A Novel Recursive Linear Estimator Based on the Frisch Scheme

Stefano Massaroli[1,†], Renato Miyagusuku[1,†], Federico Califano[2,†],
Angela Faragasso[1,†], Atsushi Yamashita[1,†], Hajime Asama[1,‡]

***Abstract*— This paper proposes an innovative identification scheme to estimate parameters constituting linear relations in time–invariant systems: the *bounding box recursive Frisch* scheme. A novel recursive version of the *Frisch scheme*, a linear estimator characterised by mild prior assumptions in the *error-in-variables* (EIV) framework, has been derived. The fast computational time and convergence in the identification of linear systems are the most relevant feature of this recursive version of the scheme. The performance of the developed algorithm has been evaluated trough numerical simulations. Results proved the effectiveness and accuracy of the proposed solution.**

I. Introduction

One of the most explored mathematical problems is the extraction of linear relations from data which are affected by additive noise. Many of the research work regarding this problem focused on finding a single solution, e.g. least squares methods, but, unfortunately, the uniqueness of the solution can be achieved only by introducing prior assumptions in the estimation scheme, whose correctness cannot be confirmed by the available data. However, according to Kalman [1], the *Frisch Scheme* (first introduced by the Nobel prize Ragnar Frisch in 1934 [2]) appears to be less prejudiced than most other schemes. The peculiarity of the scheme has been intensively investigated in [3], [4], [5]. The main feature of the *Frisch Scheme*, which remarks a positive distinction from others identification algorithms, is represented by the milder assumptions on the noisy input data that lead to a whole space of solutions fitting a given dataset, instead of a single one. In the analysis of real physical systems described by *error-in-variables* (EIV) models, the *Frisch scheme* establishes a *post-identification* degree of freedom in the choice of one set of parameters between feasible solutions which are compatible with the physical constraints of the problem, e.g. masses cannot be negative. The knowledge of these constraints provides a way to add to the "blind" estimation step, which depends only on the acquired data, a rational selection procedure of the real parameters.

Besides, once collected a data set, the family of solutions provided by the *Frisch scheme* may be very large, making the selection process meaningless. In practice, to obtain a solutions set of a useful size, one should collect a great amount of data or should perform a *well conditioned* experiment, which result to be not be suitable for online estimations or diagnostics purposes. A recursive version of the algebraic *Frisch scheme*, in which the solution space was greatly reduced in size by subdividing the data in different subsets and iteratively approximating the intersection of the resulting simplices by means of static particles, has previously been introduced by the authors in [6]. However, due to the curse of dimensionality, the number of particles involved in the algorithm increased exponentially with the dimension of the system, leading to extremely large computation times.

In this paper, to overcome dimensional problems, we propose a novel scheme whose computation complexity scales polynomially with the number of parameters. The new algorithm indeed shrinks the size of the solution space significantly. However, the improvement in computational complexity is traded with a reduced estimation accuracy with respect to [6]. This paper introduces of a new Frisch–scheme–based linear recursive estimator, which consistently restricts the searching space of the solution without loosing the trademark of a *post-identification* degree of freedom. Here, the mathematical formulation of the *bounding box recursive Frisch* (BBRF) scheme is presented and tested on a generic EIV model.

II. Background

The *Frisch scheme* is an effective method to derive models from noisy measurements through a *modification* of the observed data. The scheme relies on *a priori* assumptions which are necessary to extract linear relations from data affected by error.

A. Assumptions for EIV Estimation Schemes

Consider the linear equation

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0 \tag{1}$$

relating the n variables x_i through n scalars α_i. The *observation matrix* X is constructed with m samples of the variables as

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \quad , \quad X \in \mathbb{R}^{m \times n} \tag{2}$$

Relation (1) is equivalent to

$$XA = 0 \tag{3}$$

[1]Stefano Massaroli, Renato Miyagusuku, Angela Faragasso, Atsushi Yamashita and Hajime Asama are with the Department of Precision Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan; email: {massaroli,miyagusuku,faragasso,yamashita,asama}@robot.t.u-tokyo.ac.jp

[2]Federico Califano is with the Faculty of Electrical Engineering, Mathematics Computer Science (EWI), Robotics and Mechatronics (RAM), University of Twente, Hallenweg 23 7522NH, Enschede, The Netherlands; email: f.califano@utwente.nl

