A constraint selection technique for set membership estimation of time-varying parameters

Marco Casini, Andrea Garulli, Antonio Vicino

Abstract—This paper presents a new recursive algorithm for approximating the feasible parameter set, in set membership estimation of time-varying parameters. The novelty of the approach lies in the use of a constraint selection technique which keeps track only of a subset of the linear constraints defining the feasible set. These are chosen as the binding constraints of suitable linear programs, that are instrumental to recursively update an orthotope containing the true feasible set. It is shown through several numerical examples that the proposed technique provides an approximation which is almost as tight as the batch minimum orthotope containing the feasible set, while its computational load is much smaller than that required to propagate the exact feasible parameter set.

Index Terms—Set membership estimation; time-varying systems; recursive identification; linear programming.

I. INTRODUCTION

Tracking of time-varying parameters is a key problem in system identification and adaptive control, that has been widely studied under the assumption that both measurement noise and parameter variations can be modelled as stochastic processes (see e.g. [1], [2], [3] and references therein). A popular alternative approach is provided by the so-called set membership estimation setting [4], [5], [6], in which both noise and parameter disturbances are assumed to be unknown-but-bounded (UBB). This allows one to define a set of parameters which are consistent with the UBB assumptions, usually addressed as feasible parameter set (FPS). For example, if linear regression models with UBB signals bounded in the infinity norm are considered, the FPS is a polytope in the parameter space.

While a very large number of contributions is available in the set membership framework for identification of time-invariant systems or state estimation of dynamic systems, fewer works have been specifically dedicated to the problem of tracking time-varying parameters in the bounded error setting. The papers [7], [8] can be recalled among the first ones to propose modified versions of classic recursive schemes based on ellipsoids or parallelepipeds. Moreover, and more importantly, the conservatism introduced by the proposed approximation with respect to the exact minimum orthotope containing the FPS is negligible, in spite of the much smaller number and size (in terms of active constraints) of the LPs to be solved, compared to those necessary to keep track of the exact FPS.

The paper is organized as follows. Section II introduces the notation and the definitions used in the paper. The problem formulation is given in Section III. Section IV illustrates the proposed recursive FPS approximation procedure. Three numerical examples are presented in Section V to evaluate the quality of the approximation provided by the proposed approach, while concluding remarks are given in Section VI.

II. NOTATION AND DEFINITIONS

Let $A \in \mathbb{R}^{m \times n}$. Then, $|A|$ denotes the matrix whose entries are the absolute values of the entries of $A$. The $j$th row of $A$ is denoted by $A_j$, while if $\mathcal{I} \subset \mathbb{N}$, $A_{\mathcal{I}}$ is the matrix obtained by selecting the rows of $A$ with indexes in $\mathcal{I}$.

An orthotope, or axis-aligned box, is defined as

$$O(\theta, d) = \{\theta : \theta = \bar{\theta} + \text{diag}(d)w, \|w\|_{\infty} \leq 1\},$$

where $\bar{\theta}, d, w \in \mathbb{R}^n$, $d_i \geq 0$, $i = 1, \ldots, n$, and $\text{diag}(d)$ is a diagonal matrix with diagonal equal to $d$. We denote by $\mathcal{F}_i = \{\theta \in O : \theta_i = \bar{\theta}_i + d_i\}$ and $\mathcal{F}_{i+n} = \{\theta \in O : \theta_i = \bar{\theta}_i - d_i\}$, for $i = 1, \ldots, n$, the $(n-1)$-dimensional faces of the orthotope $O$. The weighted $\ell_\infty$ unit ball is given by

$$B_{\infty}(\varepsilon) = \{w : |w_i| \leq \varepsilon_i, \quad i = 1, \ldots, n\} = O(0, \varepsilon).$$
The symbol ⊕ denotes the Minkowski sum of sets
\[ C_1 \oplus C_2 = \{ \theta : \theta = \alpha + \beta, \alpha \in C_1, \beta \in C_2 \}. \]
Trivially, \( O(\theta_a, d_a) \oplus \ldots \oplus \Theta(t) \), according to the strategy (12)-(13), which is summarized by the following procedure.

