Optimal Worst Case Estimation for LPV-FIR Models with Bounded Errors

(Article begins on next page)

27 April 2024

Optimal worst case estimation for LPV–FIR models with bounded errors

Gustavo Belforte[a,*], Paolo Gay[b]

[a] *Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy*
[b] *Dipartimento di Economia e Ingegneria Agraria Forestale e Ambientale, Università degli Studi di Torino,
via Leonardo da Vinci 44, Grugliasco (TO), Italy*

Abstract

In this paper discrete time linear parameter varying (LPV) models with finite impulse response (FIR) dynamic structure are considered. Measurement errors are assumed to be bounded. In such condition optimal input sequences minimizing the worst case parameter uncertainties are derived. The main result of this paper consists in finding optimal worst case input sequences for LPV–FIR models. These are obtained suitably combining the optimal design results for standing alone invariant FIR models and standing alone nonlinear memoryless blocks. The quite relevant improvement obtained using optimal input sequences instead of random sequences is shown by simulation.
© 2004 Elsevier B.V. All rights reserved.

Keywords: LPV models; Parameter uncertainties; Set membership identification; Optimal experiment design; Worst case design

1. Introduction

In linear parameter varying (LPV) models the parameters of the equations linking the inputs u to the outputs y are in general functions of a (possibly vector valued) variable p, assumed to be measurable, that determines the operating condition of the system [15]. Although such class of model structures is less general than generic linear time varying models, in several applications it is perceived to be quite satisfactory for representing real systems. Moreover, its applicability to the industrial practice of gain scheduling makes it appealing for control design problems.

For identification purposes, the p variable plays the role of an auxiliary input that, beside being always measurable, in some cases is also under the designer's control. In [2,7], for example, LPV models have been used to approximate distributed parameter systems in environmental and agricultural context and the p variable has been associated to the spatial coordinate that specifies where measurements are collected so that choosing the values of input p actually corresponds to fixing the sensor allocations. In [9] the power generated by a nuclear plant (that can be set to different values within a given range) represents the operating condition and is associated to the p variable while in [10] the vector valued auxiliary input p is constituted by a set of airplane operating parameters that can be chosen to some extent.

---

* Corresponding author.

E-mail addresses: gustavo.belforte@polito.it (G. Belforte),
paolo.gay@unito.it (P. Gay).

LPV models have in general a number of parameters that is quite bigger than the number of parameters in their corresponding time invariant models and therefore also the number of measurements required for parameter estimation is correspondingly large. This fact increases the interest in getting better estimates with less measurements through optimal experiment design that consists in finding those sequences, both for input u and for variable p, that allow to derive parameter estimates with the lowest uncertainty. Applying such sequences to u and p whenever they are completely under the designer's control allows to obtain the best possible parameter estimates. When p cannot be arbitrarily settled, optimal input sequences are still of interest since they provide bounds on the best achievable parameter reliability and specify for which values of p measurements should be collected. Remark that while the best identifiability of the LPV model is ensured if optimal input sequences are applied such model results to be almost unidentifiable if poor variations of p occur.

In this paper attention is focused on discrete-time linear parameter varying (LPV) models with finite impulse response (FIR) dynamic structure. Measurement errors and uncertainties are embedded in a set-membership framework, consequently the input sequence optimality is evaluated in terms of the worst case estimate uncertainty it allows to achieve.

The main result of this paper consists in showing that optimal input sequences for LPV–FIR models are obtained combining together optimal input sequences for time invariant FIR models with those for standing alone nonlinear blocks describing the dependence of the parameters on the p variable. The dynamic part of the LPV model is restricted to be a FIR only because, to the best of our knowledge, optimal input sequences for time invariant linear dynamic models have been so far obtained for FIR models only. This occurrence is probably motivated by the circumstance that optimal input sequences for FIR models are independent of their actual parameter values and this fact, in general, does not hold true for other structures.

The paper is organized as follows: basic notation is introduced in Section 2 while in Section 3 proper input sequences $\{(u_k, p_k)\}$ of fixed length n^{opt} for the input u and for the variable p are derived. Such sequences minimize the worst case parameter uncertainty with respect to any possible error realization. Although the collection of extra measurements beside the n^{opt} optimal one cannot improve the worst case parameter uncertainty, nevertheless it can reduce the actual parameter uncertainty to an extent that is related to the obtained error realization.

The case in which more than n^{opt} measurements can be collected is then addressed in Section 4. Criteria for choosing the best input sequences in such conditions are presented and finally the performances that can be achieved using these sequences are evaluated with a simulation study in Section 5.

2. Notation and generalities

Let a linear parameter varying discrete-time model, whose dynamics is described by a FIR, be represented by

$$y_k = B(q^{-1}, p_k)u_k + e_k, \tag{1}$$

where y_k, u_k, p_k and e_k denote, respectively, the system output and input, the operating condition p and the measurement error e at sample k, while q^{-1} is the usual delay operator.