† *Member, IEEE*, ‡ *Fellow, IEEE*

where $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)^\top$. In general, $A \in \mathbb{R}^{n \times q}$ being q the number of linearly independent parameters vectors compatible with the data. Let the *sample covariance matrix* be $\Sigma = \frac{X^\top X}{m}$. Then, (3) can be rewritten as

$$\Sigma A = 0 \tag{4}$$

since Σ and X have identical null spaces. Notice that the diagonal of Σ contains the variances of the variables while the covariances are the entries of the off-diagonal elements. Let us assume that additive noise affects the variables, i.e. $x_i = \hat{x}_i + \tilde{x}_i$ where $\hat{x}_i$, $\tilde{x}_i$ are, respectively, the unknown exact value of the variable and the noise sample. The additive noise yields $\Sigma > 0$ making unsolvable the problem of extracting linear relations (i.e. A) from its null space. The only way is to modify the observations Σ. Under the usual assumptions made in EIV schemes of zero–mean value of the noise samples and orthogonality of noise and noiseless samples (see [5]), it turns out that $\Sigma = \hat{\Sigma} + \tilde{\Sigma}$, in which $\hat{\Sigma} = \frac{\hat{X}^\top \hat{X}}{m}$ and $\tilde{\Sigma} = \frac{\tilde{X}^\top \tilde{X}}{m}$, where $\hat{X}$, $\tilde{X}$ are defined as in (2) with $\hat{x}_i$ and $\tilde{x}_i$ instead of x_i. Hereafter, follows a formalization of the problem of seeking linear relations in noisy data.

Definition 1 (EIV Estimation Problem [1])**.** *Let Σ be a sample covariance matrix. Determine $\tilde{\Sigma}$ such that*

$$\hat{\Sigma} = \Sigma - \tilde{\Sigma} \geq 0 \qquad \det(\hat{\Sigma}) = 0 \tag{5}$$

Any base of $\mathrm{null}(\hat{\Sigma})$ will span a space describing a set of linear relations compatible with the data and the assumptions, i.e. it is possible to solve $\hat{\Sigma} A = 0$. Estimation schemes can be distinguished by the assumptions on the noise, i.e. on $\tilde{\Sigma}$.

The estimation schemes relevant this to this work are formally introduced in the following subsections.

B. Ordinary Least Squares (OLS) [7]

This estimator assumes noise affecting only one variable. Let the noisy variable be the i-th. Therefore,

$$\tilde{\Sigma} = \mathrm{diag}(0, \cdots, 0, \tilde{\sigma}_i^2, 0, \cdots, 0)$$

where $\tilde{\sigma}_i^2$ is the variance of the noise $\tilde{x}_i$.

Let Σ_i be obtained from Σ deleting its i-th row and column. The solution of the estimation problem is given by

$$\tilde{\sigma}_i^2 = \frac{\det(\Sigma)}{\det(\Sigma_i)} \tag{6}$$

Equation (6) defines the upper bound of noise level satisfying condition (5). Notice that, when the noisy variable is not specified, the OLS has n distinct solutions, each of which minimizes the squared norm of the estimation error, $\min_A e^\top e$:

$$e^\top e = \| [X]_i - X_i A \|^2 \tag{7}$$

where $[X]_i$ is the i-th column of X and X_i is obtained from X deleting its i-th column. $A \in \mathbb{R}^{n-1}$ describes the linear dependence of the orthogonal projection of $[X]_i$ on the subspace spanned by the columns of X_i from these generators. If we assume that the noisy variable is the i-th, a unique closed-form solution of the problem (5) is given by $A = (X_i^T X_i)^{-1} X_i^T [X]_i$. In this case A belongs to $\mathbb{R}^{n-1}$ and not to $\mathbb{R}^n$ because the i-th element is implicitly normalized to -1.

C. The Frisch Scheme

In this scheme, noise samples affecting different variables are assumed to be mutually independent, i.e.

$$\tilde{\Sigma} = \mathrm{diag}\left(\tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \cdots, \tilde{\sigma}_n^2\right)$$

Notice that for a given sample covariance matrix Σ, the n OLS solutions are also solutions of the Frisch estimation scheme. Hence, the prior assumption of the *Frisch scheme* is more general and milder than those of all the other estimators providing a single solution, e.g. OLS with a prior on the noisy variable (usually computed by pseudoinversion). Besides, closed-form solutions are more difficult to be obtained, as, in general, infinite solutions exist. Here, part of the results in [5] are employed to address this problem. In particular, we need to introduce a definition and a theorem. Let $\mathrm{Maxcor}_{\mathcal{F}}(\Sigma)$ be defined as the maximum number of linear relations which can be extracted from Σ in the context of the *Frisch scheme*. Therefore

$$\mathrm{Maxcor}_{\mathcal{F}}(\Sigma) = \max_{\Sigma \in \mathcal{D}} \left\{ \dim\left(\mathrm{null}(\hat{\Sigma}) \right) \right\}$$

where $\mathcal{D}$ is the set of all the diagonal matrices satisfying (5), i.e., all the *Frisch scheme* solutions.

