Adaptive identification of continuous-time MIMO state-space models

Chouaib Afri[1], Laurent Bako[2], Vincent Andrieu[1] and Pascal Dufour[1]

***Abstract*— This paper studies the problem of identifying linear continuous-time state-space models from input-output measurements. An adaptive identifier is developed for the online estimation of both the state and the model parameters in a deterministic framework. Our approach is a non parametric one in the sense that it provides an arbitrary realization of the system. It relies on ideas from the subspace identification literature and adaptive observer.**

I. INTRODUCTION

State-space models provide a nice and compact representation of linear multiple input multiple output (MIMO) systems. This feature probably justifies, at least partially, its tremendous and still increasing success over input-output models e.g., for multi variable control design, fault detection and isolation, performance and stability analysis. From an identification perspective, such models enjoy the advantage that their construction depends on only one structural parameter, namely the so-called McMillan degree of the system. On the other hand, estimating a black-box state-space model from input-output data is not a trivial problem due to the fact that both state and parameters are unknown.

The emergence of subspace identification methods [12], [8], [13] in the 90s has initiated a steady progress in the treatment of this problem. Since then, a number of interesting methods have been developed for efficiently recovering state-space models along with the associated state from a partial observation of the input-output behavior. Nevertheless, such techniques are mainly applied for computing discrete-time models from sampled-data collected in batch mode. A key ingredient of subspace estimation techniques is a rank factorization step, the Singular Value Decomposition (SVD) algorithm. But, since it is computationally expensive, one cannot afford to run it repeatedly in online applications. Therefore, in available extensions of subspace methods to online recursive estimation, the research effort has been essentially devoted to surmounting the need for SVD [11]. Again, apart from a few exceptions [3], this was done essentially for discrete-time models. The subspace approach can be viewed as a non parametric one since it relies on the estimation of an arbitrary basis for a subspace.

A complementary approach is the one of adaptive observers where a state basis is enforced by working on canonical parameterization [9], [14], [1]. While many impressive results have been achieved on this head, it is worth mentioning some drawbacks of canonical representations : (1) they are rarely minimal in the case of MIMO systems unless a great deal of prior knowledge is assumed (e.g. availability of Kronecker observability/controllability indexes [7]) (2) they might be numerically badly conditioned.

The current paper proceeds from the following remark. There is a need for continuous-time models which can be viewed as more natural time-description of physical systems than discrete-time models. One can then argue that since sensing devices can only acquire sampled versions of the system data, it is easier to construct a discrete-time model first and then convert it into a continuous-time one. This is indeed the most developed path of research in system identification. Note however that accuracy of such conversion is dependent of the sampling frequency which therefore must be chosen, when possible, with great care. Moreover, in case the model is to be computed online in an adaptive fashion, the cost of the conversion might not be affordable. An alternative solution is to identify directly a continuous-time model from interpolation of sampled-data [5].

In this paper, we present an adaptive identifier for linear continuous-time MIMO state-space models. The proposed method is capable of estimating online both the state and the system matrices. The paper has two main contributions : (a) a cheap, adaptive and continuous-time identifier for MIMO state-space models which does not involve any use of the costly SVD algorithm ; (b) an exponential convergence analysis under appropriate input sufficiency of richness conditions and model minimality.

The outline of the paper is as follows. In Section I, we underline the problem to deal with and set out the assumptions. The proposed solution idea is presented in Section III. In Section IV, we show that the estimates converge exponentially fast to their respective desired values. We conclude by an application of the algorithm on a linear MIMO system in Section V.

II. LINEAR SYSTEM

Consider a linear system described by a state-space model of the form

$$\mathcal{S}: \quad \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^{n_y}$, $y \in \mathbb{R}^{n_y}$ are respectively the state, the input and the output ; A, B, C, D are matrices of appropriate dimensions.

For future reference throughout the paper, we state the following assumptions :

A1. The system (1) is stable i.e., A is Hurwitz.

A2. (A, B, C) is minimal i.e., (A, B) is controllable and (A, C) is observable.

[1] C. Afri, V. Andrieu and P. Dufour are with the Université de Lyon, F-69622, Lyon, France – Université Lyon 1, Villeurbanne, France – CNRS, UMR 5007, LAGEP, France. afri@lagep.univ-lyon1.fr

[2] L. Bako is with Laboratoire Ampère – Ecole Centrale de Lyon – Université de Lyon, 69130 France. laurent.bako@ec-lyon.fr

A3. C is full row rank.

Problem : Given measurements of the input and output signals, estimate both the state and the parameters in an arbitrary state space basis.

