# Nonlinear Blind Parameter Estimation

[1,3]S. Lagrange [1]L. Jaulin [3]V. Vigneron [3]C. Jutten

[1]LISA, UFR Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France.

[2]$E^3I^2$, ENSIETA, 2 rue François Verny, 29806 Brest Cédex 09, France.

[3]LIS, INPG, 46 avenue Félix Viallet, 38031 Grenoble cedex, France.

**Abstract**

This paper deals with parameter estimation of nonlinear continuous-time models when the input signals of the corresponding system are not measured. The contribution of the paper is to show that, with simple priors about the unknown input signals and using derivatives of the output signals, one can succeed the estimation procedure. As an illustration, we will consider situations where the simple priors, *e.g.* independence or Gaussianity of the unknown inputs, is assumed.

## I. INTRODUCTION

Consider the *invertible models with a relative degree* $r$,

$$\mathbf{u}(t) = \psi_{\mathbf{p}}\left(\mathbf{y}(t), \dot{\mathbf{y}}(t), \ldots, \mathbf{y}^{(r-1)}(t), \mathbf{y}^{(r)}(t)\right), \tag{1}$$

where $t \in \mathbb{R}$ is the time, $\mathbf{u}(t) \in \mathbb{R}^m$ is the unknown (*i.e.* unobserved) input vector signal, $\mathbf{y}^{(i)}(t) \in \mathbb{R}^m$ is the $i$th derivative of the observed output vector signal $\mathbf{y}(t)$, $\mathbf{p} \in \mathbb{R}^n$ is the unknown parameter vector and $\psi_{\mathbf{p}}$ is an analytical parametric function from $(\mathbb{R}^m)^{r+1}$ to $\mathbb{R}^m$. Since all flat models [11] should satisfy (1), the class of invertible models is rather large.

*Example 1:* Consider the car represented on Figure 1. This system has two inputs : the speed $v$ of the front wheels and the angle $\delta$ of the front wheels with respect the body of the car. The outputs, denoted $(x, y)$, are the coordinates of the middle of the back axle. The state space equations of this system are given by

$$\left\{\begin{array}{llll} \text{(i)} & \dot{x} & = & v\cos\delta\cos\theta \\ \text{(ii)} & \dot{y} & = & v\cos\delta\sin\theta \\ \text{(iii)} & \dot{\theta} & = & \frac{v\sin\delta}{L} \end{array}\right. .$$

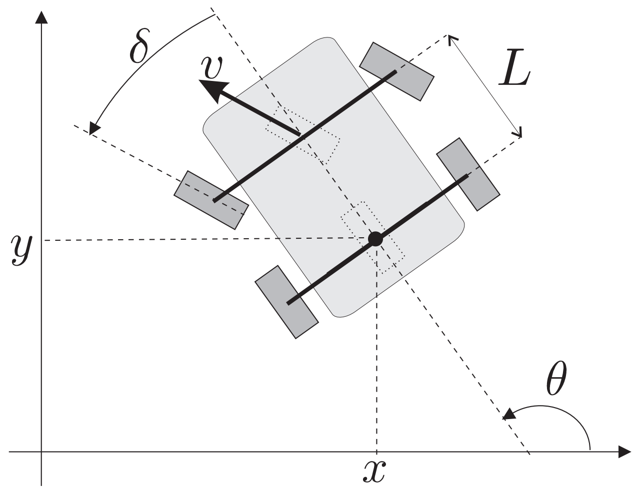


Fig. 1. Car representation in the plane.

where the parameter $L$ represents the distance between the front and the back axles. From (i) and (ii), we get $\theta = \arctan\left(\frac{\dot{y}}{\dot{x}}\right)$ and after derivation, we get $\dot{\theta} = \frac{1}{\left(\frac{\dot{y}}{\dot{x}}\right)^2+1}\left(\frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^2}\right)$. Since

$$\begin{aligned} v\cos\delta &= \frac{\dot{x}}{\cos\arctan\left(\frac{\dot{y}}{\dot{x}}\right)} && \text{(see Eq (i))} \\ v\sin\delta &= L\left(\frac{1}{\left(\frac{\dot{y}}{\dot{x}}\right)^2+1}\left(\frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^2}\right)\right) && \text{(see Eq (iii))} \end{aligned}$$

we get

$$\begin{aligned} v &= \sqrt{\left(L\left(\frac{1}{\left(\frac{\dot{y}}{\dot{x}}\right)^2+1}\left(\frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^2}\right)\right)\right)^2 + \left(\frac{\dot{x}}{\cos\arctan\left(\frac{\dot{y}}{\dot{x}}\right)}\right)^2} \\ \delta &= \arctan\left(\frac{L\left(\frac{1}{\left(\frac{\dot{y}}{\dot{x}}\right)^2+1}\left(\frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^2}\right)\right)}{\frac{\dot{x}}{\cos\arctan\left(\frac{\dot{y}}{\dot{x}}\right)}}\right) \end{aligned}$$

The car is thus an invertible system. Other invertible models can be found in [17].