Theorem 1. *[5] Let $\mathrm{Maxcor}_{\mathcal{F}}(\Sigma) = 1$. Then, the following hold:*

1. *The vectors A of all the linear relations consistent with the Frisch scheme lie (normalizing one parameter to 1) inside the simplex of the parameters space whose vertices are the n OLS solutions.*
2. *There exists a one-to-one relation linking the points of the simplex and the solutions $\tilde{Sigma}$ of the Frisch scheme.*

III. Bounding-Box Recursive Method

A. Geometrical Properties of Simplices and Bounding Boxes

From the aforementioned assumptions, it is possible to define the set of solutions of the *Frisch scheme* as a simplex of the parameter space which can be easily computed by its vertices. In order to derive a recursive version of the *Frisch scheme*, some geometrical properties of *simplices* and *bounding boxes* have to be introduced.

Definition 2 (Simplex Matrix)**.** *Let $v_i \in \mathbb{R}^n$, $i = 1, ..., n+1$ be the vertices of a n-dimensional simplex. The simplex matrix $\mathcal{S}^n$ is defined as the matrix whose columns are the $n+1$ vertices of the simplex, i.e.*

$$\mathcal{S}^n = (v_1 \ v_2 \ \cdots \ v_{n+1}) \in \mathbb{R}^{n \times n+1}$$

Because of convexity, any set of vertices uniquely identify one and only one simplex. Given a simplex matrix $\mathcal{S}^n$ the

corresponding simplex is the set of all the *convex combinations* of the columns of $\mathcal{S}^n$

$$\mathfrak{S}(\mathcal{S}^n) = \mathrm{conv}(\mathcal{S}^n) = \\ = \left\{ \sum_{i=1}^{n+1} \mu_i [\mathcal{S}^n]_i : (\mu_i \geq 0 \ \forall i) \wedge \sum_{i=1}^{n+1} \mu_i = 1 \right\}$$

where $[\mathcal{S}^n]_i$ is the i-th column of $\mathcal{S}^n$.

According to [8], the volume (i.e. the n-dimensional Lebesgue measure) of a n-simplex $\mathfrak{S}(\mathcal{S}^n)$ can be computed as

$$\lambda(\mathfrak{S}(\mathcal{S}^n)) = \frac{1}{n!} \left| \det \begin{pmatrix} \mathcal{S}^n \\ 1 \cdots 1 \end{pmatrix} \right| \tag{8}$$

Definition 3 (Bounds of a Simplex)**.** *Let $\mathcal{S}^n = \mathcal{S}^n_{ij}$ be a simplex matrix. The lower bound of the simplex corresponding to $\mathcal{S}^n$ is $l(\mathcal{S}^n) = (l_1, \cdots, l_n)$ with $l_i = \min_j(\mathcal{S}^n_{ij})$ $\forall i = 1, \ldots, n$. Analogously, the upper bound of the simplex corresponding to $\mathcal{S}^n$ is $u(\mathcal{S}^n) = (u_1, \cdots u_n)$ with $u_i = \max_j(\mathcal{S}^n_{ij})$ $\forall i = 1, \ldots, n$.*

Recalling the concepts of minimum bounding box and axis-aligned minimum bounding box (AABB) [9], we would like to define a simple set enclosing any n-dimensional simplex, e.g. the AABB given by the Cartesian product of the simplex bounds.

Definition 4 (Simplex Bounding Box)**.** *Let $\mathcal{S}^n$ be a simplex matrix and l,u be the bounds of the corresponding simplex $\mathfrak{S}(\mathcal{S}^n)$. The simplex bounding box related to $\mathcal{S}^n$ is the AABB of $\mathfrak{S}(\mathcal{S}^n)$. It is indeed unique and defined as*

$$\mathfrak{B}(\mathcal{S}^n) = b_1 \times b_2 \times \cdots \times b_n = \prod_{i=1}^{n} b_i$$

where $b_i = [l_i, u_i]$ is an interval and "$\times$" indicates the Cartesian product. It follows that $\mathfrak{S}(\mathcal{S}^n) \subseteq \mathfrak{B}(\mathcal{S}^n)$.

Hence, the volume of a *simplex bounding box* $\mathfrak{B}(\mathcal{S}^n)$ is

$$\lambda(\mathfrak{B}(\mathcal{S}^n)) = \prod_{i}^{n} |u_i - l_i|$$

Proposition 1. *Given N simplex matrices $\mathcal{S}^n_1$,$\mathcal{S}^n_2$,...,$\mathcal{S}^n_N$, it holds*

$$\bigcap_{i=1}^{N} \mathfrak{S}(\mathcal{S}^n_i) \subseteq \bigcap_{i=1}^{N} \mathfrak{B}(\mathcal{S}^n_i)$$

Proof: It is known that

$$\mathfrak{S}(\mathcal{S}^n_i) \subseteq \mathfrak{B}(\mathcal{S}^n_i) \quad \forall i = 1, 2, \ldots N.$$