III. Estimation method

A. Main idea

Let $f > n$ be an integer and introduce the notations

$$\begin{aligned} y_f &= \begin{bmatrix} y^\top & \dot{y}^\top & \cdots & y^{(f-1)^\top} \end{bmatrix}^\top \\ u_f &= \begin{bmatrix} u^\top & \dot{u}^\top & \cdots & u^{(f-1)^\top} \end{bmatrix}^\top \end{aligned} \tag{2}$$

and

$$\mathcal{O}_f \triangleq \mathcal{O}_f(A, C) = \begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{f-1})^\top \end{bmatrix}^\top$$

$$\mathcal{T}_f \triangleq \mathcal{T}_f(A, B, C, D) = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{f-2}B & CA^{f-3}B & \cdots & D \end{bmatrix}. \tag{3}$$

Proceeding from (1), we can derive the following equation

$$y_f = \mathcal{O}_f x + \mathcal{T}_f u_f. \tag{4}$$

In practice, the time derivatives of u and y are not measured in general. The only signals that can be reasonably assumed to be available are the input and output. However they are in general noisy and differentiating noisy measurements is well-known to come with some serious drawbacks. One way to avoid numerical computation of derivatives is to work with filtered data. The filter has the further advantage that when it is appropriately chosen, it is capable of attenuating the effects of noise.

Let $F(s) = 1/d(s)$ be a stable filter with $d(s) = s^f + \alpha_{f-1}s^{f-1} + \cdots + \alpha_0$ a polynomial in the Laplace operator s. The filter must be chosen to cut off potential noise. In general, the frequency spectrum of the noise is located at high frequencies. So, if we choose a linear filter with large enough eigenvalues (in absolute value) we can reduce the noise effect while preserving the dynamics of the to-be-identified-system. Note that many other choices of the filter $F(s)$ are possible, see e.g. [5]. The method to be presented is not attached to any specific choice of the filter and so remains applicable for any valid choice [1] of $F(s)$. For this reason, we will not elaborate more on this implementation aspect here.

Posing $p = F(s)y_f$, $q = F(s)u_f$, $\bar{x} = F(s)x$, we obtain

$$\begin{aligned} p &= F(s)(\xi_f(\partial) \otimes I_{n_y})y, \\ q &= F(s)(\xi_f(\partial) \otimes I_{n_u})u \end{aligned} \tag{5}$$

where $\xi_f(\partial) = \begin{bmatrix} 1 & \partial & \cdots & \partial^{f-1} \end{bmatrix}^\top$ with $\partial = \frac{d}{dt}$ denoting the time derivative operator.

1. For a different choice of the filter $F(s)$, one just needs to adjust the equations (6a)-(6b) which generate p and q ; the other steps of the identification method remain unchanged.

We can express (4) and (5) in state-space form as

$$\dot{p} = \left(\Lambda_c \otimes I_{n_y}\right) p + \left(l \otimes I_{n_y}\right) y \tag{6a}$$
$$\dot{q} = \left(\Lambda_c \otimes I_{n_u}\right) q + \left(l \otimes I_{n_u}\right) u \tag{6b}$$
$$\mathcal{O}_f \bar{x} = p - \mathcal{T}_f q \tag{6c}$$

with

$$\Lambda_c = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & \cdots & -\alpha_{f-1} \end{bmatrix}, \quad l = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Note that for any k, the pair $(\Lambda_c \otimes I_k, l \otimes I_k)$ is controllable. Using the filter, p and q can be generated from the measurements using Eqs (6a)-(6b). But we are still facing the difficulty that $\bar{x}$ in Eq. (6c) is unknown. To overcome this we will follow the method proposed in [2].

Since the system is observable, $\mathcal{O}_f$ has full column rank so that there exists a matrix $H \in \mathbb{R}^{n \times f n_y}$ such that $\operatorname{rank}(H\mathcal{O}_f) = n$. Hence $T = H\mathcal{O}_f$ is a square $(n \times n)$ non singular matrix. In fact if H is randomly generated, then the required rank condition is satisfied with probability one. But H can be chosen in many other ways. For example, if the system to be identified is MISO, then observability implies that the n first rows of $\mathcal{O}_f$ are linearly independent so that $H = [I_n \quad 0_{n \times (f-n)}]$ is a valid choice. In the MIMO case, if a set of rows of $\mathcal{O}_f$ indexed by $\{i_1, \ldots, i_n\}$ is known to be linearly independent beforehand, then $H = [e_{i_1} \quad \cdots \quad e_{i_n}]^\top$ with $e_i \in \mathbb{R}^{f n_y}$ the i-th canonical basis vector, fulfills the desired rank condition.