The dynamic behavior of the underlying system is described by the regressor

$$B(q^{-1}, p_k) = b_0(p_k) + b_1(p_k)q^{-1} + \cdots$$
$$+ b_{n_b-1}(p_k)q^{-n_b+1}. \tag{2}$$

The generic ith FIR coefficient $b_i(p_k)$ is assumed to be a nonlinear function of variable p_k expressed as a linear combination of given functions $f_j(p_k)$, $j = 1, \ldots, N$, so that

$$b_i(p_k) = \sum_{j=1}^{N} b_{i,j} f_j(p_k), \tag{3}$$

where $b_{i,j}$, $i = 0, \ldots, n_b - 1$, $j = 1, \ldots, N$, are the parameters to be actually estimated.

Relation (3) can be represented as a nonlinear static block NL_i with input p_k and output b_i as depicted in Fig. 1. The whole LPV–FIR system can then be graphically represented as reported in Fig. 2.

Parameters $b_{i,j}$ to be estimated can be rearranged in one single vector θ as

$$\theta = [b_{0,1} \ b_{0,2} \ \ldots \ b_{0,N} \ b_{1,1} \ \ldots \ b_{n_b-1,N}]. \tag{4}$$
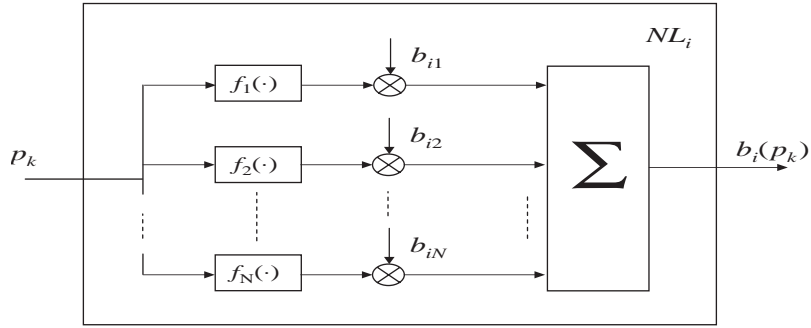
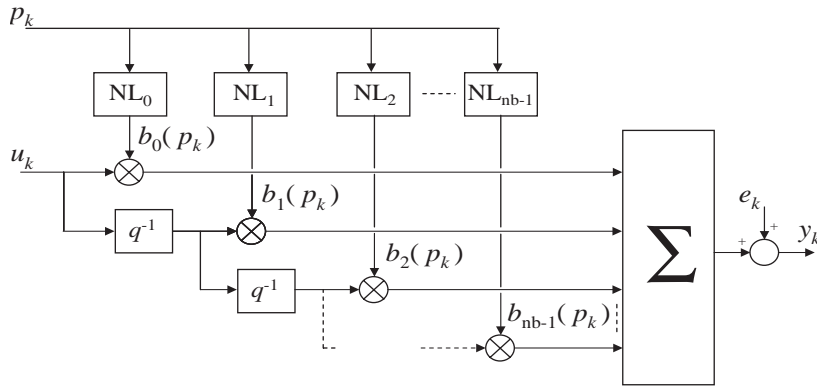Fig. 1. The NL_i nonlinear static block structure.

Fig. 2. The LPV–FIR model.

When m consecutive measurements are available, using relation (1) the following matrix equation is obtained:

$$y = A(u, p)\theta + e, \qquad (5)$$

where $A(u, p)$ is the regression matrix and $y \in R^m$, $u \in R^{m+n_b-1}$, $e \in R^m$ and $p \in R^m$ are the measurement vector, the input vector, the measurement error vector and the operating condition vector, respectively.

The regression matrix $A(u, p)$ has the following structure:

$$A(u, p) =$$
$$\begin{bmatrix} V(u_1, p_1) & V(u_0, p_1) & \dots & V(u_{-n_b+2}, p_1) \\ \vdots & \vdots & & \vdots \\ V(u_k, p_k) & V(u_{k-1}, p_k) & \dots & V(u_{k-n_b+1}, p_k) \\ \vdots & \vdots & & \vdots \\ V(u_m, p_m) & V(u_{m-1}, p_m) & \dots & V(u_{m-n_b+1}, p_m) \end{bmatrix}$$
$$(6)$$

with

$$V(u_i, p_k) = [u_i f_1(p_k)\ u_i f_2(p_k)\ \dots\ u_i f_N(p_k)]. \qquad (7)$$

In this paper the system input u_k and the operating condition parameter p_k are assumed to be bounded so that

$$\Omega_u = \{u \in R^{m+n_b-1}: |u_k| \leqslant U, k = -n_b + 2, \dots, m\} \qquad (8)$$

and

$$\Omega_p = \{p \in R^m: p_a \leqslant p_k \leqslant p_b,\ k = 1, \dots, m\}. \qquad (9)$$

Also the measurement error is assumed to be componentwise bounded so that the error vector e must belong to its membership set Ω_e defined as

$$\Omega_e = \{e \in R^m: |e_k| \leqslant E,\ k = 1, \dots, m\}. \qquad (10)$$

Relation (5) and (10) associate to each kth measurement a "strip" S_k of feasible parameters in

the parameter space

$$S_k = \{\theta \in R^{N \cdot n_b}: -E \leqslant y_k - a_k^{\mathrm{T}}\theta \leqslant +E\}, \qquad (11)$$

where a_k^{T} is the kth row of the matrix $A(u, p)$.