Therefore,

$$\mathfrak{S}(\mathcal{S}^n_i) = \mathfrak{S}(\mathcal{S}^n_i) \cap \mathfrak{B}(\mathcal{S}^n_i) \quad \forall i = 1, \ldots, N$$

Thus, it leads

$$\bigcap_{i=1}^{N} \mathfrak{S}(\mathcal{S}^n_i) = \bigcap_{i=1}^{N} [\mathfrak{S}(\mathcal{S}^n_i) \cap \mathfrak{B}(\mathcal{S}^n_i)] \\ = \bigcap_{i=1}^{N} \mathfrak{S}(\mathcal{S}^n_i) \cap \bigcap_{i=1}^{N} \mathfrak{B}(\mathcal{S}^n_i)$$

which is equivalent to

$$\bigcap_{i=1}^{N} \mathfrak{S}(\mathcal{S}^n_i) \subseteq \bigcap_{i=1}^{N} \mathfrak{B}(\mathcal{S}^n_i)$$

■

Theorem 2. *The intersection $\mathcal{T}$ of N simplex bounding boxes is an axis-aligned box, i.e., there exist n intervals ε_i ($i = 1, \ldots, n$) such that*

$$\mathcal{T} = \bigcap_{i=1}^{N} \mathfrak{B}(\mathcal{S}^n_i) = \prod_{j=1}^{n} \varepsilon_j$$

Proof: Any simplex bounding box $\mathfrak{B}(\mathcal{S}^n_i)$ is the Cartesian product of n intervals, each of which is defined by the minimal and maximal value of the corresponding coordinate for the vertices of the simplex $\mathfrak{S}(\mathcal{S}^n_i)$, namely $b_{ij} = [l_{ij}, u_{ij}]$, i.e.

$$\mathfrak{B}(\mathcal{S}^n_i) = b_{i1} \times b_{i2} \times \cdots \times b_{in} = \prod_{j=1}^{n} b_{ij} \quad \forall i = 1, 2, \ldots, N$$

Therefore,

$$\mathcal{T} = \bigcap_{i=1}^{N} \mathfrak{B}(\mathcal{S}^n_i) = \bigcap_{i=1}^{N} \prod_{j=1}^{n} b_{ij}$$

Thanks to the distributivity property of Cartesian products across intersections, we have

$$\mathcal{T} = \bigcap_{i=1}^{N} \prod_{j=1}^{n} b_{ij} = \prod_{j=1}^{n} \bigcap_{i=1}^{N} b_{ij} = \prod_{j=1}^{n} \varepsilon_j$$

where ε_j is computed as

$$\varepsilon_j = \bigcap_{i=1}^{N} b_{ij} = \left[\max_i \{\inf(b_{ij})\}, \min_i \{\sup(b_{ij})\} \right] \\ = \left[\max_i \{l_{ij}\}, \min_i \{u_{ij}\} \right]$$

Thus,

$$\mathcal{T} = \prod_{j=1}^{n} \varepsilon_j = \prod_{j=1}^{n} \left[\max_i \{l_{ij}\}, \min_i \{u_{ij}\} \right]$$

■

It follows that

$$\bigcap_{i=1}^{N} \mathfrak{S}(\mathcal{S}^n_i) \subseteq \prod_{j=1}^{n} \left[\max_i \{l_{ij}\}, \min_i \{u_{ij}\} \right] = \mathcal{T}$$

B. Bounding-Box Method for Recursive Estimation

With the same fashion proposed in [6], let us keep observing the system (1). Let $x(t_k) = (x_1(t_k), x_2(t_k), \cdots, x_n(t_k))$ be the state of the system at the k-th time instant t_k. Thus, the *observation matrix* at that time instant is

$$X(t_k) = \begin{pmatrix} x(t_{k-m}) \\ \vdots \\ x(t_{k-1}) \\ x(t_k) \end{pmatrix} \in \mathbb{R}^{m \times n} \tag{9}$$

while the *sample covariance matrix* as $\Sigma(t_k) = \frac{X(t_k)^T X(t_k)}{m}$. Thanks to Theorem 1, when $\mathrm{Maxcor}_{\mathcal{F}}(\Sigma(t_k)) = 1$, the space of solutions of the estimation problem $\Sigma(t_k)A = 0$ is the simplex lying in $\mathbb{R}^{n-1}$ whose vertices $A_j(t_k)$ correspond to the n (OLS) solutions with one normalized entry (to 1 or -1):

$$A_j(t_k): \ \mathrm{span}(A_j(t_k)) = \mathrm{null}(\Sigma(t_k) - \tilde{\Sigma}_j(t_k)) \qquad (10)$$

where $\tilde{\Sigma}_j(t_k) = \mathrm{diag}(0, \cdots, 0, \sigma_j^2(t_k), 0 \cdots, 0)$ and

$$\tilde{\sigma}_j^2(t_k) = \frac{\det \Sigma(t_k)}{\det \Sigma_j(t_k)}$$

with $\Sigma_j(t_k)$ obtained by deleting the j-th row and the j-th column of $\Sigma(t_k)$. From now on, we will always choose to normalize the last coefficient α_n to 1 .