By a change of state coordinates, let us define a new state $\tilde{x}$ as

$$\tilde{x} = T\bar{x}.$$

Accordingly, the system realization becomes $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ with $\tilde{A} = TAT^{-1}$, $\tilde{B} = TB$, $\tilde{C} = CT^{-1}$, $\tilde{D} = D$. Multiplying (6c) by H yields

$$\tilde{x} = Hp - H\mathcal{T}_f q. \tag{7}$$

On the other hand, observe from (6c) that

$$p = \mathcal{O}_f \bar{x} + \mathcal{T}_f q = \tilde{\mathcal{O}}_f \tilde{x} + \mathcal{T}_f q$$

where $\tilde{\mathcal{O}}_f = \mathcal{O}_f T^{-1}$. Inserting (7) into the foregoing expression of p leads to

$$p = M\varphi \tag{8}$$

with

$$M = \begin{bmatrix} \tilde{\mathcal{O}}_f & (I - \tilde{\mathcal{O}}_f H)\mathcal{T}_f \end{bmatrix} \in \mathbb{R}^{f n_y \times (n + f n_u)} \tag{9}$$
$$\varphi = \begin{bmatrix} Hp \\ q \end{bmatrix} \in \mathbb{R}^{n + f n_u}. \tag{10}$$

In (8) the vectors p and φ are known ; the parameter matrix to be estimated is the structured matrix M. It can be usefully remarked that (8) results from (6c) by a multiplication on the left by the projection matrix $I - \tilde{\mathcal{O}}_f H$.

B. Adaptive identifier

From the equation (8) derived above, we can proceed with the estimation of the matrix M. We use the recursive least squares method. At any time t, the estimated parameter matrix is obtained by

$$M(t) = \arg\min_{M} V_t(M)$$

where

$$V_t(M) = \frac{1}{2}e^{-\alpha t}\,\mathrm{tr}\left[(M-M_0)P_0^{-1}(M-M_0)^\top\right] + \frac{1}{2}\int_0^t \frac{e^{-\alpha(t-\tau)}}{r(\tau)^2}\,\|p(\tau)-M\varphi(\tau)\|_2^2$$

where $\alpha > 0$ is a design parameter; $r(\tau)^2 \geq 1$ is a normalizing factor which is upper-bounded, P_0 is a symmetric positive-definite matrix, $M_0 = M(0)$ is the initial value of $M(t)$, tr refers to matrix trace operator. Setting the gradient of $V_t(M)$ to zero and differentiating the resulting solution with respect to time yields the continuous-time recursive least squares estimator :

$$\dot{M} = \frac{(p-M\varphi)\,\varphi^\top P}{r(t)^2}, \quad M(0) = M_0 \tag{11}$$

$$\dot{P} = \alpha P - \frac{P\varphi\varphi^\top P}{r(t)^2}, \quad P(0) = P_0 \succ 0 \tag{12}$$

Here the notation $P_0 \succ 0$ means that P_0 is positive-definite. Simulating (11)-(12) provide an estimate of the matrix M at any time. Now the question to be addressed is how to get back to the system matrices and the state. This can be done by exploiting the structure of M as follows.

Consider partitions of M in the forms

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \\ M_{31} & M_{32} \end{bmatrix},\ M_{11} \in \mathbb{R}^{n_y \times n}, M_{31} \in \mathbb{R}^{n_y \times n},$$

$$M = \begin{bmatrix} M_1 & M_2 \end{bmatrix},\ M_1 \in \mathbb{R}^{f n_y \times n}.$$

Then we can compute the system matrices by the formulas

$$\tilde{C} = M_{11}, \quad \tilde{A} = \begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix}^\dagger \begin{bmatrix} M_{21} \\ M_{31} \end{bmatrix}. \tag{13}$$

Estimation of $\tilde{B}$ and D. We are left with estimating $\tilde{B}$ and D. To this end, note that the system output can be expressed as

$$y(t) = \begin{bmatrix} \tilde{C}e^{\tilde{A}t} & \Psi(t) & u(t)^\top \otimes I_{n_y} \end{bmatrix} \begin{bmatrix} \tilde{x}_0 \\ \mathrm{vec}(\tilde{B}) \\ \mathrm{vec}(D) \end{bmatrix} \tag{14}$$

where the matrix $\Psi(t)$ is defined by

$$\Psi(t) = \int_0^t u(\tau)^\top \otimes \left[\tilde{C}e^{\tilde{A}(t-\tau)}\right] d\tau.$$

Indeed, it is easy to see that Ψ obeys the following differential equation

$$\dot{\Psi}(t) = \Psi(t)\left[I_{n_u} \otimes \tilde{A}\right] + u(t)^\top \otimes \tilde{C}. \tag{15}$$

Hence by using the pair $(\tilde{A}(t), \tilde{C}(t))$ delivered by (13) the matrix $\Psi(t)$ can be generated at any time t.