The parameter estimation problem [24] in a blind context consists in estimating the unknown parameter vector $\mathbf{p}$ by only exploiting the observed signals $\mathbf{y}$ [1]. One only assumes weak statistical assumptions on the unknown input signals, *e.g.* independency or Gaussianity, and the model $\psi_{\mathbf{p}}$ is known (except its parameters).

For instance, consider Example 1 where only the outputs $(x(t), y(t))$ are measured (for instance using a GPS localization system). If assuming that the unknown inputs $(v(t), \delta(t))$ are Gaussian (or independent) is realistic, then the parameter $L$ could be obtained from the knowledge of the

outputs without any other measurements.
For this king of borderline problem where priors knowledge is poor, our goal is to prove that the estimation of $\mathbf{p}$ is possible.

### A. Signal assumptions

We assume that the input vector signal $\mathbf{u}$ belongs to $\mathcal{S}^m$, where $\mathcal{S}$ denotes the set of all *stationary, ergodic and smooth random signals i.e.*, whose the $i$th derivative with respect to $t$ is defined for any $i \in \mathbb{N}$. As a consequence, $\dot{\mathbf{u}}, \ddot{\mathbf{u}}, \dddot{\mathbf{u}}, \dots$ also belong to $\mathcal{S}^m$. Remark that white noise is not a suitable input since it is not differentiable [20].
Note that this assumption does not imply that $\mathbf{y}$ belongs to $\mathcal{S}^m$. For instance, when the model is unstable, $\mathbf{y}$ is not stationary and cannot belong to $\mathcal{S}^m$.

We also assume that $\mathbf{u}$ satisfies a few statistical properties, that can be described by statistical moments. A *generalized moment* $\mu$ (*or moment* for short) of $\mathbf{u} \in \mathcal{S}^m$, is a function from $\mathcal{S}^m$ to $\mathbb{R}$ which can be written as $\mu(\mathbf{u}) = E(u_{j_1}^{(i_1)} u_{j_2}^{(i_2)} \dots u_{j_s}^{(i_s)})$, where $s \geq 1$, $j_1, \dots, j_s \in \{1, \dots, m\}$ is the input index and $i_1, \dots, i_s \in \mathbb{N}$ is the derivative order. The set of all moments will be denoted by $\mathcal{M}$. The integer $s$ is called the order of $\mu$. For instance $E\left(\dot{u}_1^3 \ddot{u}_2\right) = E(\dot{u}_1 \dot{u}_1 \dot{u}_1 \ddot{u}_2)$ is a moment of $\mathcal{S}^2$ with order $4$ (where $j_1, \dots, j_3 = 1$ and $j_4 = 2$). To be consistent with the literature, when $s \geq 3$, $\mu$ will be said to be of higher (than 2) order.

### B. Estimating functions

An *estimating function* [4] is a function from $\mathcal{S}^m$ to $\mathbb{R}^q$, the components of which are functions of generalized moments belonging to $\mathcal{M}$. For instance, the function

$$h : \begin{cases} \mathcal{S}^2 & \rightarrow \quad \mathbb{R} \\ (u_1, u_2) & \rightarrow \quad E(u_1 u_2) - E(u_1) E(u_2) \end{cases} \tag{2}$$

is an estimating function. In this paper, these functions will be designed in order to vanish when some statistical assumptions on the inputs are satisfied. For example, if the signals $u_1$ and $u_2$ are assumed to be decorrelated, the estimating function (2) could be used.

We are now able to formalize the blind parameter estimation problem to be considered in this paper as follows. Given a parametric model $\mathbf{u} = \psi_{\mathbf{p}}\left(\mathbf{y}, ..., \mathbf{y}^{(r)}\right)$, where $\mathbf{y}$ is measured whereas

$\mathbf{u}$, assumed to belong to $\mathcal{S}^m$, is unknown, and an estimating function $\mathbf{h}$, the blind parameter estimation problem consists in characterizing the set

$$\mathbb{P} = \left\{ \mathbf{p} \in \mathbb{R}^n \mid \mathbf{h}\left(\psi_{\mathbf{p}}\left(\mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)\right) = \mathbf{0} \right\}. \tag{3}$$

Estimating $\mathbb{P}$ is then equivalent to solve a set of nonlinear equations in $\mathbf{p}$.