The *simplex matrix* of the corresponding simplex of solutions at the k-th iteration is

$$\mathcal{S}^{n-1}(t_k) = \begin{pmatrix} A_1(t_k) & A_2(t_k) & \cdots & A_n(t_k) \end{pmatrix} \in \mathbb{R}^{(n-1)\times n}$$

Furthermore, the corresponding simplex bounding box $\mathfrak{B}(\mathcal{S}^{n-1}(t_k))$ is

$$\mathfrak{B}(\mathcal{S}^{n-1}(t_k)) = \prod_{i=1}^{n} [l_i(t_k), u_i(t_k)] \quad \text{where}$$

$$l_i(t_k) = \min_j \mathcal{S}_{ij}^{n-1}(t_k) \quad \text{and} \quad u_i(t_k) = \max_j \mathcal{S}_{ij}^{n-1}(t_k)$$

As a matter of fact, inside the simplex obtained at the k-th time instant, the value of each parameter α_i will be bounded by $l_i(t_k)$ and $u_i(t_k)$. Indeed, this must true for all the simplices obtained in different time instants, i.e,

$$l_i(t_k) \leq \alpha_i \leq u_i(t_k) \quad \forall k \geq 0, k \in \mathbb{N}$$

Since, the true vector of parameters A is contained in all the simplices $\mathfrak{S}(\mathcal{S}^{n-1}(t_k))$, it is also contained in their intersection and, as ensured by Porposition 1, in the intersection of their bounding boxes:

$$A \in \bigcap_{j=0}^{k} \mathfrak{S}(\mathcal{S}^{n-1}(t_j)) \subseteq \bigcap_{j=0}^{k} \mathfrak{B}(\mathcal{S}^{n-1}(t_j))$$

Although the *Frisch scheme* allows an easy computation of the individual simplices in the parameters space, the computation of $\bigcap_{j=0}^{k} \mathfrak{S}(\mathcal{S}^{n-1}(t_j))$ from the knowledge of the vertices is not trivial. In addition the number of intersection points will keep increase with the number of iterations. Therefore, we cannot setup a recursive *Frisch scheme* method based on seeking a solution inside the intersection of simplices obtained at different iterations.

However, the results shown in Section 2 allow a simple computation of the intersection of simplices bounding boxes. For this reason, in the *bounding box recursive Frisch scheme* we consider as set of solutions the intersection of the bounding boxes rather than the intersection of the simplices.

It must be pointed out that in the proposed method we might include in the solutions set some points which do not lie in the intersection of the simplices, and therefore violate the *Frisch scheme* hypotheses.

C. Updating Rules of the Solutions Set

Definition 5 (Solutions Bounds)**.** *The lower solution bound of each parameter α_i is recursively defined as*

$$\gamma_i(t_k) = \max\{\gamma_i(t_{k-1}), l_i(t_k)\} \quad \forall i = 1, \ldots, n-1 \quad (11)$$

computed at each iteration. Similarly, the upper solution bound of α_i is

$$\mu_i(t_k) = \min\{\mu_i(t_{k-1}), u_i(t_k)\} \quad \forall i = 1, \ldots, n-1 \quad (12)$$

Notice that for any integer $k \geq 0$, and $\forall i = 1, \ldots, n-1$

$$\gamma_i(t_k) = \max_k\{l_i(t_k)\}, \ \mu_i(t_k) = \min_k\{u_i(t_k)\} \qquad (13)$$

Moreover, by definition, γ_i are monotonically increasing and the μ_i are monotonically decreasing, i.e., for all $i = 1, \ldots, n-1$

$$\gamma_i(t_k) \geq \gamma_i(t_{k-1}) \text{ and } \mu_i(t_k) \leq \mu_i(t_{k-1}) \quad \forall k > 0$$

Hence, by denoting with $\lambda([l_i(t_k), u_i(t_k)])$ the *Lebesgue measure* of the interval $[l_i(t_k), u_i(t_k)]$, it holds

$$\lambda([\gamma_i(t_k), \mu_i(t_k)]) \leq \lambda([\gamma_i(t_{k-1}), \mu_i(t_{k-1})]) \ \forall k > 0 \quad (14)$$

Hereafter we will refer to $\mathcal{T}(t_k)$ as the intersection of all the bounding boxes until the k-th time instant and we name it *solutions box*:

$$\mathcal{T}(t_k) = \bigcap_{j=0}^{k} \mathfrak{B}(\mathcal{S}^{n-1}(t_j))$$

$\mathcal{T}(t_k)$ is the *set of solutions* of the recursive *Frisch scheme*.