Writing Eq. (14) in the linear regression form

$$y(t) = \tilde{C}e^{\tilde{A}t}\tilde{x}_0 + \Phi(t)\theta, \tag{16}$$

with

$$\Phi(t) = \begin{bmatrix} \Psi(t) & u(t)^\top \otimes I_{n_y} \end{bmatrix} \tag{17}$$

$$\theta = \begin{bmatrix} \mathrm{vec}(\tilde{B})^\top & \mathrm{vec}(D)^\top \end{bmatrix}^\top, \tag{18}$$

we can derive a recursive least squares algorithm for estimating θ as follows,

$$\dot{\theta} = Q\Phi^\top \frac{(y-\Phi\theta)}{\rho(t)^2}, \quad \theta(0) = \theta_0 \tag{19}$$

$$\dot{Q} = \beta Q - \frac{Q\Phi^\top\Phi Q}{\rho(t)^2}, \quad Q(0) = Q_0 \succ 0 \tag{20}$$

where $\rho(t)^2 = 1 + \mathrm{tr}\left(\Phi(t)Q(t)\Phi(t)^\top\right)$.

Estimation of the state. The state is obtained as

$$\tilde{x} = H(p - \mathcal{T}_f q).$$

This is indeed a filtered version of the transformed state Tx since $\tilde{x} = F(s)(Tx)$. Note that $\tilde{x}$ might be more desirable in practice than the state Tx for the following reason. Since the noise spectrum is generally located in a high frequency range, if the filter $F(s)$ is properly chosen, then $\tilde{x}$ is likely to contain less noise than Tx. But in case one is interested in the state Tx, note that it can be generated by constructing an observer based on the identified realization $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$.

IV. Exponential convergence

In this section, we prove exponential convergence of the recursive identifier (11)-(12) and (19)-(20). To proceed, we need to introduce some specific definitions of persistence of excitation and sufficiency of richness.

***Definition** 1 (persistence of excitation):* A locally integrable signal $\varphi : \mathbb{R}_+ \to \mathbb{R}^d$ is said to be persistently exciting (PE) if there exist constant numbers $\alpha_1, \alpha_2, T > 0$ and $t_0 \geq 0$ such that

$$\alpha_1 I \preceq \int_t^{t+T} \varphi(\tau)\varphi(\tau)^\top d\tau \preceq \alpha_2 I \quad \forall t \geq t_0.$$

Here the notation $S \preceq R$ is used to mean that $R - S$ is positive semi-definite.

Inspired by [10], we also define sufficiency of richness of an input signal as follows.

***Definition** 2 (Sufficiency of richness):* A smooth signal $u : \mathbb{R}_+ \to \mathbb{R}^{n_u}$ is said to be sufficiently rich (SR) of order k if

- $u^{(i)}(0) = 0$ for $i = 0, \dots, k-1$ and
- for any $\gamma > 0$, the signal $\bar{u}_k \in \mathbb{R}^{kn_u}$ defined by

$$\bar{u}_k = \frac{1}{(s+\gamma)^k}\begin{bmatrix} u^\top & \dot{u}^\top & \cdots & (u^{(k-1)})^\top \end{bmatrix}^\top$$

  is PE.

It can be shown that if u is SR of order k, then u is also SR of order ℓ for any $\ell \leq k$.

Now we relate sufficiency of excitation of the input and persistence of excitation of the state under controllability assumption.

***Lemma** 1:* Let $x \in \mathbb{R}^n$ be a signal generated by

$$\dot{x} = Ax + Bu, \quad x(0) = 0 \tag{21}$$

where it is assumed that A is Hurwitz and (A, B) is controllable. Then the signal x is PE if u is sufficiently rich of order m where $m = \deg(m_A(\lambda))$, with $m_A(\lambda)$ denoting the minimal polynomial of A.

Proof: Let $m = \deg(m_A(\lambda))$ and $m_A(\lambda) = \lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0$. Differentiating m times the first equation in (1) yields $x^{(m)} = A^m x + \mathcal{C}_m(A, B)u_m$, where u_m is defined from u using the notational convention (2) and $\mathcal{C}_m(A, B) = [A^{m-1}B \quad \cdots \quad AB \quad B]$. By definition of the minimal polynomial $m_A(A) = 0$ so that $A^m = -c_0 I - c_1 A - \cdots - c_{m-1}A^{m-1} = -(c^\top \otimes I_m)\mathcal{O}_m(A, I)$, where $c = [c_0 \quad \cdots \quad c_{m-1}]^\top$. It follows that