When m consecutive measurements are considered, their joint information has to be stored. Such information is summarized by the parameter feasible set $\mathscr{D}(u, p)$

$$\mathscr{D}(u, p) = \{\theta \in R^{N \cdot n_b}: y = A(u, p)\theta + e, \ e \in \Omega_e\}$$

$$= \bigcap_{k=1}^{m} S_k \qquad (12)$$

that collects all the parameter vectors θ consistent with model (5), the actual measurements y and the errors e. From relation (12) it follows that the feasible set $\mathscr{D}(u, p)$ is a polytope described by a subset of the $2m$ inequalities

$$y_k - E \leqslant a_k^{\mathrm{T}}\theta \leqslant y_k + E, \qquad k = 1, \ldots, m \qquad (13)$$

defining the boundaries of S_k.

Any vector in $\mathscr{D}(u, p)$ could be the one that generated the data and therefore the reliability of the estimate is evaluated through the "size" of $\mathscr{D}(u, p)$. Such "size" is usually evaluated in terms of the width of the parameter uncertainty intervals $W(PUI_{b_{i,j}})$ defined as

$$W(PUI_{b_{i,j}}(u, p)) = b_{i,j}^M - b_{i,j}^m, \qquad (14)$$

where

$$b_{i,j}^m = \min_{\theta \in \mathscr{D}(u,p)} b_{i,j} \quad \text{and} \quad b_{i,j}^M = \max_{\theta \in \mathscr{D}(u,p)} b_{i,j}. \qquad (15)$$

An alternative measure of the "size" of $\mathscr{D}(u, p)$ is its radius $r_h(\mathscr{D}(u, p))$ that represents the maximum achievable estimation error in some suitable l_h norm and is given by

$$r_h(\mathscr{D}(u, p)) = \inf_{\vartheta \in R^{N \cdot n_b}} \sup_{\theta \in \mathscr{D}(u,p)} \|\vartheta - \theta\|_h. \qquad (16)$$

Throughout this paper the l_∞ norm radius $r_\infty(\mathscr{D}(u, p))$ will be considered.

Remark that $\mathscr{D}(u, p)$, beside being a function of the input vector u and the operating condition p, also depends on the output vector y and thus on the unknown error realization. In order to get rid of this last dependance it is common to refer to the maximum achievable "size" with respect to any possible error realization, referred to as the *worst case estimation error*.

The largest width of the parameter uncertainty intervals (denoted in the following as $W(P\hat{U}I_{b_{i,j}}(u, p))$) and the largest radius (denoted in the following as $r(\hat{\mathscr{D}}(u, p))$) with respect to any error realization can be easily computed a priori, for linear in the parameter systems (see e.g. [11,13,14]) from relations (14)–(16) in which, however, the set $\mathscr{D}(u, p)$ is replaced by the set $\hat{\mathscr{D}}(u, p)$ defined as

$$\hat{\mathscr{D}}(u, p) = \{\theta \in R^{N \cdot n_b}: A(u, p)\theta = e, \ e \in \Omega_e\}. \qquad (17)$$

Remark that, once the error realization dependance has been discarded considering worst case uncertainties, $W(P\hat{U}I_{b_{i,j}}(u, p))$ and $r_\infty(\hat{\mathscr{D}}(u, p))$ depend only on the applied input sequences u and p that determine the set of measurements defining $\hat{\mathscr{D}}(u, p)$. Such inputs are constrained by relations (8) and (9) only. It would be convenient to derive the lowest (or optimal) worst case uncertainties that can be obtained whatever the input strings are. Such quantities, denoted by $W_{b_{i,j}}^{\mathrm{opt}}$ and r_∞^{opt}, are defined as

$$W_{b_{i,j}}^{\mathrm{opt}} \doteq \min_{\substack{u \in \Omega_u \\ p \in \Omega_p}} W(P\hat{U}I_{b_{i,j}}(u, p)),$$

$$r_\infty^{\mathrm{opt}} \doteq \min_{\substack{u \in \Omega_u \\ p \in \Omega_p}} r_\infty(\hat{\mathscr{D}}(u, p)). \qquad (18)$$

It can be shown that $W_{b_{i,j}}^{\mathrm{opt}}$ and r_∞^{opt} depend on a rather low number n^{opt} of optimal measurements [1]. Whenever the m measurements of a generic feasible parameter set $\mathscr{D}(u, p)$ include the n^{opt} optimal measurements the corresponding worst case parameter uncertainties and worst case radius are $W_{b_{i,j}}^{\mathrm{opt}}$ and r_∞^{opt}, respectively. It can be shown that n^{opt} is bounded between n_p and n_p^2 where $n_p = N \cdot n_b$ is the total number of parameters to be estimated [1]. The interest is then in finding suitable input strings $\{(u_k, p_k)\}$ that allow to collect a set of n^{opt} optimal measurements.

Indeed in general it is not ensured that optimal input strings of minimal length, that allow to collect only n^{opt} optimal measurements, exist.

Models for which the optimal worst case estimation error can be obtained with a set of only n_p suitable measurements are referred to as *minimal worst case* models. Some sufficient condition for a model to be minimal worst case can be found in literature [1,12].

