Identification of Time-Non Local Models under Diffusive Representation

Céline Casenave

Identification of Time-non local Models under Diffusive Representation

C. Casenave *

* *CNRS; LAAS; 7 avenue du colonel Roche, F-31077 Toulouse, France, Université de Toulouse; UPS, INSA, INP, ISAE; LAAS; F-31077 Toulouse, France (e-mail: casenave@laas.fr).*

Abstract: In this paper, we show how the so-called diffusive representation can be used in order to identify nonlinear Volterra models of the form $H(\partial_t)X = f(u, X) + v$. Several methods are described, all being based on a suitable parameterization of $H(\partial_t)$ by means of its γ-symbol. Following this idea, the complex dynamic nature of $H(\partial_t)$ can be summarized by a few parameters on which the identification of the dynamic part of the model will focus. For illustration, we implement the methods on a concrete numerical example.

Keywords: Least squares identification, Sampled data, Operators, Convolution integral, Continuous time systems, Dynamic systems, Distributed-parameter systems

1. INTRODUCTION

Numerous models of physics involve dynamic convolution operators of the form:

$$u \mapsto H(\partial_t)u,\ [H(\partial_t)u](t) = \int_0^t h(t-s)\,u(s)\,ds, \qquad (1)$$

where h is the impulse response of $H(\partial_t)$. Identifying such an operator means identifying one of its characteristic quantities. As the operator to be identified is linear, a convenient and rather general approach consists in working in the frequency domain where any causal operator can be well-defined by its symbol $H(\mathrm{i}\omega)$, that is the Fourier transform of its impulse response. Then, the problem of identifying $H(\mathrm{i}\omega)$ can be classically solved from physical measurements by means of Fourier techniques. Note however that purely frequency identification presents some well-known shortcomings. In particular, the so-identified symbol $H(\mathrm{i}\omega)$ is in general ill adapted to the construction of efficient time-realizations. This is partly due to excessive numerical cost of quadrature approximations resulting from the intrinsic convolution nature of the associated operator, sometimes with long memory (Rumeau et al. [2006]) or even delay-like behaviors (Montseny [2007]). Another shortcoming is that frequency methods are incompatible with real-time identification (and so with pursuit when the symbol has the ability to evolve slowly). But above all, the number of unknown parameters is excessive, what makes the problem excessively sensitive to measurement noises.

In opposite, time domain identification techniques do not present such drawbacks. However, their scope is in general not so wide. See for example (Monin and Salut [1996]) for an interesting optimal method based on ARMA lattices.

By means of *diffusive representation*, it is possible to use the notion of symbol in the time-domain to identify a small number of unknown parameters, sufficient to get a good accuracy. Indeed, the diffusive representation theory is devoted to state realizations of causal integral operators which can be characterized by a suitable distribution called the γ-symbol. Instead of identifying the symbol $H(\mathrm{i}\omega)$ of the operator, we then identify its γ-symbol μ, which is linearly involved.

Identifying the γ-symbol of a convolution operator presents numerous advantages:

- an input-output stable differential formulation;
- recursive identification algorithms, compatible with real-time identification or even pursuit, can be build;
- as for frequency methods, the rational and non rational operators are considered without distinction and can be identified by the same identification process.

The paper is organized as follows. In section 2, we present a simplified version of the diffusive representation. In section 3, we describe the method of identification of a dynamic operator by means of its γ-symbol. Then in section 4, we present two methods of identification of Volterra models. Finally, in section 5, we implement the methods described on a concrete example.

2. DIFFUSIVE FORMULATION OF CAUSAL CONVOLUTION OPERATORS

A complete statement of diffusive representation will be found in (Montseny [2005]). Various applications and questions relating to this approach will be found for example in (Audounet et al. [1998], Carmona and Coutin [1998], Degerli et al. [Ap. 1999], Garcia and Bernussou [1998], Lenczner and Montseny [2005], Levadoux and Montseny [2003], Montseny [2007], Mouyon and Imbert [2002], Rumeau et al. [2006]).