$$x^{(m)} = -(c^\top \otimes I_m)\mathcal{O}_m(A, I)x + \mathcal{C}_m(A, B)u_m$$

On the other hand, we have

$$x_m = \mathcal{O}_m(A, I)x + \mathcal{T}_m(A, B, I, 0)u_m$$

with $x_m = [x^\top \quad \dot{x}^\top \quad \cdots \quad x^{(m-1)}{}^\top]^\top$ and $\mathcal{T}_m(A, B, I, 0)$ defined similarly as in (3). We get after some calculations,

$$x^{(m)} + c_{m-1}x^{(m-1)} + \cdots + c_1\dot{x} + c_0 x = \mathcal{C}_m(\mathcal{K} \otimes I_{n_u})u_m \tag{22}$$

where

$$\mathcal{K} = \begin{bmatrix} 1 & & & & \\ c_1 & 1 & & & \\ c_2 & c_1 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ c_{m-1} & c_{m-2} & \cdots & c_1 & 1 \end{bmatrix}.$$

Let $\gamma > 0$ be an arbitrary number and $\bar{u}_m(t) = 1/(s+\gamma)^m u_m(t)$. By applying the Laplace transform to (22), it follows that [2]

$$x(t) = \frac{(s+\gamma)^m}{m_A(s)}\mathcal{C}_m(\mathcal{K} \otimes I_{n_u})\bar{u}_m(t)$$

Since (A, B) is controllable, $\mathcal{C}_m$ is full row rank. As a consequence, $\mathcal{C}_m(\mathcal{K} \otimes I_{n_u})$ is also full row rank because $\mathcal{K} \otimes I_{n_u}$ is a square non singular matrix. We can therefore apply a similar reasoning as in [10, Theorem 4.1,Theorem 4.2] to conclude. ■

A. Convergence of the (A, C)-estimates

To prove convergence of matrices $\tilde{A}$ and $\tilde{C}$ we must prove the convergence of matrix M. So, let us start with the following proposition.

***Proposition** 1:* Assume that the matrices A and Λ_c from (1) and (6) are Hurwitz. If the pair (A, B) of system (1) is controllable and the input u is SR of order $f + m$, then the signal φ defined in (10) is PE.

2. There is a slight abuse of notation here : time and Laplace variable s are simultaneously used.

Now we state the exponential convergence result of M towards the true value M^o.

***Theorem** 1 (Exponential convergence):* Let Assumptions A1 and A2 hold. If the input u in system (1) is sufficiently rich of order $f + m$, then M defined in (11)-(12) converges exponentially fast to M^o.

Proof: The proof is similar to the proof of Corollary 4.3.2 in [6]. ■

B. Convergence of the (B, D)-estimates

To study the convergence of (19)-(20), we may need some more notations. For the sake of clarity, denote with $(\tilde{A}^o, \tilde{B}^o, \tilde{C}^o, \tilde{D}^o)$ the "true" realization of the system (with respect to the same basis as the estimated one). Indeed such a realization is obtainable from any other "true" minimal realization (A, B, C, D) by the formulas

$$\tilde{A}^o = (H\mathcal{O}_f(A, C))A(H\mathcal{O}_f(A, C))^{-1},\ \tilde{B}^o = (H\mathcal{O}_f(A, C))B,$$
$$\tilde{C}^o = C(H\mathcal{O}_f(A, C))^{-1},\ \tilde{D}^o = D. \tag{23}$$

It must be noted that in spite of the appearance, for a given matrix H, the expressions $(\tilde{A}^o, \tilde{B}^o, \tilde{C}^o, \tilde{D}^o)$ in (23) are independent of any specific realization. Pose $\Phi^o(t) = \Phi(t, \tilde{A}^o, \tilde{C}^o)$, that is, $\Phi^o(t)$ is the matrix arising from (15)-(17) if the true matrices were used for the generation of Ψ. Similarly, let $\Phi(t) = \Phi(t, \tilde{A}(t), \tilde{C}(t))$. By letting $\theta^o = \left[\mathrm{vec}(\tilde{B}^o)^\top \quad \mathrm{vec}(D^o)^\top\right]^\top$, define the estimation error as $e(t) = \theta(t) - \theta^o$, where $\theta(t)$ is defined by (19)-(20). Then

$$\dot{e} = -\frac{Q\Phi^\top\Phi}{\rho^2}e + v \tag{24}$$

with

$$v = \frac{Q\Phi^\top}{\rho^2}\left[\tilde{C}^o e^{\tilde{A}^o t}\tilde{x}_0 + (\Phi^o - \Phi)\,\theta^o\right]. \tag{25}$$

If we can prove that

(i) the homogeneous part of (24) is exponentially stable and

(ii) $v(t)$ vanishes as $t \to \infty$

then, we can conclude on the convergence of the estimation error e towards zero. This argument is supported by the following lemma (see, e.g., [4, chap 4]).