Consider the linear program (LP)
\[
\max c^T x \\
\text{s.t.} \quad Ax \leq b,
\]
A constraint \( A_i x \leq b_i \) is said to be active if there exists a feasible \( x \) for which \( A_i x = b_i \). A constraint \( A_i x \leq b_i \) is a binding constraint of the LP if there exists a solution \( x^* \) of (4) such that \( A_i x^* = b_i \). The region defined by the binding constraints of an LP is called binding set and denoted by \( A \) (obviously, \( A \) contains the entire feasible region of the LP).

### III. PROBLEM FORMULATION

Consider the time-varying linear regression model
\[
y(t) = \varphi^T(t)\theta(t) + e(t),
\]
where \( \theta(t) \in \mathbb{R}^n \) is the parameter vector to be estimated, \( \varphi(t) \) is a known vector containing past values of the system input and output signals, and \( e(t) \) is an unknown-but-bounded noise such that
\[
|e(t)| \leq \delta(t), \quad \forall t,
\]
where \( \delta(t) \) is a known nonnegative sequence. The parameter vector varies in time according to
\[
\theta(t+1) = \theta(t) + w(t),
\]
where \( w(t) \in \mathbb{R}^n \) is UBB in the weighted \( \ell_{\infty} \) norm, i.e.
\[
|w_i(t)| \leq \epsilon_i(t), \quad \forall t, \ i = 1, \ldots, n,
\]
and \( \epsilon_i(t) \) are known nonnegative sequences.

The evolution of the FPS \( \Theta(t) \) is described according to the following recursion
\[
\Theta^+(t+1) = \Theta(t) \oplus B_{\infty}(\epsilon(t)) \quad \text{(9)}
\]
\[
\Theta(t+1) = \Theta^+(t+1) \cap S(t+1) \quad \text{(10)}
\]
where
\[
S(t) = \{ \theta : |y(t) - \varphi^T(t)\theta| \leq \delta(t) \}\]
is the measurement feasibility set at time \( t \).

The FPS \( \Theta(t) \) is a generic polytope, whose number of faces tends to grow in time. For this reason, a wide variety of recursive set approximation techniques have been proposed in the literature. They are usually based on a family of sets \( \mathcal{R} \) of fixed complexity, and proceed by subsequent approximations satisfying the inclusions
\[
\mathcal{R}^+(t+1) \supseteq \mathcal{R}(t) \oplus B_{\infty}(\epsilon(t)) \quad \text{(12)}
\]
\[
\mathcal{R}(t+1) \supseteq \mathcal{R}^+(t+1) \cap S(t+1). \quad \text{(13)}
\]

By choosing an initial approximation \( \mathcal{R}(0) \supseteq \Theta(0) \), where \( \Theta(0) \) is the initial FPS (usually derived from a priori information on the system parameters), one has the guarantee that \( \Theta^+(t) \subseteq \mathcal{R}(t) \) and \( \Theta(t) \subseteq \mathcal{R}(t) \), at every time instant \( t \). Typical choices of the approximating regions \( \mathcal{R} \) include orthotopes, ellipsoids, parallelepipeds, zonotopes, limited complexity polyhedra. The computation of \( \mathcal{R}^+(t+1) \) and \( \mathcal{R}(t+1) \) in (12)-(13), can be based on various criteria, such as volume minimization or other indicators assessing the overall parametric uncertainty.

In this paper, the sets \( \mathcal{R} \) are orthotopes, defined as in (1). It is well known that if this class of sets is plugged in the recursive scheme (12)-(13), minimizing the volume at each step, the resulting approximation of the FPS is extremely coarse. On the other hand, the minimum volume orthotope containing the FPS, \( \Theta^*(\Theta(t)) \), can be computed by solving 2n LPs at each time \( t \), whose number of constraints increases with \( t \). In the next section, a constraint selection technique is proposed which allows one to compute a recursive orthotopic approximation of the FPS by solving a much smaller number of LPs, with a limited number of constraints.

### IV. CONSTRAINT SELECTION TECHNIQUE

In order to devise a procedure for recursively approximating the FPS \( \Theta(t) \), we consider separately the time update step (12) and the measurement update step (13).

#### A. Time update

Let us introduce the following result which is instrumental to the time update (12).