Proposition 2. *$\mathcal{T}(t_k)$ depends only on the solutions bounds of each parameter at that time instant. In particular,*

$$\mathcal{T}(t_k) = \prod_{i=1}^{n-1} [\gamma_i(t_k), \mu_i(t_k)]$$

Proof: Thanks to Theorem 2 we have

$$\begin{aligned} \mathcal{T}(t_k) &= \bigcap_{j=0}^{k} \mathfrak{B}(\mathcal{S}^{n-1}(t_j)) \\ &= \prod_{i=1}^{n-1} [\max\{l_i(t_0), l_i(t_1), \ldots, l_i(t_k)\}, \\ &\qquad \min\{u_i(t_0), u_i(t_1), \ldots, u_i(t_k)\}] \end{aligned}$$

Thus, from (13), it yields

$$\mathcal{T}(t_k) = \prod_{i=1}^{n-1} [\gamma_i(t_k), \mu_i(t_k)]$$

■

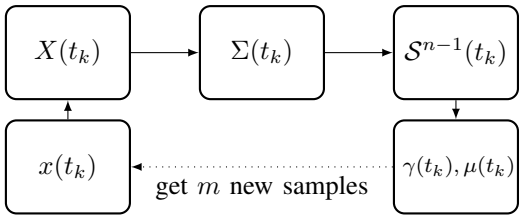

Fig. 1. Schematic representation: block diagram of the overall procedure.

D. Convergence of the Algorithm

It's known that the (hyper)volume of a simplex of solutions can be seen as a measure of the *estimation precision* [5]. In general this criteria can be extended to any solutions set, e.g. to the simplex bounding box. We now want to formalize that, during the execution of the BBRF scheme the volume of the set of solutions is monotonically decreasing.

Theorem 3. *The measure (volume) of $\mathcal{T}$ is monotonically decreasing:*

$$\lambda(\mathcal{T}(t_k)) \leq \lambda(\mathcal{T}(t_{k-1}))$$

Proof: $\mathcal{T}(t_k)$ can be obtained as

$$\begin{aligned}\mathcal{T}(t_k) &= \bigcap_{j=0}^{k} \mathfrak{B}(\mathcal{S}^{n-1}(t_j)) \\ &= \mathfrak{B}(\mathcal{S}^{n-1}(t_k)) \cap \bigcap_{j=0}^{k-1} \mathfrak{B}(\mathcal{S}^{n-1}(t_j)) \\ &= \mathfrak{B}(\mathcal{S}^{n-1}(t_k)) \cap \mathcal{T}(t_{k-1})\end{aligned}$$

Therefore, $\mathcal{T}(t_k) \subseteq \mathfrak{B}(\mathcal{S}^{n-1}(t_k)) \cap \mathcal{T}(t_{k-1})$.
Thanks to the monotonicity property of the Lebesgue measure, it holds

$$\lambda(\mathcal{T}(t_k)) \leq \lambda(\mathcal{T}(t_{k-1}))$$

■

The *measure* of $\mathcal{T}(t_k)$ can be computed as follow

$$\lambda(\mathcal{T}(t_k)) = \prod_{i=1}^{n-1} |\mu_i(t_k) - \gamma_i(t_k)| \tag{15}$$

and, thus, the proof of Theorem 3 comes directly from (14). We have shown that set of solutions can be easily obtained iteratively as it only depends on $\lambda_i(t_k)$ and $\mu_i(t_k)$ which are updated directly from the new simplex matrix. In addition, we have proved that the *size* of the *set of solutions* is monotonically decreasing. However, it is not possible to define a region of convergence of these values, i.e. a lower or upper bounds in the value of $\lambda(\mathcal{T}(t_k))$ as it is totally dependent on the data. A schematic overview of the procedure is reported in Fig. 1.

IV. Experimental Simulations and Discussion

The proposed algorithm has been tested considering the following linear model:

$$\alpha_1 x_1 + \alpha_2 x_2 + x_3 = 0$$

where the following values have been assigned to the parameters: $\alpha_1 = 1.5, \alpha_2 = 1.2$. In a Monte Carlo simulation of

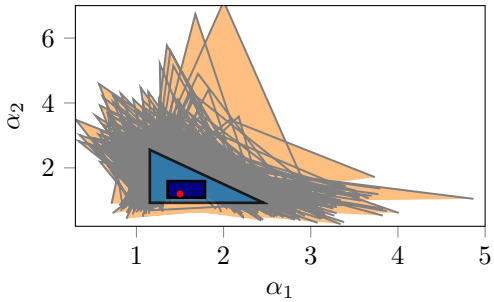

Fig. 2. Comparison between the solution box (dark blue), the total simplex (light) blue and the simplices computed at different itarations (orange). The red dot indicates the position of the true parameters.