***Lemma** 2:* Consider a time-varying system

$$\dot{z}(t) = A(t)z(t) + v(t)$$

with exponentially stable homogeneous part $\dot{z}(t) = A(t)z(t)$. Assume that $v(\cdot)$ is bounded and $\lim_{t\to\infty} v(t) = 0$. Then $\lim_{t\to\infty} z(t) = 0$.

By a similar reasoning as in the proof of Theorem 1, proving assertion (i) above boils down essentially to showing that Φ is PE, which we do in Lemmas 3 and 4 below.

***Lemma** 3:* Let Assumptions A1–A3 hold for system (1) and let u be sufficiently rich of order m. Then $\Phi^o = \Phi(t, \tilde{A}^o, \tilde{C}^o)$ defined as in (15)-(17) is PE.

Proof: Let $F \in \mathbb{R}^{n \times n_u}$, $G \in \mathbb{R}^{n_y \times n_u}$ be nonzero matrices and let

$$\theta = \left[\mathrm{vec}(F)^\top \quad \mathrm{vec}(G)^\top\right]^\top.$$

We first show that for a given θ with $\|\theta\|_2 = 1$, $w(t) \triangleq \Phi^o(t)\theta$ is PE with lower bound $\alpha_1(\theta) > 0$. The so-defined

w is indeed the output of the following system

$$\begin{cases} \dot{z} = \tilde{A}^o z + Fu, \quad z(0) = 0 \\ w = \tilde{C}^o z + Gu \end{cases} \tag{26}$$

Without loss of generality, we can assume that the realization $(z, \tilde{A}^o, F, \tilde{C}^o, G)$ is minimal [3] with state dimension $d \leq n$. Note that if w is PE when $d = n$, then it will also be PE when $d < n$. It is therefore enough to study the case $d = n$. Arguing as in the proof of Lemma 1, it can be shown that

$$w(t) = \frac{(s+\gamma)^m}{m_A(s)} \Omega(s) \bar{u}_m(t)$$

where

$$\Omega(s) = \tilde{C}^o \mathcal{C}_m(\tilde{A}^o, F)(\mathcal{K} \otimes I_{n_u}) + m_A(s)(e_1^\top \otimes G)$$

with $e_1 = [1 \quad 0 \quad \cdots \quad 0]^\top$. The rest of the proof consists in applying [10, Theorem 4.2]. For this purpose, we must verify that $\Omega(s)$ is full row rank, that is, no γ can be found such that $\gamma^\top \Omega(s) = 0$ for all s. This follows easily from the controllability property of $(\tilde{A}^o, F)$ and the full row rank assumption on $\tilde{C}^o$.

We have hence shown that for any nonzero θ with $\|\theta\|_2 = 1$, $w(t) = \Phi^o(t)\theta$ is PE, that is, there are T, t_1 and $\alpha_1(\theta) > 0$ and $\alpha_2 > 0$ such that

$$\alpha_1(\theta) I \preceq \int_t^{t+T} w(\tau) w(\tau)^\top d\tau \preceq \alpha_2 I \quad \forall t \geq t_1.$$

As a consequence, by taking the trace, we get $\theta^\top R(t)\theta \geq \alpha_1(\theta) n_\theta > 0$ for all $t \geq t_1$, where n_θ is the dimension of θ and

$$R(t) = \int_t^{t+T} \Phi^o(\tau)^\top \Phi^o(\tau) d\tau.$$

On the other hand,

$$\inf_{t \geq t_1} \theta^\top R(t)\theta = \inf_{t \geq t_1} \lambda_{\min}[R(t)] \|\theta\|_2^2 \geq \alpha_1(\theta) n_\theta.$$

It follows that $\inf_{t \geq t_1} \lambda_{\min}[R(t)] > n_\theta \inf_{\theta \neq 0, \|\theta\|_2 = 1} \alpha_1(\theta) > 0$ so that Φ^o is PE. ■

Under the conditions of Theorem 1 we know that $(\tilde{A}(t), \tilde{C}(t))$ converges exponentially fast to $(\tilde{A}^o, \tilde{C}^o)$ as $t \to \infty$. We then call rate of convergence the largest number c such that

$$\sup_{k \in \mathbb{N}} \left\| \tilde{C}(t)\tilde{A}(t)^k - \tilde{C}^o(\tilde{A}^o)^k \right\|_2 \leq \lambda e^{-ct} \tag{27}$$

for a certain constant $\lambda > 0$. Here, the notation $\|\cdot\|_2$ refers to matrix 2-norm (largest singular value).