For the special case of time invariant FIR systems with n_b parameters, that however, can also be represented by Eqs. (1)–(5) assuming, in relation (3), $N = 1$

and $f_1(p) = 1$, it has been proved [6] that such systems are *minimal worst case*, hence the worst case estimation uncertainty can be obtained with a set of only n_b suitable measurements. Moreover, it has been shown that input sequences that allow to get n_b subsequent measurements leading to the worst case estimation error exist. One of these optimal input series is the one in which

$$\begin{cases} u_i = U, & i = 1, \ldots, n_b, \\ u_i = -U, & i = n_b + 1, \ldots, 2n_b. \end{cases} \tag{19}$$

This sequence is also "cyclic" in the sense that $u_{k+n_b} = -u_k$ and can therefore be cyclically applied to the input u. For such kind of optimal input series it turns out that $2n_b - 1$ input samples must first be applied in order to collect n_b measurements (the FIR filter needs to be loaded) thereafter any new set of n_b subsequent inputs allows to collect a new set of n_b measurements that alone allows to determine again the worst case estimation error. The collection of more than n_b measurements, although it does not affect the worst case estimation error, nevertheless can reduce the actual estimation error. Remark that in general there are several optimal "cyclic" input series that can be found by exhaustive search. Among these it is in general convenient to choose the one that leads to the smallest volume of the corresponding $\hat{\mathscr{D}}(u)$ set as discussed in [6].

For what concerns the worst case error radius $r_\infty(\hat{\mathscr{D}}(u, \cdot))$ of the resulting time invariant FIR system it has been proved [6] that

$$\inf_{u \in \Omega_u} r(\hat{\mathscr{D}}(u, \cdot)) = E/U \tag{20}$$

and always from the results in [6], although it is not explicitly stated therein, it can be easily shown that the minimal worst case parameter uncertainty interval width $W_{b_{i,1}}^{\text{opt}}$ is

$$W_{b_{i,1}}^{\text{opt}} = \min_{u \in \Omega_u} (W(P\hat{U}I_{b_{i,1}}(u, \cdot))) = 2E/U,$$

$$i = 0, \ldots, n_b - 1. \tag{21}$$

Consider now a generic nonlinear static sub-system NL_i as depicted in Fig. 1, whose behavior is described by relation (3). For such standing alone block it is possible to derive a set of N^{opt} optimal input levels $\{p_1^{\text{opt}}, \ldots, p_{N^{\text{opt}}}^{\text{opt}}\}$ that, whenever applied in sequence

to the input p, allow the optimal worst case identification [1] of parameters $b_{i,j}$, $j = 1, \ldots, N$.

In fact, for any family of linear independent functions $\{f_1(p), \ldots, f_N(p)\}$ on the interval $[p_a, p_b]$, the N^{opt} optimal input levels $\{p_1^{\text{opt}}, \ldots, p_{N^{\text{opt}}}^{\text{opt}}\}$ can always be numerically derived and $N \leqslant N^{\text{opt}} \leqslant N^2$, as discussed in [1]. However, for some particular families of functions, like complete Taylor polynomial bases where

$$f_j(p_k) = p_k^{j-1}, \quad j = 1, \ldots, N, \tag{22}$$

closed form analytical results are available [5]. Taylor polynomial models are minimal worst case, so that $N^{\text{opt}} = N$, and the set of optimal input levels $\{p_1^{\text{opt}}, \ldots, p_{N^{\text{opt}}}^{\text{opt}}\}$ over the interval $[p_a, p_b]$ is constituted (see [4,8]) by

$$p_1^{\text{opt}} = p_a,$$

$$p_i^{\text{opt}} = p_a + \tfrac{1}{2}(p_b - p_a)(1 + \cos \tfrac{i}{N-1}\pi),$$

$$i = 1, \ldots, N - 2,$$

$$p_N^{\text{opt}} = p_b \tag{23}$$

that are the values for which the 1st kind Chebyshev polynomial $T_{N-1}(p)$ of degree $N - 1$ associated with the interval $[p_a, p_b]$ achieves its extremal (maximal and minimal) values.

Remark that, since in relation (3) functions $f_j(p)$, $j = 1, \ldots, N$, are the very same for all the N nonlinear blocks, the N^{opt} optimal input levels are the same for each block and $N \leqslant N^{\text{opt}} \leqslant N^2$. It follows that the worst case widths of the parameter uncertainty intervals result to be

$$W_j^{\text{opt}} = \inf_{p \in \Omega_p} W(P\hat{U}I_{b_{0,j}}(\cdot, p))$$

$$= \inf_{p \in \Omega_p} W(P\hat{U}I_{b_{1,j}}(\cdot, p)) = \cdots$$

$$= \inf_{p \in \Omega_p} W(P\hat{U}I_{b_{n_b-1,j}}(\cdot, p)). \tag{24}$$

Such optimal worst case width of the parameter uncertainty intervals for the standing alone static nonlinear blocks NL_i, $i = 1, \ldots, n_b$, are told to be normalized when $E = 1$. In such condition they are denoted with $\bar{W}_j^{\text{opt}}$, $j = 1, \ldots, N$.

---

[1] When identifying a standing alone nonlinear block as the one depicted in Fig. 1 an error e (not reported in figure) is assumed to exist so that relation (3) becomes $b_i(p_k) = \sum_{j=1}^{N} b_{i,j} f_j(p_k) + e_k$.