**Proposition 1:** Let \( \mathcal{C} = \{ \theta : A\theta \leq b \} \) be a polytope in \( \mathbb{R}^n \) and let \( \Theta^*(\mathcal{C}) = O(\bar{\theta}, d) \). Then,
\[
\mathcal{C} \oplus B_{\infty}(\epsilon) = \mathcal{T} \bigcap O(\bar{\theta}, d + \epsilon) \quad \text{(14)}
\]
where
\[
\mathcal{T} = \{ \theta : A\theta \leq b + |A|\epsilon \} \quad \text{(15)}
\]
and
\[
O(\bar{\theta}, d + \epsilon) = \Theta^*(\mathcal{C} \oplus B_{\infty}(\epsilon)). \quad \text{(16)}
\]

**Proof:** See [17].

Proposition 1 states that the time update (9) can be performed by computing the intersection between two sets: the enlarged version of the polytope \( \Theta(t) \), according to (15), and the enlarged bounding box \( \Theta^*(\Theta(t)) \oplus B_{\infty}(\epsilon(t)) \). This suggests a way to recursively update an approximating orthotope \( \mathcal{R}(t) \supseteq \Theta(t) \), according to the strategy (12)-(13), which is summarized by the following procedure.
Step T0. At a generic time $t$, let
- $\mathcal{C}(t) = \{\theta : A(t)\theta \leq b(t)\}$ be a polytope such that $\mathcal{C}(t) \supseteq \Theta(t)$;
- $\mathcal{O}(t) = \mathcal{O}(\tilde{\theta}(t), d(t))$ be an orthotope such that $\mathcal{O}(t) \supseteq \Theta(t)$;
- $v^{(i)}(t) \in \mathcal{C}(t)$, $i = 1, \ldots, 2n$ be $2n$ elements such that $v^{(i)}(t)$ belongs to the $i$th face of the orthotope $\mathcal{O}(t)$.

Step T1. According to (15), define the enlarged polytope
\[ \mathcal{C}^+(t+1) = \left\{ \theta : A(t+1)\theta \leq b^+(t+1) \right\} \] (17)
where
\[
A(t+1) = A(t) \quad \text{and} \quad b^+(t+1) = b(t) + |A(t)|e(t). \quad (18)
\]
Moreover, define the enlarged orthotope
\[ \mathcal{O}^+(t+1) = \mathcal{O}(t) \cup B_{\infty}(\epsilon) = \mathcal{O}(\tilde{\theta}^+(t+1), d^+(t+1)) \] (20)
where
\[
\tilde{\theta}^+(t+1) = \tilde{\theta}(t), \quad d^+(t+1) = d(t) + \epsilon(t). \quad (21)
\]

Step T2. For each $v^{(i)}(t)$, $i = 1, \ldots, 2n$, find a set of $n$ indexes
\[ \mathcal{I}_i = \{j_h \in \mathbb{N} : A_{jh}(t)v^{(i)}(t) = b_{jh}(t), \ h = 1, \ldots, n\} \]
such that $A_{\mathcal{I}_i}$ is nonsingular, and set
\[ \hat{v}^{(i)}(t+1) = [A_{\mathcal{I}_i}(t+1)]^{-1}b_{\mathcal{I}_i}^+(t+1). \quad (23) \]

It is worth observing that in the procedure T0-T2, the sets $\mathcal{C}^+(t+1)$ and $\mathcal{O}^+(t+1)$ in (17) and (20) play the role of the sets $\mathcal{C}$ and $\mathcal{O}(\tilde{\theta}, d+\epsilon)$ in the right hand side of (14). Therefore, the following result is a straightforward application of Proposition 1 to (9) and (17)-(20).

**Proposition 2:** Assume that at time $t$, $\mathcal{C}(t) = \Theta(t)$ and $\mathcal{O}(t) = \mathcal{O}^*(\Theta(t))$. Then,
\[ \Theta^+(t+1) = \mathcal{C}^+(t+1) \cap \mathcal{O}^+(t+1), \]
where $\mathcal{C}^+(t+1)$ and $\mathcal{O}^+(t+1)$ are given by (17) and (20), respectively.