2.1 Mathematical framework

We consider a causal convolution operator defined, on any continuous function $u : \mathbb{R}^+ \to \mathbb{R}$, by:

$$u \mapsto \left(t \mapsto \int_0^t k(t-s)\, u(s)\, ds \right). \tag{2}$$

We denote K the Laplace transform of k and $K(\partial_t)$ the convolution operator defined by (2).

Let $u^t(s) = \mathbf{1}_{]-\infty,t]}(s)\, u(s)$ be the restriction of u to its past and $u_t(s) = u^t(t-s)$ the so-called "history" of u. From causality of $K(\partial_t)$, we deduce:

$$[K(\partial_t)(u-u^t)](t) = 0 \text{ for all } t; \tag{3}$$

then, we have for any continuous function u:

$$[K(\partial_t)\, u](t) = \left[\mathcal{L}^{-1}\left(K\, \mathcal{L}u\right)\right](t) = \left[\mathcal{L}^{-1}\left(K\, \mathcal{L}u^t\right)\right](t). \tag{4}$$

We define $\Psi_u(t,p) := \mathrm{e}^{p\,t}\left(\mathcal{L}u^t\right)(p) = \left(\mathcal{L}u_t\right)(-p)$; by computing $\partial_t \mathcal{L}u_t$, Laplace inversion and use of (4):

Lemma 1. Ψ_u is solution of the differential equation:

$$\partial_t \Psi(t,p) = p\, \Psi(t,p) + u, \;\; t>0, \;\; \Psi(0,p) = 0, \tag{5}$$

and for any $b \geqslant 0$:

$$\left[K(\partial_t)\, u\right](t) = \frac{1}{2i\pi} \int_{b-i\infty}^{b+i\infty} K(p)\, \Psi_u(t,p)\, dp. \tag{6}$$

Let γ be a closed [1] simple arc in $\mathbb{C}^-$; we denote Ω_γ^+ the exterior domain defined by γ, and Ω_γ^- the complementary of $\overline{\Omega_\gamma^+}$. Assume there exists $\alpha_\gamma \in]\frac{\pi}{2}, \pi[$ and $a \in \mathbb{R}$ such that:

$$e^{\mathrm{i}[-\alpha_\gamma,\, \alpha_\gamma]}\mathbb{R}_+ + a \;\subset\; \Omega_\gamma^+. \tag{7}$$

By use of standard techniques (Cauchy theorem, Jordan lemma), it can be shown:

Lemma 2. For γ such that K is holomorphic in Ω_γ^+, if $K(p) \to 0$ when $p \to \infty$ in Ω_γ^+, then:

$$\left[K(\partial_t)\, u\right](t) = \frac{1}{2i\pi} \int_{\tilde{\gamma}} K(p)\, \Psi_u(t,p)\, dp, \tag{8}$$

where $\tilde{\gamma}$ is any closed simple arc in Ω_γ^+ such that $\gamma \subset \Omega_{\tilde{\gamma}}^-$.

We now suppose that γ, $\tilde{\gamma}$ are defined by functions of the Sobolev space [2] $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R};\mathbb{C})$, also denoted γ, $\tilde{\gamma}$ and such that $\gamma(0) = 0$. We use the convenient notation $\langle \mu, \psi \rangle = \int \mu\, \psi\, d\xi$; in particular, when μ is atomic that is $\mu = \sum_k a_k\, \delta_{\xi_k}$, we have: $\langle \mu, \psi \rangle = \sum_k a_k\, \psi(\xi_k)$.