### C. *Content of the paper*

In this paper, we propose a new methodology for solving the blind estimation of nonlinear parametric invertible models. However, defining the rigorous conditions, for which the methodology will succeed remains beyond the scope of this paper. Moreover, we shall assume that the measured signals are noiseless and all derivatives $\mathbf{y}^{(i)}, i \geq 0$ of $\mathbf{y}$ are available[1]. In practice, this is never the case, but, even in such an ideal context, there is no general method for blindly estimating $\mathbf{p}$.

Note that the estimation of $\mathbf{p}$ leads to the knowledge of $\mathbf{u}$, via the relation (1). Thus, the Blind Parameter Estimation (BPE) is very closed to Blind Source Separation (BSS) (or Independent Components Analysis (ICA) when the input are assumed to be independent) which consists in estimating the unknown input signals [15][7][6][2][5][12], using simple statistical priors. Indeed, unlike BSS (or ICA) which recovered the inputs with some indeterminacies [8], BPE can be viewed as ”perfect” ICA where no indeterminacies is allowed on the inputs.

This paper is organized as follows. Section 2 introduces definitions and properties of random signals and of their derivatives. Section 3 shows how these properties can be used for solving the blind estimation problem. A simple example is proposed in Section 4, before the conclusion.

## II. Some results related to derivatives of random signals

In the blind parameter estimation problem, one only assumes that statistical features (such as independence or Gaussianity) are satisfied by the unknown input signals $\mathbf{u}$. These conditions can be written into a set of equations involving generalized moments of $\mathbf{u}$. In order to be able to write these equations, we first introduce a few properties satisfied by generalized moments. These properties will then be used to build suited estimation functions $\mathbf{h}$ in Section 3.

[1]The estimation of the derivatives $\mathbf{y}^{(i)}$ from $\mathbf{y}$ is not addressed in this paper. The reader interested by this topic is referred to [23][10][19][9]

### A. *Statistical independence*

This Subsection introduces some properties verified by the derivatives of independent random signals [21]. First, let us present the definition of statistically independent signals.

*Definition 2:* Denote $\mathcal{F}$ the set of all functions defined from $\mathcal{S}$ to $\mathbb{R}$. The random signals $u_1, \ldots,\ u_m$ of $\mathcal{S}^m$ are statistically independent if $\forall f_1, f_2, \ldots, f_m \in \mathcal{F}$, the random variables $x_1 = f_1\left(u_1\right), \ldots, x_m = f_m\left(u_m\right)$ are statistically independent, i.e.,

$$p_{x_1,\ldots,x_m}(.) = \prod_{i=1}^{m} p_{x_i}(.), \tag{4}$$

where $p_{x_i}(.)$ and $p_{x_1,\ldots,x_m}(.)$ denote the marginal and joint probability density functions of the random variables $x_i$ and of the random vector $(x_1, \ldots, x_m)^T$ respectively.

The following proposition gives relationship between independent signal and their derivatives that will be exploited in next Section.

*Proposition 3:* Let $u_1, \ldots, u_m$ and $y$ be random signals of $\mathcal{S}$. If $u_1, \ldots, u_m$ are independent then $\forall\left(k_1, \ldots, k_m\right) \in \mathbb{N}^m$, $u_1^{(k_1)}, \ldots, u_m^{(k_m)})$ are independent.

*Proof:* Suppose that the signals $u_1, \ldots, u_m$ are independent. Take $m$ functions $f_1, \ldots, f_m \in \mathcal{F}$ and define

$$g_i : \begin{cases} \mathcal{S} & \rightarrow \quad \mathbb{R} \\ u_i & \rightarrow \quad f_i(u_i^{(k_i)}) \end{cases}, \quad i = 1, \ldots, m. \tag{5}$$

From Definition 2, since $u_1, \ldots, u_m$ are independent, the variables $g_1\left(u_1\right), \ldots, g_m\left(u_m\right)$ are independent, *i.e.* the random variables $f_1\left(u_1^{(k_1)}\right), ..., f_m\left(u_m^{(k_m)}\right)$ are independent. Since no assumption has been made on $f_i$, we get, for all functions $f_1, \ldots, f_m \in \mathcal{F}$, independent variables $f_1\left(u_1^{(k_1)}\right), \ldots, f_m\left(u_m^{(k_m)}\right)$. Therefore, the signals $u_1^{(k_1)}, \ldots, u_m^{(k_m)}$ are independent. ■

### B. *Gaussianity*

Let us introduced the definition of Gaussian random signals in order to establish properties of their derivatives.