150 runs, random sequences of x_1 and x_2 have been generated from independent zero-mean Gaussian distributions with unitary variance, while x_3 has been obtained as $x_3 = -(\alpha_1 x_1 + \alpha_2 x_2)$. Hence, $x(t_k) = (x_1(t_k), x_2(t_k), x_3(t_k))$ for all t_k and $X(t_k)$ has been constructed according to (9). The height of X has been chosen to be $m = 20$, while the total number of samples m_{tot} was $2 \cdot 10^7$, resulting in $N = 10^6$ BBRF iterations. In order to respect the ideal condition of uncorrelation of noise, a diagonal noise covariance matrix $\tilde{\Sigma} = \text{diag}(\tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \tilde{\sigma}_3^2)$, has been added to the sample covariance matrix at each iteration. The noise variances $\tilde{\sigma}_i^2$ $(i = 1, 2, 3)$ has been assigned as $\tilde{\sigma}_i^2 = \delta \cdot \hat{\sigma}_i^2$, where $\hat{\sigma}_i^2$ refers to the variance of the i-th variable and δ has been chosen to be 0.3. The first m samples has been used to initialize the procedure, i.e. $X(t_0)$. The simplex of solutions have been obtained, at each iteration, normalizing to 1 the last entry of each OLS solution, having the prejudice on the value of the parameter related to x_4. Then the solution bounds have been derived as in (11) and (12), having initialized γ_i, μ_i (i = 1,2), with the bounds of the first simplex. In order to compare the BBRF results, all the m_{tot} data points have been used to compute the *total simplex* $\mathfrak{S}_{tot}$. It represents the classical way of performing the *Frisch scheme* estimation on a given set of measurements. The whole procedure was implemented in *Python*[1]. The results of one Monte Carlo run, are pictured in Fig. 2. It can be noticed that the solution box is much smaller than the total simplex or all the simplices computed at different iterations. In Fig. 3, the time evolution of the solution bounds are shown.

The measure (volume) $\lambda_{sol} = \lambda(\mathcal{T}(t_k))$ of the solution space (i.e. the solution box), which is computed according to (15), has been chosen as metric for the evaluation of the performance of the proposed method. This value has been compared with the measure λ_{tot} of $\mathfrak{S}_{tot}$ and the measure λ_k of the simplices obtained at different iterations computed, in both case,s as in (8). The rapidly descent in time of λ_{sol} and its comparison with λ_{tot} and λ_k are shown in Fig. 4.

A. Performance Analysis

The results of the simulations show the reliability of the estimation bounds even before having collected and processed a great amount of data. In fact, as shown in Fig. 4, after about 10 iterations, the size of the solution box becomes

[1]The source code is available at `https://github.com/massastrello/Recursive-Estimation`

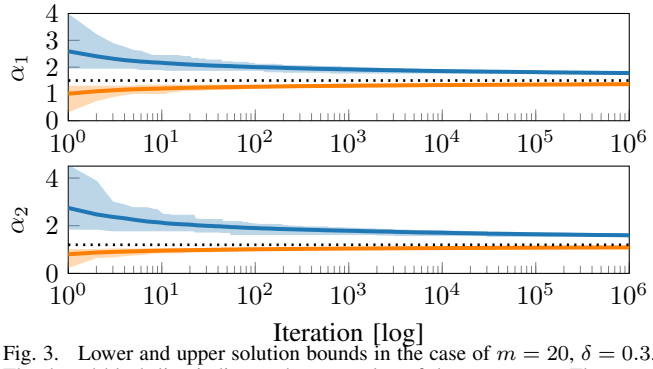

Fig. 3. Lower and upper solution bounds in the case of $m = 20$, $\delta = 0.3$. The dotted black line indicates the true value of the parameters. The mean of all Monte Carlo runs of μ_i is plotted with a solid blue line while the one of γ_i with a solid orange line. The solid lines are bounded by the minimum and maximum values obtained among all the Monte Carlo iterations.

smaller than the one of the total simplex, which is, in average, comparable to the one of the simplices computed at different iterations. Although the convergence of the bounds to specific values cannot be formally proven, the results show an interesting phenomenon. It is possible to notice that the solution bounds are observed to get closer reaching, eventually, a constant values. Consequently and with the same fashion, the λ_{sol} will become constant. Moreover, the BBRF overtakes in terms of computational efficiency the previous recursive Frisch method based on particles. The computational complexity results to be $O(n^3)$ because of the null space extraction in (10). Figure 5 clearly shows how the computational time increases polynomially, i.e. $O(n^3)$. This result has been obtained using a machine equipped with an Intel® Xeon E3-1240v5. These evidences demonstrate that, whenever a set of measurements of a time-invariant linear system is accessible, the BBRF estimator is more accurate than the classical Frisch. Furthermore, thanks to the low computational time and its rapid (observed) convergence, the proposed algorithm can be adopted and implemented for real time applications. Conversely, the BBRF estimation would be inevitably compromised in presence of time varying parameters. In addition, the solution box is redundant with respect to the intersection of all the simplices. This makes the BBRF scheme less accurate than the particle version of the recursive Frisch scheme previous implemented [6]. Other