Under hypothesis of lemma 2, we have (Montseny [2005]):

Theorem 3. If the possible singularities of K on γ are simple poles or branching points such that $|K \circ \gamma|$ is locally integrable in their neighborhood, then:

1. with $\tilde{\mu} = \frac{\tilde{\gamma}'}{2i\pi}\, K \circ \tilde{\gamma}$ and $\tilde{\psi}(t,.) = \Psi_u(t,.) \circ \tilde{\gamma}$:

$$\left[K(\partial_t)\, u\right](t) = \left\langle \tilde{\mu}, \tilde{\psi}(t,.) \right\rangle; \tag{9}$$

2. with [3] $\tilde{\gamma}_n \to \gamma$ in $W^{1,\infty}_{\mathrm{loc}}$ and $\mu = \frac{\tilde{\gamma}'}{2i\pi} \lim K \circ \tilde{\gamma}_n$ in the sense of measures:

$$\left[K(\partial_t)\, u\right](t) = \left\langle \mu, \psi(t,.) \right\rangle, \tag{10}$$

where $\psi(t,\xi)$ is solution of the following evolution problem on $(t,\xi) \in \mathbb{R}^{*+} \times \mathbb{R}$ (of diffusive type):

$$\partial_t \psi(t,\xi) = \gamma(\xi)\, \psi(t,\xi) + u(t),\; \psi(0,\xi) = 0. \tag{11}$$

[1] Possibly at infinity

[2] $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R};\mathbb{C})$ is the topological space of measurable functions $f : \mathbb{R} \to \mathbb{C}$ such that $f, f' \in L^\infty_{\mathrm{loc}}$ (that is f and f' are locally bounded in the almost everywhere sense).

[3] This convergence mode means that on any bounded set P, $\tilde{\gamma}_{n|P} - \gamma_{|P} \to 0$ and $\tilde{\gamma}'_{n|P} - \gamma'_{|P} \to 0$ uniformly.

Definition 4. The measure μ defined in theorem 3 is called γ-symbol of operator $K(\partial_t)$. The function ψ solution of (11) is called the γ-representation of u.

Remark 1. Note in particular that thanks to (10), the Dirac measure δ is clearly a γ-symbol of the operator $u \mapsto \int_0^t u(s)\, ds$, denoted ∂_t^{-1}. We indeed have $(\partial_t^{-1} u)(t) = \langle \delta, \psi(t,.) \rangle = \psi(t,0)$, with $\partial_t \psi(t,0) = u$, $\psi(0,0) = 0$.

Beyond the measure framework, the general space of γ-symbols is a quotient space of distributions, denoted Δ'_γ; it is the topological dual of the space $\Delta_\gamma \ni \psi(t,.)$ (Montseny [2005]). The composition product of operators has an equivalent in Δ'_γ, denoted $\sharp_\gamma$ or simply $\sharp$: if μ and ν are respective γ-symbols of $H(\partial_t)$ and $K(\partial_t)$, then $\mu \sharp_\gamma \nu$ is a γ-symbol of $H(\partial_t) \circ K(\partial_t)$. Note that the product $\sharp_\gamma$ is inner, commutative and continuous [4] in Δ'_γ and so $(\Delta'_\gamma, \sharp_\gamma)$ is an algebra (of γ-symbols) isomorphic to a commutative algebra of causal convolution operators.

Formulation (11,10) can be extended to operators $H(\partial_t)$ of the form $H(\partial_t) = K(\partial_t) \circ \partial_t^n$ where $K(\partial_t)$ admits a γ-symbol in Δ'_γ; such operators are said to be γ-diffusive of degree n. Formally we have:

$$[K(\partial_t) \circ \partial_t^n u](t) = \left\langle \mu, \partial_t^n \psi(t,.) \right\rangle, \tag{12}$$

with $\psi(t,\xi)$ solution of (11) and μ the γ-symbol of $K(\partial_t)$. In the same way, Δ'_γ can be extended to an algebra, denoted Σ_γ, composed of the γ-symbols of operators γ-diffusive of degree $n \in \mathbb{N}$, the γ-symbols ν of $K(\partial_t) \circ \partial_t^n$ being characterized by the relation:

$$\mu = \nu \sharp \delta^n \tag{13}$$

where $\delta^n = \underbrace{\delta \sharp \delta \sharp ... \sharp \delta}_{n \text{ times}} \in \Delta'_\gamma$ is a γ-symbol of ∂_t^{-n}.