***Lemma** 4:* Let Assumptions A1 and A2 hold for system (1) and let u be sufficiently rich of order at least $f + m$. If the rate of convergence defined in (27) satisfies $c > 1$, then $\Phi = \Phi(t, \tilde{A}(t), \tilde{C}(t))$ is PE, i.e., there exist $T, \beta_1, \beta_2 > 0$ such that

$$\beta_1 I \preceq \int_t^{t+T} \Phi(\tau)^\top \Phi(\tau) d\tau \preceq \beta_2 I.$$

for any t larger than a certain T_1.

3. Indeed since $F \neq 0$ one can always reduce the considered realization to a minimal one.

We turn now to checking assertion (ii), that is, we question whether v defined in (25) may vanish as $t \to \infty$.

***Lemma** 5:* Under the assumptions of Lemma 4, it holds that $\lim_{t\to\infty} v(t) = 0$.

V. Simulation

As mentioned in Section II our goal here is to give a linear model from the knowledge of outputs and inputs. In other words, the determination of the model (1) matrices A, B, C and D in an arbitrary basis. To illustrate the performance of the proposed identification algorithm, we estimate the MIMO linear system described by the following matrices :

$$A = \begin{bmatrix} -1.4 & 0.3 & -0.6 & -0.3 \\ 0.3 & -0.9 & -0.9 & 0.1 \\ -0.6 & -0.9 & -2.6 & -1.0 \\ -0.3 & 0.1 & -1.0 & -1.8 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 \\ -0.4 & 1.3 \\ 0.6 & 0.2 \\ 0 & 0.6 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.9 & 0 & 0 & 1.3 \\ -1.1 & -0.8 & 0.9 & 0.4 \end{bmatrix}, D = \begin{bmatrix} 0.8 & -1.5 \\ 0 & 0 \end{bmatrix}.$$

A white noise was selected as an input signal $u(t)$ which is sufficiently rich in frequency and the algorithm initial values are given as : $M(0) = 0$, $\theta(0) = 0$, $P(0) = I$, $\psi(0) = 0$ and $Q(0) = I$.

We recall that the algorithm is composed of two parts. The first part is given by Eqs (6a)-(6b) and (11)-(12). The second part is given by Eqs (15), (19) and (20) ; this latter estimates $(\tilde{B}, D)$-matrices once the first part of the algorithm has converged. The matrix H used to calculate the vector φ is generated in a random way.

The state basis of the identified realization is not necessarily the same as that of the real system. It follows that the estimated matrices are not numerically identical but have common characteristics such as eigenvalues that are left unchanged under a similar transformation. Fig. 1 shows the convergence of the matrix $\tilde{A}$ eigenvalues to those of the true matrix A. It can be seen that estimated and true eigenvalues coincide perfectly in the noise-free scenario. Fig. 2 is the bode diagram of the transfer function between the first output and the first input ; it shows that the estimated frequency spectrum converges perfectly to the real one and the same can be seen for the phase diagram. The expression $\tilde{O}_f\tilde{B} = O_f B$ is also invariant through a similar transformation and is used to compute the error norm shown in Fig.3 and 4. The relative error norm

$$\varepsilon_r = \frac{\|\tilde{O}_f\tilde{B} - O_f B\|_F}{\|O_f B\|_F}$$

vanishes in Fig.3 which means that the first f Markov parameters of the system are well estimated and in Fig.4 it differ a little in the presence of a moderate amount of noise.

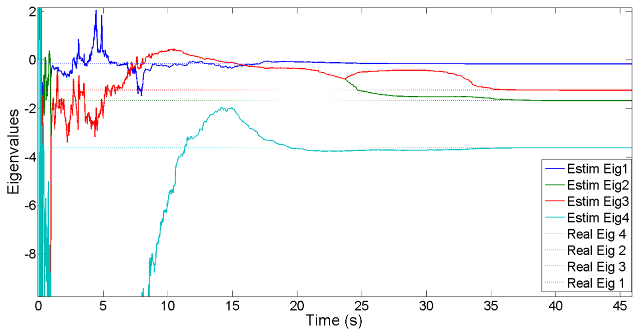

Fig. 1. $\tilde{A}$ eigenvalues convergence without output noise ($\alpha = 4.10^{-1}$, $\beta = 8.10^{-2}$).