3. Optimal worst case experiment design

In this section the main result of the paper is presented. It consists in showing that the optimal worst case experiment design for *LPV–FIR* models is obtained combining together the optimal design results of standing alone invariant *FIR* models and standing alone nonlinear memoryless blocks. Optimal input sequences for *LPV–FIR* models can be chosen to be "cyclic" and whenever the standing alone nonlinear block is minimal worst case also the corresponding *LPV–FIR* model is minimal worst case.

The following theorem allows to find optimal input sequences $\{(u_k, p_k)\}$ both for the input u and for the operating condition p.

Theorem 1. *Consider a LPV–FIR model as described by relations* (1)–(3).

Let $\bar{u}$ be any "cyclic" sequence that is optimal for the time invariant FIR system with n_b parameters.

Let $\{p_1^{\mathrm{opt}}, \ldots, p_{N^{\mathrm{opt}}}^{\mathrm{opt}}\}$ be the set of N^{opt} optimal input levels for the identification of any standing alone static nonlinear block NL_i described by relation (3) *and $\bar{W}_j^{\mathrm{opt}}$, $j = 1, \ldots, N$, be the normalized optimal worst case uncertainty of its parameters.*

Then, for the LPV–FIR model, the following statements hold:

- *the lowest worst case uncertainty $W_{b_{i,j}}^{\mathrm{opt}}$ of parameter $b_{i,j}$ is given by*
$$W_{b_{i,j}}^{\mathrm{opt}} = \bar{W}_j^{\mathrm{opt}} \frac{E}{U}, \quad i = 0, \ldots n_b - 1, \quad j = 1, \ldots, N.$$
- *the optimal result is achieved collecting $n^{\mathrm{opt}} = n_b \cdot N^{\mathrm{opt}}$ optimal measurements generated with the following sequences for input u and for variable p:*
  - *the input u is a sequence of $n_b(N^{\mathrm{opt}} + 1)$ samples such that $u_{k+n_b} = -u_k$ and the first n_b samples are equal to $\bar{u}$ (a "cyclic" optimal sequence for time invariant FIR),*
  - *the variable p is a sequence of $n_b \cdot N^{\mathrm{opt}}$ samples such that $p_{i+n_b(k-1)} = p_k^{\mathrm{opt}}$, $i = 1, \ldots, n_b$, $k = 1, \ldots, N^{\mathrm{opt}}$.*

Proof. When the input sequences described in the theorem statement are applied to a $LPV - FIR$ model, each one of the N^{opt} sets of n_b measurements obtained with $p = p_k^{\mathrm{opt}}$ can be regarded as a set of n_b optimal measurements for a time invariant FIR derived from

the *LPV–FIR* model setting the operating condition p to the constant value p_k^{opt} and applying to the input u an optimal "cyclic" sequence $\bar{u}$. Each one of the N^{opt} sets of n_b measurements then allows to derive a set of n_b values $b_i(p_k^{\mathrm{opt}})$, $i = 0, \ldots, n_b - 1$, whose uncertainty intervals are all equal to $\pm E/U$ according to the previously recalled results in [6].

The worst case parameter bounds widths $W(P\hat{U}I_{b_{i,j}})$ of the *LPV–FIR* model parameters $b_{i,j}$ can then be derived rearranging the $b_i(p_k^{\mathrm{opt}})$ values into n_b sets of the N^{opt} elements $b_i(p_k^{\mathrm{opt}})$, $k = 1, \ldots, N^{\mathrm{opt}}$.

Each of these sets can be regarded as a set of N^{opt} optimal measurements for the identification of the standing alone nonlinear block NL_i, in fact from relation (3) it is possible to derive the following set of N^{opt} equations

$$b_i(p_k^{\mathrm{opt}}) = \sum_{j=1}^{N} b_{i,j} f_j(p_k^{\mathrm{opt}}), \quad k = 1, \ldots, N^{\mathrm{opt}}. \quad (25)$$

Since the uncertainty intervals of the $b_i(p_k^{\mathrm{opt}})$ are all equal to $\pm E/U$ and the p_k^{opt} values are the optimal ones for the standing alone nonlinear block identification, it follows that

$$W(P\hat{U}I_{b_{i,j}}) = \bar{W}_j^{\mathrm{opt}} \frac{E}{U}, \quad j = 1, \ldots, N.$$

This is obtained from relation (25), substituting the $b_i(p_k^{\mathrm{opt}})$, $k = 1, \ldots, N^{\mathrm{opt}}$ values with their uncertainties $\pm E/U$ and reminding that $\bar{W}_j^{\mathrm{opt}}$, $j = 1, \ldots, N$, are the normalized optimal worst case widths of the parameters for the standing alone nonlinear blocks when the uncertainty is equal to ± 1.

If any measure is added to the previously defined set of optimal measurements, it is easy to show that no improvement can be achieved. In fact, if the new added measure is collected with a p level equal to any of the optimal ones $\{p_1^{\mathrm{opt}}, \ldots, p_{N^{\mathrm{opt}}}^{\mathrm{opt}}\}$, then the new measure combined with the other ones with the same value of $p = p_k^{\mathrm{opt}}$ will lead, in the worst error context, to the same value $b_i(p_k^{\mathrm{opt}})$ and no improvement on the $LPV - FIR$ parameters will be obtained.