The inversion of γ-symbols cannot be defined in Δ'_γ because this algebra is not unitary; this operation is nevertheless well-defined in Σ_γ. If $\mu \in \Sigma_\gamma$ is a γ-symbol of $K(\partial_t)$ such that $K(\partial_t)^{-1} \circ \partial_t^{-n}$ has a γ-symbol $\nu \in \Delta'_\gamma$, then $\nu = \mu^{-1} \sharp \delta^n$ and we have:

$$[K(\partial_t)^{-1} u](t) = \left\langle \mu^{-1} \sharp \delta^n, \partial_t^n \psi(t,.) \right\rangle, \tag{14}$$

with ψ solution of (11). Note in particular that operator ∂_t^{-1} defined above is the unique inverse of the derivative operator ∂_t. See (Casenave [2009]) for details about the inversion of γ-symbols.

2.2 Summary

Let γ be a closed simple arc in $\mathbb{C}^-$ verifying the property (7). An operator $H(\partial_t)$ γ-diffusive of degree n admits the following state-realization:

$$\partial_t \psi(t,\xi) = \gamma(\xi) \psi(t,\xi) + u(t),\; \psi(0,.) = 0, \tag{15}$$

$$(H(\partial_t) u)(t) = < \mu, \partial_t^n \psi(t,.) >_{\Delta'_\gamma, \Delta_\gamma}, \tag{16}$$

where $\mu \in \Delta'_\gamma$ is the γ-symbol of $K(\partial_t) := H(\partial_t) \circ \partial_t^{-n}$; the essential conditions the operator has to satisfy to admit such a representation are:

- K holomorphic in Ω_γ^+, (17)
- $K(p) \to 0$ when $|p| \to +\infty$ in Ω_γ^+, (18)

[4] In the sense of the strong topology of Δ'_γ.

K being the Laplace-symbol of $K(\partial_t)$ given by:

$$K(p) = \frac{H(p)}{p^n}. \tag{19}$$

Thanks to the sector condition (7) verified by γ, the state realization is of diffusive type: so, cheap and precise numerical approximations of (15,16) can be easily built.

2.3 About numerical approximations

The state equation (11) is infinite-dimensional. To get numerical approximations, we consider $\mathcal{M}_L$ a sequence of L-dimensional spaces of atomic measures on suitable meshes $\{\xi_l^L\}_{l=1:L}$ on the variable ξ; L-dimensional approximations μ^L of the γ-symbol $\mu \in \Delta'_\gamma$ are then defined in the sense of atomic measures, that is:

$$\mu^L = \sum\nolimits_{l=1}^{L} \mu_l^L \, \delta_{\xi_l^L}, \quad \mu_l^L \in \mathbb{C}. \tag{20}$$

If $\cup_L \mathcal{M}_L$ is dense in the topological space Δ'_γ (that is, concretely, if $\cup_L \{\xi_l^L\}$ is dense in $\mathbb{R}$), then we can have (Montseny [2005]):

$$\left\langle \mu^L, \psi \right\rangle \underset{L\to+\infty}{\longrightarrow} \langle \mu, \psi \rangle \;\; \forall \psi \in \Delta_\gamma; \tag{21}$$

so, we have the following L-dimensional approximate state formulation of $K(\partial_t)$ (with γ-symbol μ):

$$\begin{cases} \partial_t \psi(t,\xi_l^L) = \gamma(\xi_l^L)\,\psi(t,\xi_l^L) + u(t),\; l = 1:L,\; \psi(0,\xi_l^L) = 0 \\ [K(\partial_t)\,u](t) \simeq \sum\nolimits_{l=1}^{L} \mu_l^L \, \psi(t,\xi_l^L). \end{cases} \tag{22}$$

One of the properties of the approach presented above is that most of non rational operators encountered in practice can be closely approximate with small L (see for example (Montseny [Nov. 2004])). In the context of identification of dynamic operators, this will be a great advantage because only a few numerical parameters μ_l^L will have to be identified from experimental data, while the property (21) will ensure the well-posedness and the robustness of the problem as soon as the operator to be identified admits a γ-symbol in Σ_γ.

