Vicino and G. Zappa (1999b). Reliability of projection algorithms in conditional estimation. *J. Optim. Theory Appl.* **101**(1), 1–14.
- Giarrè, L., B. Z. Kacewicz and M. Milanese (1997). Model quality evaluation in set membership identification. *Automatica* **33**(6), 1133–1139.
- Kacewicz, B. Z. (1999). Worst-case conditional system identification in a general class of norms. *Automatica* **35**, 1049–1058.
- Mäkilä, P. M., J. R. Partington and T. K. Gustafsson (1995). Worst-case control-relevant identification. *Automatica* **31**(12), 1799–1819.
- Milanese, M. and A. Vicino (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica* **27**(6), 997–1009.
- Van den Hof, P. M. J., P. S. C. Heuberger and J. Bokor (1995). System identification with generalized orthonormal basis functions. *Automatica* **31**(12), 1821–1834.
- Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Trans. Automat. Contr.* **36**, 551–562.
- Wahlberg, B. (1994). System identification using Kautz models. *IEEE Trans. Automat. Contr.* **39**, 1276–1282.