Loosely speaking, Proposition 2 guarantees that the time update step is not conservative, if applied to the true FPS $\Theta(t)$ and to its corresponding minimum outer orthotope $\mathcal{O}^*(\Theta(t))$.

**B. Measurement update**

The measurement update procedure is summarized by the following steps.

**Step M0.** Let $\mathcal{C}^+(t+1)$, $\mathcal{O}^+(t+1)$, $\hat{v}^{(i)}(t+1) \in \mathcal{C}^+(t+1)$, $i = 1, \ldots, 2n$ be given, according to procedure T0-T2, and let $\mathcal{S}(t+1)$ be the new measurement set at time $t+1$. For $i = 1, \ldots, 2n$, define the sets
\[ C^+_i(t+1) = \{\theta : A_{\mathcal{I}_i}(t+1)\theta \leq b_{\mathcal{I}_i}^+(t+1)\} \]
with $\mathcal{I}_i$ given by (23). Notice that this is a subset of constraints of $\mathcal{C}^+(t+1)$ that are active at $\hat{v}^{(i)}(t+1)$.

**Step M1.** For each $i = 1, \ldots, 2n$, if $\hat{v}^{(i)}(t+1) \notin \mathcal{S}(t+1)$ set
\[ v^{(i)}(t+1) = \text{arg max} \min \quad e_i^T\theta \quad \text{s.t.} \quad \theta \in C^+_i(t+1) \cap \mathcal{S}(t+1) \]
and $C_i(t+1) = A_{\mathcal{I}_i}(t+1)$, where $A_{\mathcal{I}_i}(t+1)$ is the binding set of the LP (24). Otherwise, if $\hat{v}^{(i)}(t+1) \in \mathcal{S}(t+1)$, set $v^{(i)}(t+1) = \hat{v}^{(i)}(t+1)$ and $C_i(t+1) = C^+_i(t+1)$.

**Step M2.** Set
\[ C(t+1) = \bigcap_{i=1}^{2n} C_i(t+1), \]
and compute $\mathcal{O}^*(\mathcal{C}(t+1)) = \mathcal{O}(\tilde{\theta}(t+1), \hat{d}(t+1))$, where
\[
\tilde{\theta}_{i}(t+1) = \frac{v^{(i)}(t+1) + v_{i}^{(i+n)}(t+1)}{2},
\]
\[
\hat{d}_{i}(t+1) = \frac{v^{(i)}(t+1) - v_{i}^{(i+n)}(t+1)}{2},
\]
for $i = 1, \ldots, n$.

**Step M3.** Set
\[ \mathcal{O}(t+1) = \mathcal{O}^*(\mathcal{C}(t+1)) \cap \mathcal{O}^+(t+1) = \mathcal{O}(\tilde{\theta}(t+1), \hat{d}(t+1)), \]
where
\[
\tilde{\theta}(t+1) = \frac{u+l}{2}, \quad \hat{d}(t+1) = \frac{u-l}{2},
\]
with
\[
u_i = \min \left\{ \tilde{\theta}_{i}(t+1) + \hat{d}_{i}(t+1), \tilde{\theta}_{i}^+(t+1) + d_{i}^+(t+1) \right\},
\]
\[ l_i = \max \left\{ \tilde{\theta}_{i}(t+1) - \hat{d}_{i}(t+1), \tilde{\theta}_{i}^+(t+1) - d_{i}^+(t+1) \right\}. \]

Then, proceed with the next time update at time $t+1$.

By construction, $\mathcal{C}(t+1)$, $\mathcal{O}(t+1)$ and the elements $v^{(i)}(t+1)$ satisfy the same properties of their counterparts at time $t$, and hence the procedure can be iterated.

The key idea underlying the measurement update procedure, which allows to limit the number of constraints in the set $\mathcal{C}(t)$, is to keep track only of the binding constraints whenever a new LP is solved in (24). The role of the elements $v^{(i)}(t)$ is twofold: they are used in step M1 to decide whether to perform an LP or not, and in step M2 to update the approximating orthotope. Notice that the approach

\[1\text{In (24), the notation max(min) means that max holds for } i = 1, \ldots, n, \text{ while min holds for } i = n + 1, \ldots, 2n.\]
proposed in [16] is adapted here to cope with the different structure of the approximation scheme. In fact, in [16] the approximating orthotope is just $O^+(C(t))$, where $C(t)$ is a subset of the constraints of the FPS $\Theta(t)$. On the contrary, in this paper the approximating orthotope is such that, in general, $O(t) \subset O^+(C(t))$.