3. IDENTIFICATION OF AN OPERATOR BY MEANS OF ITS γ-SYMBOL

We consider here the problem of identification of an integral operator $H(\partial_t)$:

$$X \longmapsto v = H(\partial_t)X. \tag{23}$$

Let v_m and X_m be some (direct or indirect and possibly noisy) measurements of the output v and the associated input X; the problem consists then in identifying the γ-symbol of $H(\partial_t)$ from these data.

3.1 Principle

Assume that $H(\partial_t)$ is γ-diffusive of degree n and let μ be the γ-symbol of $K(\partial_t) = H(\partial_t) \circ \partial_t^{-n}$. We then have (see section 2):

$$v = H(\partial_t)X = \langle \mu, \partial_t^n \psi_X \rangle = \left\langle \mu, \psi_{\partial_t^n X} \right\rangle. \tag{24}$$

By denoting $\mathcal{A}_{\partial_t^n X}$ the operator defined by:

$$\mathcal{A}_{\partial_t^n X} : \mu \longmapsto \left\langle \mu, \psi_{\partial_t^n X} \right\rangle, \tag{25}$$

we get a new formulation, linear with respect to μ:

$$v = \mathcal{A}_{\partial_t^n X}\, \mu, \tag{26}$$

and which can also be written:

$$v_m = \mathcal{A}_{\partial_t^n X_m}\, \mu + \epsilon(\mu), \tag{27}$$

where $\epsilon(\mu)$ is the error equation. The least squares estimator $\hat{\mu}$ of μ is then given by:

$$\hat{\mu} = \arg\min_{\mu\in\mathcal{E}} \|\epsilon(\mu)\|_{\mathcal{F}}^2, \tag{28}$$

where $\mathcal{E}$ is an Hilbert subspace of Δ'_γ and $\mathcal{F}$ another Hilbert space chosen *a priori*. The solution of (28) is formally obtained by:

$$\hat{\mu} = \mathcal{A}^\dagger_{\partial_t^n X_\mathrm{m}}\, v_\mathrm{m}, \tag{29}$$

where $\mathcal{A}^\dagger_{\partial_t^n X_\mathrm{m}}$ designates the pseudo-inverse of $\mathcal{A}_{\partial_t^n X_\mathrm{m}}$ (Ben-Israel and Greville [2003]).
In the sense of the hilbertian norm of $\mathcal{F}$, the estimator $\hat{\mu}$ is optimal. When the measurement X_m is noisy, the estimator $\hat{\mu}$ is also *biased* because $\mathcal{A}^\dagger_{\partial_t^n X_\mathrm{m}}$ depends on the measurement noise. To mitigate this problem, it can be interesting to consider some classical bias reduction methods as the ones used in time-continuous system identification.

By denoting $\hat{K}(\partial_t)$ the operator of γ-symbol $\hat{\mu}$ and $\hat{H}(\partial_t) = \hat{K}(\partial_t) \circ \partial_t^n$, the identified model is then written $v = \hat{H}(\partial_t)X$, or, under diffusive formulation:

$$\begin{cases} \partial_t \psi = \gamma\,\psi + \partial_t^n X,\; \psi(0,.) = 0 \\ v = \left\langle \hat{\mu}, \psi_{\partial_t^n X} \right\rangle. \end{cases} \tag{30}$$