*Definition 4:* Denote by $\mathcal{L}$ the set of all linear forms defined from $\mathcal{S}^m$ to $\mathbb{R}$. The random vector signal $\mathbf{u} = (u_1, \ldots, u_m)$ is said to be Gaussian if for all $f \in \mathcal{L}$, the random variable $f(\mathbf{u})$ is Gaussian.

The following proposition shows that the differentiation preserves the Gaussianity.

*Proposition 5:* If $\mathbf{u} \in \mathcal{S}^m$ is a Gaussian random signal then its derivatives $\mathbf{u}^{(k)}$ $(\forall k = 1, 2, \dots)$ are Gaussian.

*Proof:* According to Definition 4, the signal $\mathbf{u}^{(k)}$ is Gaussian if (and only if), for all $f \in \mathcal{L}$, the random vector $f(\mathbf{u}^{(k)})$ is Gaussian.
Now, since $\mathbf{u}$ is Gaussian, and since

$$g : \begin{cases} \mathcal{S}^m & \rightarrow & \mathbb{R} \\ \mathbf{u} & \rightarrow & f(\mathbf{u}^{(k)}) \end{cases} \tag{6}$$

is linear, $g(\mathbf{u})$ is Gaussian, *i.e.* $f(\mathbf{u}^{(k)})$ is Gaussian. Thus, since no assumption has been made on $f$, we get, for all $f \in \mathcal{L}$, $f(\mathbf{u}^{(k)})$ is Gaussian. Therefore, $\mathbf{u}^{(k)}$ is Gaussian. ■

## III. Estimating functions

This section points out how estimating functions $\mathbf{h}$ could be built when statistical assumptions on the input random signals $\mathbf{u}$ are available. First, we shall consider the case where inputs are assumed to be independent. Then, the case where a few inputs are assumed to be Gaussian will be treated.

### A. Case of independent inputs

Let us introduce some statistical properties of generalized moments of random vector signal $\mathbf{u} = (u_1, \dots, u_m) \in \mathcal{S}^m$.

*Proposition 6:* Define

$$m_{i,j}^{(k)(\ell)}(\mathbf{u}) = E(u_i^{(k)} u_j^{(\ell)}) - E(u_i^{(k)}) E(u_j^{(\ell)}), k \geq 0, \ell \geq 0, i, j \in \{1, \dots, m\}, \tag{7}$$

where $\mathbf{u} = (u_1, \dots, u_m) \in \mathcal{S}^m$, we have

$$(i) \quad k + \ell \geq 1 \quad \Rightarrow m_{i,j}^{(k)(\ell)}(\mathbf{u}) = E(u_i^{(k)} u_j^{(\ell)}), \tag{8}$$

$$(ii) \quad k_1 + \ell_1 = k_2 + \ell_2 \quad \Rightarrow \quad m_{i,j}^{(k_1)(\ell_1)}(\mathbf{u}) = (-1)^{\ell_1 - \ell_2} m_{i,j}^{(k_2)(\ell_2)}(\mathbf{u}). \tag{9}$$

$$(iii) \quad m_{i,j}^{(k)(\ell)}(\mathbf{u}) = (-1)^k m_{j,i}^{(k)(\ell)}(\mathbf{u}), \tag{10}$$

*Proof:* The proof can be done considering that

$$E\left(u_i^{(k)}\right) = 0, \forall k \geq 1, \tag{11a}$$

$$E\left(u_i^{(k)}(t)\, u_j^{(\ell)}(t - \tau)\right) = (-1)^\ell \frac{d^{k+\ell} \gamma_{u_i u_j}(\tau)}{d^{k+\ell} \tau}, \tag{11b}$$

where $\gamma_{u_i u_j}(\tau)$ is the correlation function between $u_i$ and $u_j$ (see [20]). ■

Assume that $I_s(u_1, \ldots, u_m)$. Then, for all $k \geq 0$ and $l \geq 0$, we have

$$m_{i,j}^{(k)(\ell)}(\mathbf{u}) = 0 \text{ if } i \neq j. \tag{12}$$

However, the Proposition 6 (ii) and (iii) points out relationships between moments. As a consequence, in order to cancel the redundant terms, we can choose the following estimating function:

$$\mathbf{h} : \begin{cases} \mathcal{S}^m & \rightarrow & \mathbb{R}^{\frac{(m-1)m}{2}\frac{(q+1)q}{2}} \\ \mathbf{u} & \rightarrow & \left(m_{i,j}^{(k)(\ell)}(\mathbf{u})\right)_{i=1,\ldots,m;\ j=i+1,\ldots,m}^{k=0,\ldots,q\ ;\ \ell=k,\ldots,q} \end{cases} \tag{13}$$

where the integer number $q$ is the maximum derivative order of the input signal. Practically, $q$ must be chosen so that the number of estimating equations is equal or larger to the unknown number. For $q$ large, one has a large number of estimating functions with respect to the unknown number and one can hope to achieve a robust resolution, but with a higher computational cost.