If on the contrary the new added measure is collected with a p value different from the optimal ones its uselessness can be proved supposing to collect not only one, but n_b measurements with the same value of p. This set of measurements allows to derive a new set of n_b values $b_i(p)$, $i = 0, \ldots, n_b - 1$, that have, according to the previously recalled results of

[6], an uncertainty interval of $\pm E/U$. This information however does not improve the worst case uncertainty of the parameters $b_{i,j}$ since it only adds a set of n_b relations of the form

$$b_i(p) = \sum_{j=1}^{N} b_{i,j} f_j(p)$$

that, by definition, is useless with respect to the determination of the worst case width of the parameter uncertainty intervals.

Finally, it is easy to see that if any of the previously recalled optimal $n_b \cdot N^{\text{opt}}$ measurements is omitted, the worst case parameter uncertainty interval width is not achieved. In fact, if one of the optimal measurements is omitted there will no longer be a set of N measurements for one of the optimal levels p_k^{opt} so that the bounds of the corresponding parameters $b_i(p_k^{\text{opt}})$, $i = 0, \ldots, n_b - 1$, will be larger than $\pm E/U$ thus leading to larger bounds on at least one set of the parameters $b_{i,j}$. $\square$

The previous results hold indeed also for a vector valued p variable.

Finally, remark that these results can be extended to the case in which (3) is substituted by

$$b_i(p) = \sum_{j=1}^{N} b_{i,j} f_{i,j}(p) \tag{26}$$

so that different sets of functions are used to represent the dependance on p of each one of the FIR parameters.

In this case the set of optimal levels of p consists of the union of the input sets that are optimal for the n_b static blocks represented by relation (26) with $i = 0, \ldots, n_b - 1$.

4. Non-minimal input sequences design

Although the worst case identification error cannot be improved adding extra measurements, the actual identification error can be reduced adding more measurements provided that their error realization is not the worst one. Since this occurrence is quite common in several practical cases, in this section we analyze the case in which more than $n_b \cdot N^{\text{opt}}$ measurements can be collected and their error realization is unlikely to be the worst one. To formalize

this last statement it is convenient to introduce some mild statistical assumption for the measurement error e that, beside being bounded over the domain $[-E, +E]$ is also characterized by an unknown distribution. With only this assumption it is still difficult to derive general results therefore the error distribution is restricted to be *not-impulsive* so that it associates probability zero to the event of obtaining the same error realization in two different measurements.

Under such conditions, whenever the considered model is minimum worst case, the collection of extra measurements beside the first $n^{\text{opt}} = n_b \cdot N$ leads to a reduction of the actual measurement error according to the following Theorem.

Theorem 2. *Given a minimum worst case model, if the bounded measurement error is* not-impulsive, *then for any optimal input sequence* $\{(u_k, p_k)\}$ *providing* $n_b \cdot N^{\text{opt}}$ *measurements, the probability of reducing the* $W(PUI_{b_{i,j}})$ *of all the parameters replicating one time only the already performed* $n_b \cdot N^{\text{opt}}$ *measurements is one.*

Proof. Recalling relation (12), the $\mathscr{D}(u, p)$ set, obtained with $n_b \cdot N^{\text{opt}}$ measurements, is defined as the intersection of the $n_b \cdot N^{\text{opt}}$ strips S_k, $k = 1, \ldots, n_b \cdot N^{\text{opt}}$,

$$S_k = \{\theta \in R^{n_b \cdot N^{\text{opt}}}: \ -E \leqslant y_k - a_k^{\text{T}}\theta \leqslant E\}. \tag{27}$$

Since, by assumption, the probability of obtaining the same error realization performing two measurements is zero, the replication of an existing measurement leads, with probability one, to a new measurement $y_k' \neq y_k$, whose associated strip is

$$S_k' = \{\theta \in R^{n_b \cdot N^{\text{opt}}}: \ -E \leqslant y_k' - a_k^{\text{T}}\theta \leqslant E\}. \tag{28}$$

Since the two strips in relation (27) and (28) are oriented along the same direction (they share the same regressor a_k^{T}) in the parameter space, their information can be combined in an equivalent strip $\bar{S}_k$

$$\bar{S}_k = S_k \cap S_k'$$
$$= \{\theta \in R^{n_b \cdot N^{\text{opt}}}: \ -E_k \leqslant \bar{y}_k - a_k^{\text{T}}\theta \leqslant E_k\} \tag{29}$$

relative to the equivalent measurement $\bar{y}_k$

$$\bar{y}_k = \frac{y_k + y_k'}{2} \tag{30}$$

with an associated error bound E_k, such that

$$2E_k = \min\{(y_k + E), (y'_k + E)\}$$
$$- \max\{(y_k - E), (y'_k - E)\}. \qquad (31)$$