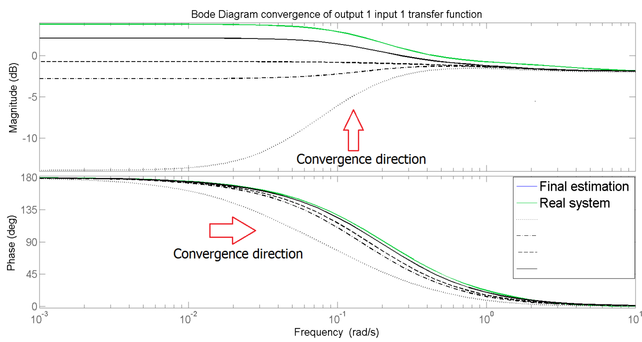

Fig. 2. Bode diagram of output 1 input 1 transfer function

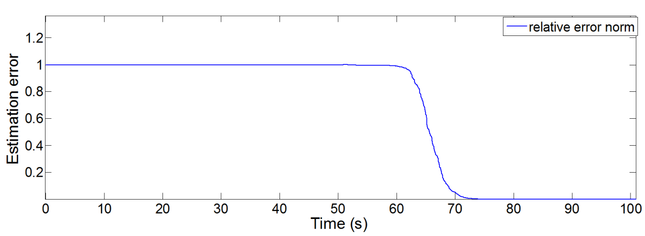

Fig. 3. Error norm without noise ($\alpha = 4.10^{-1}$, $\beta = 8.10^{-2}$).

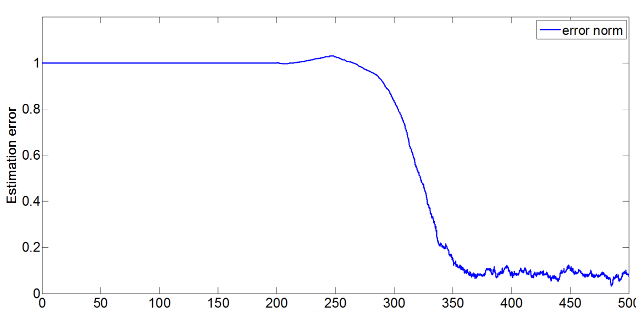

Fig. 4. Error norm ε_r with output noise of 50 dB ($\alpha = 10^{-2}$, $\beta = 8.10^{-3}$).

VI. Conclusion

In this paper, we have developed a new identification algorithm for adaptively identifying linear continuous-time MIMO systems in state-space representation. The main challenges with such a problem are (i) estimate time derivatives of input-output signals (ii) provide a cheap update process for the system matrices while the state is unknown (iii) analyze convergence of the whole algorithm. The contribution of the paper lies essentially in the last two points. Simulation results tend to show that the algorithm performs very well in a deterministic context. However further work might be needed to strengthen its robustness to noise. In particular, the H-matrix based projection step can probably be improved by choosing the matrix H to be the transpose of an off-line prior estimate of the extended observability matrix $\mathcal{O}_f$. Also, there is still a room for improvement concerning the filtering of the input-output signals.

References

[1] C. Afri, V. Andrieu, L. Bako, and P. Dufour. Identification of linear systems with nonlinear luenberger observers. In *IEEE American Control Conference, Chicago*, accepted, 2015.

[2] L. Bako, G. Mercère, and S. Lecoeuche. A least squares approach to the subspace identification problem. In *IEEE Conference on Decision and Control, Cancun, Mexico*, 2008.

[3] M. Bergamasco and M. Lovera. Continuous-time predictor-based subspace identification using laguerre filters. *IET Control Theory and Applications*, 5 :856–867, 2011.

[4] R. W. Brockett. *Finite dimensional linear systems*. John Wiley & Sons, 1970.

[5] H. Garnier and L. Wang (Eds). *Identification of Continuous-time Models from Sampled Data*. Springer, 2008.

[6] Petros Ioannou and Jing Sun. *Robust Adaptive Control*. Dover Publications, 2012.

[7] T. Kailath. *Linear systems*. Prentice Hall, 1979.

[8] Tohru Katayama. *Subspace Methods for System Identification*. Springer London Ltd, 2005.

[9] G. Kreisselmeier. Adaptive observers with exponential rate of convergence. *IEEE Transactions on Automatic Control, 22, pp 2-8*, 1977.

[10] I. M. Y. Mareels and M. Gevers. Persistency of excitation criteria for linear, multivariable, time-varying systems. *Mathematics of Control, Signals and Systems*, (1) :203–206, 1988.

[11] G. Mercère, L. Bako, and S. Lecœuche. Propagator-based methods for recursive subspace model identification. *Signal Processing 88 (3), 468-491*, 2008.

[12] P. van Overschee and B. De Moor. *Subspace Identification for Linear Systems : Theory - Implementation - Applications*. Springer, 1996.

[13] M. Verhaegen and V. Verdult. *Filtering and System Identification : A Least Squares Approach*. Cambridge University Press, 2007.

[14] Q. Zhang. Adaptive observer for multiple-input multiple-output (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control*, 2002, 47,525-529.