Note that in the estimating functions (13), we only took second order moments. Of course, higher order moments could also be used.

### *B. Case of Gaussian inputs*

Assume now the inputs are Gaussian, the divergence to Gaussianity can be measured using the Kurtosis, which is equal to zero for Gaussian random variables. Moreover, according to Proposition 5 and Equality (11a), the input derivatives are centered and still Gaussian. As a consequence, the following estimating function could be a good candidate for measuring the (divergence to) Gaussianity of the signal $\mathbf{u}$.

$$\mathbf{h} : \begin{cases} \mathcal{S}^m & \rightarrow & \mathbb{R}^{q-1} \\ \mathbf{u} & \rightarrow & \begin{pmatrix} E\left([u-E(u)]^4\right) - 3E\left([u-E(u)]^2\right)^2 \\ E(\dot{u}^4) - 3E\left(\dot{u}^2\right)^2 \\ \vdots \\ E\left(\left[u^{(q)}\right]^4\right) - 3E\left(\left[u^{(q)}\right]^2\right)^2 \end{pmatrix} \end{cases}. \tag{14}$$

As previously, the integer $q$ (which is the maximum derivative order) must be chosen so that the number $n$ of estimating equations (components of $\mathbf{h}$) is equal or greater than the unknown number.

*Remark 7:* If the invertible system is linear (*i.e.* $\mathbf{u} = \sum_{i=0}^{i=r} \mathbf{A}_i \mathbf{y}^{(i)}$) and the inputs are known to be Gaussian then, the outputs $\mathbf{y}$ are also Gaussian [20]. Thus, all matrices $\mathbf{A}_i$ should be considered as feasible, since they lead to a Gaussian input signal $\mathbf{u}$. All parameters (the entries of $\mathbf{A}_i$) should thus be considered as acceptable and the blind estimation is impossible without priors. In fact, additional information, like non-stationarity [22] or temporal correlation [3] of sources could be used in this context like in blind source separation.

### C. Approximating the estimating function

The set $\mathbb{P}$ is defined by the following vector equations (see Equation (3)) :

$$\mathbf{g}(\mathbf{p}) \stackrel{\text{def}}{=} \mathbf{h}(\psi_{\mathbf{p}}\left(\mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)) = \mathbf{0}. \tag{15}$$

In practice, it is not possible to get an analytical expression for $\mathbf{g}$ since only signal expectations can be estimated. We shall now explain, on a simple example, how an empirical estimate of $\mathbf{g}$ can be obtained. Take for instance the scalar model described by

$$u = \psi_p\left(y, \dot{y}\right) = \dot{y} + p \sin y, \tag{16}$$

and assume that the estimating function is chosen as $h(u) = E(u\dot{u})$. Since $u$ is stationary,

$$\begin{aligned} g(p) &= h(\psi_p\left(y, \dot{y}\right)) \\ &= h(\dot{y} + p \sin y) \\ &= E\left((\dot{y} + p \sin y)(\ddot{y} + p\dot{y} \cos y)\right) \\ &= E\left(\dot{y}\ddot{y} + p\dot{y}^2 \cos y + p\ddot{y} \sin y + p^2 \dot{y} \sin y \cos y\right). \end{aligned}$$

On the other hand, define the empirical estimator $\hat{E}\left(x\right)$ for $E(x)$ of a signal as

$$\hat{E}(x) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{k=0}^{N} x(k\tau), \tag{17}$$

where $N$ is the number of available samples and $\tau$ is the sampling period. Thus $g(p)$ can be approximated by

$$\begin{aligned} \hat{g}(p) &= \hat{E}\left(\dot{y}\ddot{y} + p\dot{y}^2 \cos y + p\ddot{y} \sin y + p^2 \dot{y} \sin y \cos y\right) \qquad (18) \\ &= \hat{E}\left(\dot{y}\ddot{y}\right) + p\hat{E}\left(\dot{y}^2 \cos y\right) + p\hat{E}\left(\ddot{y} \sin y\right) + p^2 \hat{E}\left(\dot{y} \sin y \cos y\right). \end{aligned}$$

The function $g(p)$ is thus approximated by the second degree polynomial $\hat{g}(p)$, the coefficient of which are computed from the knowledge of the signal $y(t)$.