Since $y'_k \neq y_k$ with probability one, also $E_k < E$ with probability one. Assuming $E_k = E$ whenever the kth measurement has not been replicated, from the sign structure of the inverse of the regression matrix $A(u, p)$ [1,12] it results

$$\begin{bmatrix} \theta_0^M - \theta_0^m \\ \theta_1^M - \theta_1^m \\ \vdots \\ \theta_{n_b \cdot N^{\mathrm{opt}} - 1}^M - \theta_{n_b \cdot N^{\mathrm{opt}} - 1}^m \end{bmatrix}$$

$$= 2abs(A^{-1}(u, p)) \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n_b \cdot N^{\mathrm{opt}}} \end{bmatrix}. \qquad (32)$$

Therefore, the reduction of any of the error bounds E_k reduces the $W(PUI_{b_{i,j}})$ of all the parameters. ☐

Since the $W(PUI_{b_{i,j}})$ obtained using the input sequences of Theorem 1 are the smallest that can be recovered from $n_b \cdot N^{\mathrm{opt}}$ measurements only, it follows that the best strategy consists in replicating such optimal measurements (i.e. to provide the same input sequences).

Collecting a new set of the measurements, forcing the same input sequence, is not the unique way to reduce, with probability one, the *PUI* of the parameters. In fact, other input sequences, whose associate measurements reduce with probability equal to one the *PUI* of all the parameters may exist, but their existence and location depend on the error realization of the first $n_b \cdot N^{\mathrm{opt}}$ measurements. Since the computation of the location of these measurements requires the knowledge of the true parameter vector, they can not be practically computed [3].

It appears that the *not-impulsive* error hypothesis alone does not allow to derive results about the *PUI* reduction if more than $2(n_b \cdot N^{\mathrm{opt}})$ measurements can be performed. In such conditions it is however reasonable to conjecture that more than two replications of the measurements (i.e. the collection of the measurement applying two times the same input sequences) are more convenient with respect to the collection of measurement with generic input sequences, different from the optimal ones of Theorem 1. In order to test this conjecture, a numerical simulation has been performed and is presented in the next section.

5. Numerical example

In this section the effectiveness of the proposed optimal input design is illustrated with a simulation study. An LPV system consisting of a 15th degree FIR, whose parameters are fifth-degree polynomials in the p variable was considered. The actual values of the parameters used to generate output data were randomly obtained and are reported in Table 1. It was assumed that the input variable u and the operating condition parameter p were allowed to vary over the intervals $[-1, 1]$ and $[0, 1]$, respectively.

Two possible input sequences were considered: the first one is the optimal sequence in Theorem 1 with $\bar{u}$ as in relation (19), "cyclically" replicated; the second one is constituted by random sequences, both for u and p, uniformly distributed over their support intervals. For both possible input sequences 100 measurement sets were generated (100 sets with optimal input and 100 with random input). Each measurement set consisted of 465 input samples (the first 15 of which were used for loading the FIR) and 450 output samples.

Measurements were obtained as follows:

- *Optimal measurement sets*: the optimal input sequence was applied 100 times and a random error uniformly distributed over the interval $[-1, +1]$ was added to the resulting output sequences.
- *Random measurement sets*: 100 different random sequences were applied to the system input and random errors uniformly distributed over the interval $[-1, +1]$ were added to the resulting outputs.

From each measurement set five subsets of different length were obtained considering the first $90 \cdot r$, $r = 1, \ldots, 5$, measurements. Remark that $90 = n_b \cdot N^{\mathrm{opt}}$ is the minimum number of measurements required for identifying the system, so that, when the optimal input sequence is considered the five mentioned subsets are

Table 1
Parameters $b_{i,j}$ used for the simulation

$b_{i,j}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 1$	2.9003	1.8114	1.0305	2.0056	1.9931	2.4055
$i = 2$	1.4623	2.8709	2.4936	2.4189	2.7995	2.0931
$i = 3$	2.2137	2.8338	1.8902	1.8578	2.6433	1.8898
$i = 4$	1.9720	1.8205	2.8636	1.6092	2.2898	2.3891
$i = 5$	2.7826	2.7873	1.9320	1.3793	2.6359	2.2426
$i = 6$	2.5242	1.1158	1.8373	1.3869	2.3205	2.5896
$i = 7$	1.9129	1.7057	2.6924	2.3644	1.6839	2.9137
$i = 8$	1.0370	2.6263	2.0503	1.6055	1.5795	2.0452
$i = 9$	2.6428	1.0197	1.4053	2.0833	1.6824	2.7603
$i = 10$	1.8894	1.2778	2.3443	1.3017	2.0682	1.3459
$i = 11$	2.2309	1.4055	2.6762	2.3958	2.4542	2.9595
$i = 12$	2.5839	1.3974	1.0393	1.7567	1.6186	1.5429
$i = 13$	2.8436	2.2076	2.3626	2.7200	2.6770	1.5047
$i = 14$	2.4764	1.5444	1.7590	2.7073	2.1361	2.7515
$i = 15$	1.3525	1.3976	2.6636	2.1871	1.7408	2.4746

Table 2
Performance indexes Min_r and Avg_r as functions of r

Criteria	r (used measurements = $90r$)				
	1	2	3	4	5
Avg_r	961.07	4.95	3.28	2.77	2.53
Min_r	54.42	3.62	2.45	2.00	1.72

derived applying the optimal input sequence r times ($r = 1, \ldots, 5$) as discussed in Section 4.