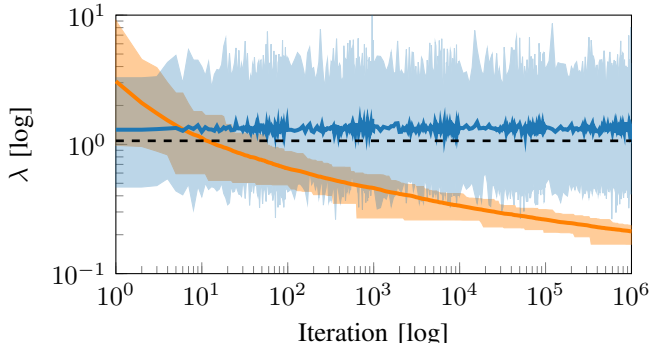

Fig. 4. Evolution in time of the measure of the solution box λ_{sol} (solid orange line), the measure λ_k of each simplex (solid blue line) and the measure λ_{tot} of the total simplex (dashed black line), for $m = 20$, $\delta = 0.3$. The solid lines represent the mean among the Monte Carlo runs of the values and are enclosed by the minimum and maximum values obtained among all the Monte Carlo iterations.

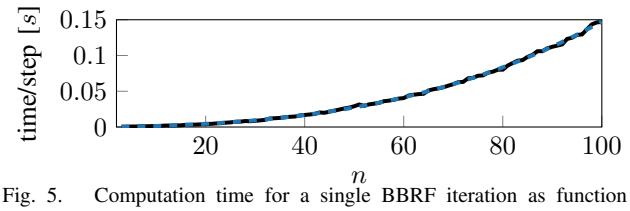

Fig. 5. Computation time for a single BBRF iteration as function of the dimension of the system. The solid black line represents the ground truth while the dashed blue line indicates the fitting obtained with a cubic function.

problems which may rise in practical applications, such as the validity of the assumption of unitary $\mathrm{Maxcor}_{\mathcal{F}}(\Sigma(t_k))$ in each iteration and the correlation of noises affecting different variables, implying a not guaranteed inclusion of the true solution in all the simplices. These problems will be explored in future work.

V. Conclusions And Future Works

This paper presented a novel recursive estimation scheme developed from the *Frisch scheme*, the BBRF scheme. This scheme is an iterative procedure which benefit of the structural properties of simplices bounding boxes to finds a solution set for the identification problem. In order to introduce the concept of *simplex bounding box*, we have firstly analyzed the geometrical properties of the *locus* of solutions of the standard version of the *Frisch scheme*. Then we have presented the BBRF and analytically proved some of its properties. Finally, the effectiveness of the proposed method has been tested on a practical and constructive case of study of relevant interest. Although the method allows a simple recursive computation of the solutions set, there is a consistent redundancy of the *solutions box* with respect to the intersection of all the simplices. In future works, we will formalize a proof of stability of this scheme and explore its extension to the case of time-varying systems.

References

[1] R. E. Kalman, "Identification from real data," in *Current developments in the interface: Economics, Econometrics, Mathematics*. Springer, 1982, pp. 161–196.

[2] R. Frisch, *Statistical confluence analysis by means of complete regression systems*. Universitetets Økonomiske Instituut, 1934, vol. 5.

[3] R.P. Guidorzi, "Identification of the maximal number of linear relations from noisy data," *Systems & Control Letters*, vol. 24, no. 3, pp. 159 – 165, 1995.

[4] R. Guidorzi, "Certain models from uncertain data: the algebraic case," *Systems & control letters*, vol. 17, no. 6, pp. 415–424, 1991.

[5] R. Guidorzi, R. Diversi, and U. Soverini, "The frisch scheme in algebraic and dynamic identification problems," *Kybernetika*, vol. 44, no. 5, pp. 585–616, 2008.

[6] S. Massaroli, R. Miyagusuku, F. Califano, C. Melchiorri, A. Yamashita, and H. Asama, "Recursive algebraic Frisch scheme: A particle based approach," in *Preprints of the 9th IFAC Symposium on Robust Control Design (ROCOND2018)*, 2018, pp. 599–605.

[7] R. Guidorzi, *Multivariable system identification: from observations to models*. Bononia University Press, 2003.

[8] M. Henk, J. Richter-Gebert, and G. M. Ziegler, "16 basic properties of convex polytopes," *Handbook of discrete and computational geometry*, pp. 255–382, 2004.

[9] J. O'Rourke, "Finding minimal enclosing boxes," *International Journal of Parallel Programming*, vol. 14, no. 3, pp. 183–199, 1985.