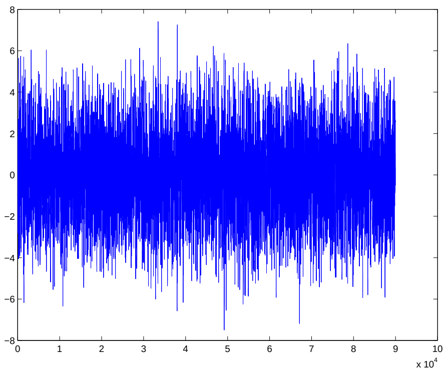

Fig. 2. Signal $y_1(t)$ (90000 samples)

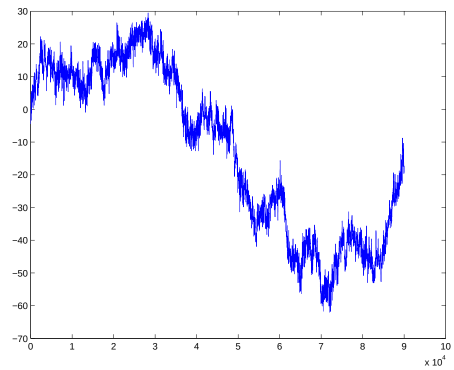

Fig. 3. Signal $y_2(t)$ (90000 samples)

## IV. A SIMPLE EXAMPLE

In this section, we consider a simple example, where the input and observed signals are related to an inverse model. We shall assume the statistical independence of the inputs and we will design an estimating function in order to estimate the unknown parameters.

Consider two independent colored Gaussian signals $u_1$ and $u_2$ of $\mathcal{S}$ obtained as a filtering of a white Gaussian noise ($N$ samples). These two signals feed a parametric nonlinear model described by the two following relations:

$$\left\{ \begin{array}{l} u_1 = \frac{1}{p_1}\dot{y}_1 + \frac{p_1}{p_2}\cos y_2 - y_1 \\ u_2 = \dot{y}_2 + p_1^2 p_2 \sin y_1, \end{array} \right. \tag{19}$$

where the two parameters are $p_1^* = -2$ and $p_2^* = -0.9$. Since this model is unstable, $\mathbf{y}$ is not stationary (see Figures 2 and 3) even if $\mathbf{u}$ is.

Assume that $u_1$ and $u_2$ are independent. No other assumption are taken into account: especially, the Gaussianity is not exploited. According to Subsection III-A, we choose the following estimating function $\mathbf{h}$:

$$\mathbf{h} = \begin{pmatrix} E\left(u_1 u_2\right) - E\left(u_1\right) E\left(u_2\right) \\ E\left(\dot{u}_1 u_2\right) \end{pmatrix} \tag{20}$$

which provides two equations (as many as the number of parameters); $h(u_1, u_2)$ vanishes when $u_1$ and $u_2$, and $\dot{u}_1$ and $u_2$ become uncorrelated.

The condition $\mathbf{h}(\mathbf{u}) = \mathbf{0}$ can be approximated by

$$\left\{ \begin{array}{rcl} \hat{E}\left(u_1 u_2\right) - \hat{E}\left(u_1\right)\hat{E}\left(u_2\right) & = & 0 \\ \hat{E}\left(\dot{u}_1 u_2\right) & = & 0 \end{array} \right. \tag{21}$$

where the operator $\hat{E}\left(\cdot\right)$ is defined by (17). From the model (19), and since

$$\dot{u}_1 = \frac{1}{p_1}\ddot{y}_1 - \frac{p_1}{p_2}\dot{y}_2 \sin y_2 - \dot{y}_1 \tag{22}$$

we have

$$\begin{array}{rcl} \hat{E}\left(u_1\right) & = & \hat{E}\left(\frac{1}{p_1}\dot{y}_1 + \frac{p_1}{p_2}c_2 - y_1\right) = \frac{1}{p_1}\hat{E}\left(\dot{y}_1\right) + \frac{p_1}{p_2}\hat{E}\left(c_2\right) - \hat{E}\left(y_1\right) \\ \hat{E}\left(u_2\right) & = & \hat{E}\left(\dot{y}_2 + p_1^2 p_2 s_1\right) = \hat{E}\left(\dot{y}_2\right) + p_1^2 p_2 \hat{E}\left(s_1\right) \\ \hat{E}\left(u_1 u_2\right) & = & \frac{1}{p_1}\hat{E}\left(\dot{y}_1\dot{y}_2\right) + p_1 p_2 \hat{E}\left(\dot{y}_1 s_1\right) + \frac{p_1}{p_2}\hat{E}\left(\dot{y}_2 c_2\right) \\ & & + p_1^3 \hat{E}\left(c_2 s_1\right) - \hat{E}\left(y_1 \dot{y}_2\right) - p_1^2 p_2 \hat{E}\left(y_1 s_1\right) \\ \hat{E}\left(\dot{u}_1 u_2\right) & = & \frac{1}{p_1}\hat{E}\left(\ddot{y}_1\dot{y}_2\right) + p_1 p_2 \hat{E}\left(\ddot{y}_1 s_1\right) - \frac{p_1}{p_2}\hat{E}\left(\dot{y}_2^2 s_2\right) \\ & & - p_1^3 \hat{E}\left(\dot{y}_2 s_2 s_1\right) - \hat{E}\left(\dot{y}_1\dot{y}_2\right) - p_1^2 p_2 \hat{E}\left(\dot{y}_1 s_1\right). \end{array} \tag{23}$$