From each measurement subset the system parameters were estimated and the estimates' reliability was evaluated in terms of the ℓ_∞ radius of the corresponding parameter feasible set $\mathscr{D}$. For each possible input sequence (optimal or random) and for each length of measurement subset ($90r$, $r = 1, \ldots, 5$) 100 different evaluations of the ℓ_∞ radius denoted with $\rho_r^{\text{opt}}(i)$, $r = 1, \ldots, 5$, $i = 1, \ldots, 100$, and $\rho_r^{\text{rand}}(i)$, $r = 1, \ldots, 5$, $i = 1, \ldots, 100$, respectively were therefore available.

The minimum and the average of these sets of 100 radiuses have been computed and the performance of the optimal input with respect to the random one was evaluated in terms of the ratios between the minimal radiuses and the average radiuses obtained with the two possible input sequences (random and optimal), that is with the two following indexes Min_r and Avg_r,

$r = 1, \ldots, 5$, defined as

$$\text{Min}_r = \frac{\min_i \rho_r^{\text{rand}}(i)}{\min_i \rho_r^{\text{opt}}(i)}, \quad \text{Avg}_r = \frac{1/100 \sum_i \rho_r^{\text{rand}}(i)}{1/100 \sum_i \rho_r^{\text{opt}}(i)},$$

$$r = 1, \ldots, 5.$$

Results are reported in Table 2. Results show that optimal input sequences perform quite better than random input sequences.

6. Conclusions

In this paper the problem of the worst case identification of discrete time linear parameter varying (LPV) models with finite impulse response (FIR) dynamic structure in the case of unknown but bounded errors was studied. In particular optimal input design for the worst case identification was addressed. Suitably combining available results for optimal worst case

identification of time invariant FIR systems with those relative to optimal worst case identification of nonlinear static blocks, optimal input sequences that allow lowest worst case parameter uncertainties for LPV–FIR systems were derived.

The case in which it is possible to collect more measurement than the minimum required was addressed as well, outlining strategies for optimally choosing such measurements. The effectiveness of the proposed approach was shown through a simulation study whose numerical results show that optimal input sequences perform quite better than random ones.

Acknowledgements

References

[1] G. Belforte, B. Bona, S. Frediani, Optimal sampling schedule for parameter estimation of linear models with unknown but bounded measurement errors, IEEE Trans. Automat. Control AC-32 (2) (1987) 179–182.

[2] G. Belforte, F. Dabbene, P. Gay, LPV approximation of distributed parameter systems in environmental modeling, Proceedings of iEMSs 2002 Biennial Conference of the International Environmental Modelling and Software Society, Vol. 2, Lugano, Switzerland, 24–27 June, 2002, pp. 66–71.

[3] G. Belforte, P. Gay, Optimal non minimal experiment design for static non-linear systems corrupted by unknown but bounded errors, Proceedings of 14th World Congress of IFAC, Beijing, China, July 1999, pp. 325–330.

[4] G. Belforte, P. Gay, Optimal experiment design for regression polynomial models identification, Internat. J. Control 75 (15) (2002) 1178–1189.

[5] G. Belforte, P. Gay, G. Monegato, Some new properties of Chebyshev polynomials, J. Comput. Appl. Math. 117 (2000) 175–181.

[6] G. Belforte, T.T. Tay, Optimal input design for worst case system identification in $l_1/l_2/l_\infty$, Systems Control Lett. 20 (1993) 273–278.

[7] F. Dabbene, P. Gay, C. Tortia, Modelling and control of steam soil disinfestation processes, BioSystems Eng. 84 (3) (2003) 247–256.

[8] P. Gay, Nonlinear (static) blocks in dynamics systems: identification and optimal experiment design in the case of set membership errors, Ph.D. Thesis, Politecnico di Torino, 1998.

[9] M.V. Kothare, B. Mettler, M. Morari, P. Bendotti, C. Falinower, Level control in the steam generator of a nuclear power plant, IEEE Trans. Control Systems Technol. 8 (2000) 55–69.

[10] C.H. Lee, M.H. Shin, M.J. Chung, A design of gain-scheduled control for a linear parameter varying system: an application to flight control, Control Eng. Practice 9 (2001) 11–21.

[11] A.G. Marchuk, K.Y. Oshipenko, Best approximation of functions specified with an error at a finite number of points, Mat. Zametki 17 (1975) 359–368 (in Russian); Transl. Math. Notes 17 (1975) 207–212 (in English).

[12] C.A. Micchelli, Optimal sampling design for parameter estimation and p-widths under stochastic and deterministic noise, in: M. Milanese, et al., (Eds.), Robustness in Identification and Control, Plenum Press, New York, 1988, pp. 25–40.

[13] C.A. Micchelli, T.J. Rivlin, A survey of optimal recovery, in: C.A. Micchelli, T.J. Rivlin (Eds.), Optimal Estimation in Approximation Theory, Plenum Press, New York, 1977, pp. 1–54.

[14] M. Milanese, R. Tempo, Optimal algorithms theory for robust estimation and prediction, IEEE Trans. Automat. Control AC-30 (1985) 730–738.

[15] J. Shamma, D. Xiong, Set valued methods for linear parameter varying systems, Automatica 35 (1999) 1081–1089.