V. NUMERICAL EXAMPLES

In this section, the proposed procedure for estimation of time-varying parameters is compared to other recursive approaches taken from the literature. In particular, three algorithms will be considered.

a) A recursive outer bounding ellipsoidal estimator (ROBE), performing the measurement update according to the algorithm proposed in [18] (minimum volume ellipsoid containing the intersection between an ellipsoid and the measurement set $S(t)$), and the time update by first computing the minimum ellipsoid containing $B_\infty(\varepsilon)$ and then applying the algorithm in [19] which provides the optimal bounding ellipsoid for the vector sum of two ellipsoids.

b) A recursive outer bounding parallelopetotic estimator (ROBP), corresponding to the algorithm proposed in [20].

c) The minimum volume orthoptic estimator, which computes the exact $O^+(\Theta(t))$ at each time $t$, by keeping track of all active constraints in $\Theta(t)$. This requires to solve $2n$ LPs whenever at least one of the constraints in $S(t+1)$ intersects the time updated polytope $\Theta^+(t+1)$.

A. Example 1

The first case study is taken from [21]. Consider the FIR model

$$y(t) = b_1(t)u(t) + b_2(t)u(t-1) + e(t)$$

where the time-varying parameters $b_1(t)$, $b_2(t)$ have been generated according to

$$b_1(t) = 1.5 + \sin(2\pi t/3000);$$
$$b_2(t) = 0.5 + \sin(2\pi t/1500).$$

and $e(t)$ is unknown-but-bounded as in (6). The input signal $u(t)$ has been randomly generated within the interval $[-1, 1]$. The bounds on the parameter variations have been considered constant and set to $c(t) = [2\pi/3000, 2\pi/1500]$, $\forall t$. The noise bound $\delta(t)$ is constant and equal to 1, and the noise signal $e(t)$ has been generated according to two different distributions (uniform and Gaussian). All results are averaged over 200 identification experiments with 4000 data points each.

In order to compare the performance of the considered recursive algorithms, we report the value of uncertainty associated to each parameter. This corresponds to the half length of the orthotope sides for the orthotopic approximations, while the minimum volume orthotope containing the ellipsoidal and parallelopetotic approximating set has been considered for the ROBE and ROBP algorithms. Figure 1 shows these uncertainties for the parameters $b_1(t)$ and $b_2(t)$ for a uniformly distributed noise $e(t)$, while Figure 2 does the same for the case of Gaussian noise. It can be observed that the proposed technique provides a tighter approximation of the parametric uncertainty with respect to the ROBE and ROBP algorithms. Moreover, and more importantly, such approximation is almost indistinguishable from that given by the minimum orthotope $O^+(\Theta(t))$, thus testifying that the constraint selection technique has captured the essential constraints which contribute to define the true FPS. However, the computational burden required by the proposed technique is much smaller than that necessary to keep track of the true set $O^+(\Theta(t))$. This is shown by Table I, reporting the average number of active constraints and of LPs to be performed per time step.

B. Example 2

Consider the time-varying ARX model

$$y(t) + a_1(t)y(t-1) + a_2(t)y(t-2) =$$

$$= b_1(t)u(t-1) + b_2(t)u(t-2) + e(t)$$

where the parameter vector $\theta(t) = [a_1(t) \ a_2(t) \ b_1(t) \ b_2(t)]^\prime$ changes in time according to model (7), with process noise $w(t)$ satisfying (8), with

$$e(t) = [1/50 0 1/100 0]^\prime, \ \forall t$$

and $\theta(0) = [0.2 \ 0.5 \ 0.5 \ 0.25]^\prime$ (notice that $a_2(t)$ and $b_2(t)$ are constant parameters, while $a_1(t)$ and $b_1(t)$ are time-varying, with different variation rates). The signals $w_1(t)$ and $w_3(t)$ have been generated from a uniform random distribution, within the bounds (8). The input $u(t)$ and the measurement noise $e(t)$ have been generated as in Example 1, with $\delta(t) = 0.2, \forall t$. Results are averaged over 100 different realizations of $u(t), e(t)$ and $w(t)$.