where $s_i = \sin y_i$ and $c_i = \cos y_i$. The system (21) becomes:

$$\Sigma_I(N) : \left\{ \begin{array}{l} \frac{1}{p_1}\hat{E}\left(\dot{y}_1\dot{y}_2\right) + p_1 p_2 \hat{E}\left(\dot{y}_1 s_1\right) + \frac{p_1}{p_2}\hat{E}\left(c_2\dot{y}_2\right) + p_1^3 \hat{E}\left(c_2 s_1\right) - \hat{E}\left(y_1\dot{y}_2\right) \\ -p_1^2 p_2 \hat{E}\left(y_1 s_1\right) - \left[\frac{1}{p_1}\hat{E}\left(\dot{y}_1\right) + \frac{p_1}{p_2}\hat{E}\left(c_2\right) - \hat{E}\left(y_1\right)\right] \cdot \left[\hat{E}\left(\dot{y}_2\right) + p_1^2 p_2 \hat{E}\left(s_1\right)\right] = 0 \\ \frac{1}{p_1}\hat{E}\left(\ddot{y}_1\dot{y}_2\right) + p_1 p_2 \hat{E}\left(\ddot{y}_1 s_1\right) - \frac{p_1}{p_2}\hat{E}\left(\dot{y}_2^2 s_2\right) \\ -p_1^3 \hat{E}\left(\dot{y}_2 s_2 s_1\right) - \hat{E}\left(\dot{y}_1\dot{y}_2\right) - p_1^2 p_2 \hat{E}\left(\dot{y}_1 s_1\right) = 0 \end{array} \right. \tag{24}$$

It consists of two nonlinear equations with two unknowns $p_1$ and $p_2$. The coefficients of this systems $(\overline{\dot{y}_1\dot{y}_2}, \overline{\dot{y}_1 s_1}, \overline{c_2\dot{y}_2}, \ldots)$ depend on $N$. We solved this system for different values of $N$ using an interval method (see [13], [18]). For each $N$, we have found only one solution vector $\hat{\mathbf{p}}(N)$. The table of Figure (4) represents the estimation square error estimation for different values of $N$. One can check $\hat{\mathbf{p}}(N)$ tends to $\mathbf{p}^*$, the true parameter vector, as $N$ tends to infinity.

| $N$ | 100 | 600 | 1000 | 6000 | 10000 | 16000 |
|---|---|---|---|---|---|---|
| $\hat{E}\left([\hat{p}_1(N) - p_1^*]^2\right)$ | $4,8.10^{-2}$ | $1,1.10^{-2}$ | $1,1.10^{-2}$ | $10^{-4}$ | $2,5.10^{-5}$ | $10^{-5}$ |
| $\hat{E}\left([\hat{p}_2(N) - p_2^*]^2\right)$ | $7,8.10^{-2}$ | $3,5.10^{-2}$ | $1,4.10^{-2}$ | $10^{-3}$ | $9.10^{-5}$ | $3.10^{-6}$ |

Fig. 4. Parametric square error for blind estimation based on independence prior.

Note that, in the presence of noise, the proposed method is ill-suited since differentiation leads to noise amplification. In this context, our approach fails to give any reliable results without a serious adaptation. However there exists practical situations where the number of outputs is larger than that of outputs and/or where more statistical properties of the inputs signal are available. For such cases, where noise is involved, our approach can also be used as an effective method for estimating the parameters of a system [16][14].

*Remark 8:* If the inputs are not assumed to be independent anymore and if the only first input $u_1$ is assumed to be Gaussian, then experiments have shown that our approach is able to identify the parameter vector in a unique way, using estimation functions implementing Gaussianity of $u_1$, $\dot{u}_1$ etc.

## V. Conclusion

In this paper, a new blind parameter estimation approach based on statistical assumptions on the unknown input signals has been proposed. We have shown that for a large class of parametric models, the knowledge of a few statistical cross-moments of outputs can be used for estimating the unknown parameters. Although no proof of parameter uniqueness is given, the proposed method returns the set of compatible parameters (with the estimating function).