Figures 3 and 4 show the uncertainties for the auto-regressive parameters $a_1(t)$ and $a_2(t)$, when $e(t)$ is generated according to a uniform or a Gaussian distribution, respectively (similar results are obtained for the input parameters $b_1(t)$ and $b_2(t)$). It can be observed that the proposed technique outperforms the ROBE and ROBP algorithm, and provides a very small overestimation of the actual parametric uncertainty given by the minimum orthotope $O^+(\Theta(t))$. Indeed, the two orthotopic estimates are almost superimposed for the time-invariant parameter $a_2$, while some more conservatism is observed in the approximation of the time varying parameter $a_1$. As long as the computational burden is concerned, Table II shows that the average number of active constraints and solved LPs per time step propagated by the proposed technique is significantly smaller than the corresponding figures for the propagation of the true feasible

---

**TABLE I**

<table>
<thead>
<tr>
<th></th>
<th>Unif. noise</th>
<th>Gauss. noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(t)$</td>
<td>4.98</td>
<td>5.80</td>
</tr>
<tr>
<td>active LPs</td>
<td>0.25</td>
<td>0.51</td>
</tr>
<tr>
<td>$O^+(\Theta(t))$</td>
<td>5.62</td>
<td>8.74</td>
</tr>
</tbody>
</table>
Fig. 1. Example 1, uniform noise: uncertainty associated to parameters $b_1(t)$ (a) and $b_2(t)$ (b): minimum orthotope $O^*(\Theta(t))$ (black), proposed approximating orthotope $O(t)$ (red), ROBP (green); ROBE (cyan).

Fig. 2. Example 1, Gaussian noise: uncertainty associated to parameters $b_1(t)$ (a) and $b_2(t)$ (b): minimum orthotope $O^*(\Theta(t))$ (black), proposed approximating orthotope $O(t)$ (red), ROBP (green); ROBE (cyan).

TABLE II

<table>
<thead>
<tr>
<th></th>
<th>Unif. noise</th>
<th>Gauss. noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(t)$</td>
<td>12.75</td>
<td>1.94</td>
</tr>
<tr>
<td>$O^*(\Theta(t))$</td>
<td>16.14</td>
<td>4.98</td>
</tr>
</tbody>
</table>

set $\Theta(t)$. By comparing Tables I and II, it may be argued that the computational benefits of the proposed technique tend to increase with the dimension of the parameter space. This is confirmed by numerical tests performed on higher dimensional models (see [17]).

VI. CONCLUSIONS

The proposed technique provides an effective trade-off between quality of the approximation and required computational load, for set membership tracking of time-varying parameters. In fact, while the exact propagation of the true feasible set eventually leads to a polytope with a large number of active constraints, the presented numerical examples have demonstrated that only a small subset of such constraints is sufficient to compute an orthotopic approximation which is almost as tight as the minimum one. Moreover, while the latter may require the solution of up to $2n$ LPs per time step, the proposed constraint selection technique allows one to solve much fewer LPs, and therefore can be effectively employed in real-time estimation problems. Finally, the resulting approximation turns out to be significantly tighter than that provided by standard recursive techniques based on ellipsoidal or paralleleotopic approximations.

The extension of the proposed recursive approximating scheme to set membership state estimation of dynamic systems is the subject of ongoing research.

REFERENCES

approximating orthotope $O$.

Fig. 3. Example 2, uniform noise: uncertainty associated to parameters $a_1(t)$ (a) and $a_2(t)$ (b): minimum orthotope $O^\ast(\Theta(t))$ (black), proposed approximating orthotope $\hat{O}(t)$ (red), ROBP (green); ROBE (cyan).

Fig. 4. Example 2, Gaussian noise: uncertainty associated to parameters $a_1(t)$ (a) and $a_2(t)$ (b): minimum orthotope $O^\ast(\Theta(t))$ (black), proposed approximating orthotope $\hat{O}(t)$ (red), ROBP (green); ROBE (cyan).