As we explained, our goal is not to provide a robust estimation method, but simply to prove that borderline problems, basically considered as unsolvable, can be solved using weak statistical priors on the input signals. Moreover, this paper points out the interest of using statistics of the derivatives of signals.

This approach seems related to second-order blind source separation (BSS) method [3] based on delayed variance-covariance matrices, in which delays instead of derivatives are used. Of course, when observations are noisy, when the number of samples is not large enough or when the model is not perfectly known (*e.g.* the order $r$ is unknown), the method proposed in this paper does not lead to reliable estimations. Further investigations include the existence condition and the uniqueness issues, the relationships with BSS method using delayed variance-covariance matrices, and robust algorithms especially for noisy observations or small size samples.

## References

[1] A. Akhenak, M. Cahldi, D. Maquin, and J. Ragot. Sliding mode multiple observer for fault detection and isolation. In *42nd IEEE Conf. on Decision and Control*, volume 1, pages 953–958, Hawaii, USA, december 2003.

[2] T. Bell and T. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. In *Neural Computation*, volume 7, pages 1004–1034, 1995.

[3] A. Belouchrani, K. Abed Meraim, J.F. Cardoso, and E. Moulines. A blind source separation technique using second order statistics. *IEEE Transactions on Signal Processing*, 45:210–220, 1997.

[4] J. F. Cardoso. Infomax and maximum likelihood for blind source separation. *IEEE Signal Processing Letters*, 4(4), 1997.

[5] J. F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, 44(12):3017–3030, December 1996.

[6] A. Cichocki and S. Amari. *Adaptative Blind Signal and Image Processing : Learning algoritms and applications*. John Wiley, 2002.

[7] P. Comon. Independent component analysis. In *International Signal Proceedings Workshop on High-Order Statistics*, pages 29–38, 1991.

[8] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, April 1994.

[9] S. Diop, J. W. Grizzle, P.E. Moraal, and A. Stephanopoulou. Interpolation and numerical differentiation for observer design. In *Proceedings of the American Control Conference*, pages 1329–1333, Baltimore, June 1994.

[10] S. Diop, J.W. Grizzle, and F. Chaplais. On numerical differentiation. algorithms for nonlinear estimation. In *Proceedings of the 2000 IEEE Conference on Decision and Control*, volume 2, pages 1133–1138, Sydney Australia, 2000.

[11] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon. Flatness and defect of non linear systems : Introductory theory and examples. In *Int J. Control*, number 61, pages 1327–1361, 1995.

[12] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley, 2001.

[13] J. Jaulin, M. Kieffer, D. Didrit, and E. Walter. *Applied Interval Analysis*. Springer, 2001.

[14] M. Joho, H. Mathis, and R. H. Lambert. Overdetermined blind source separation: Using more sensors than source signals in a noisy mixture. In *Independent Component Analysis and Blind Signal Separation ICA 2000*, pages 81–86, Helsinki, Finland, June 2000.

[15] C. Jutten and J. Hérault. Blind separation of sources, part i : an adaptative algorithm based on neuromimetic architecture. *IEEE Transactions on Signal Processing*, 24:1–10, 1991.

[16] A. Koutras, E. Dermatas, and G. Kokkinakis. Improving simultaneous speech recognition in real room environments using overdetermined blind source separation. In *Eurospeech 2001*, Scandinavia, 2001.

[17] Ph. Martin, R.M. Murray, and P. Rouchon. Flat system, equivalence and trajectory generation. Technical report, ensmp, April 2003. *http:/www.cds.caltech.edu/∼murray/papers/2003d_mmr03-cds.html*.

[18] R.E. Moore. *Methods and Applications of Interval Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 1979.

[19] F. Palestan and J.W. Grizzle. Synthesis of nonlinear observers via structural analysis and numerical differentiation. In *In Proceedings of the 1999 European Control Conference*, Karlsruhe, Germany, September 1999.

[20] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, third edition edition.

[21] E. Parzen. *Stochastic Processes*. Holden-Day: San Francisco, 1962.

[22] D. T. Pham and J. F. Cardoso. Blind separation of instantaneous mixtures of non stationary sources. *IEEE Transaction on Signal Processing*, 9(49):1837–1848, September 2001.

[23] H. Sira-Ramirez and M. Fliess. On the output feedback control of a synchronous generator. In *Proceedings of the 43 rd IEEE Conf. Decision Control*, Bahamas, 2004.

[24] E. Walter and L. Pronzato. *Identification of parametric models from experimental data*. Springer, 1997.