Remark 2. Note that in the case where $v_m = v$ (no output measurement noise), we can get an unbiased estimation by identifying the operator $H(\partial_t)^{-1}$ of input v and output X instead of $H(\partial_t)$. We suppose $H(\partial_t)^{-1}$ is γ-diffusive of degree n' and we denote $\nu \in \Delta'_\gamma$ the γ-symbol of $K(\partial_t) = H(\partial_t)^{-1} \circ \partial_t^{-n'}$. We then have:

$$X = \mathcal{H}(\partial_t)^{-1} v = \mathcal{A}_{\partial_t^{n'} v}\, \nu, \tag{31}$$

and so get, after identification:

$$\hat{\nu} = \mathcal{A}^\dagger_{\partial_t^{n'} v}\, X_\mathrm{m}. \tag{32}$$

The identified model then writes $v = \hat{K}(\partial_t)^{-1} \circ \partial_t^{-n'} X$, that is, under diffusive formulation:

$$\begin{cases} \partial_t \psi = \gamma\,\psi + \partial_t^n X,\; \psi(0,.) = 0 \\ v = \left\langle (\hat{\nu})^{-1} \sharp \delta^{n'+n}, \psi_{\partial_t^n X} \right\rangle. \end{cases} \tag{33}$$

The computation of $(\hat{\nu})^{-1} \sharp \delta^{n'+n}$ can be numerically performed as shown in (Casenave [2009]).

More information about this identification method will be found in (Casenave and Montseny [2009b, 2010]).

3.2 Prefiltering with an invertible convolution operator

The identification model (26) can be equivalently transformed by composition with any invertible causal convolution operator $Q(\partial_t)$:

$$Q(\partial_t)v = Q(\partial_t) \circ H(\partial_t)X = H(\partial_t) \circ Q(\partial_t)X. \tag{34}$$

By denoting $\tilde{v} = Q(\partial_t)v$ and $\tilde{X} = Q(\partial_t)X$, the model is then written:

$$\tilde{v} = H(\partial_t)\tilde{X}. \tag{35}$$

When applying the identification method to model (35), the estimator of μ is written:

$$\hat{\mu} = \mathcal{A}^{\dagger}_{\partial_t^n \tilde{X}_\mathrm{m}} \, \tilde{v}_\mathrm{m}, \tag{36}$$

with $\tilde{X}_\mathrm{m} = Q(\partial_t) X_\mathrm{m}$ and $\tilde{v}_\mathrm{m} = Q(\partial_t) v_\mathrm{m}$.

When $n \neq 0$, this transformation is necessary to make the identification process possible, otherwise, the amplification of the noise measurement of the term $\partial_t^n X_\mathrm{m}$ in (28) and (29) would make the identification impossible. So, the operator $Q(\partial_t)$ has to be chosen so that the measurement noise is not amplified by the term $\partial_t^n \tilde{X}_m$. It has to sufficiently attenuate high frequencies (i.e. $|Q(\mathrm{i}\omega)| \underset{H.F}{\sim} \frac{1}{\omega^n}$), without amplifying low and middle ones (i.e. $|Q(\mathrm{i}\omega)| \underset{L.F}{\sim} 1$). Basically, it behaves like a nth order low-pass filter; here we simply consider the filter $Q(\partial_t)$ with transfer function:

$$Q(p) = \frac{\sigma^n}{(p+\sigma)^n}, \tag{37}$$

where $\sigma > 0$ (the cutoff frequency) will be chosen in such a way that $\|\epsilon(\mu)\|^2_{\mathcal{F}}$ is as "small" as possible.

Remark 3. The transfer function $Q(p)$ could also be optimized in order to minimize the estimation error.

3.3 Case of multiple trajectories

Consider a set of input trajectories X^j, $j = 1 : J$ and the associated output trajectories $v^j = H(\partial_t) X^j$. Let v^j_m and X^j_m be some measurements of v^j and X^j. Then, without any change of notations, model (26) can be extended to the general case of multiple trajectories simply by defining:

$$\begin{aligned} v &= (v^1, ..., v^J)^T, \; X = (X^1, ..., X^J)^T, \\ v_\mathrm{m} &= (v^1_\mathrm{m}, ..., v^J_\mathrm{m})^T, \; X_\mathrm{m} = (X^1_\mathrm{m}, ..., X^J_\mathrm{m})^T, \\ \text{and } \mathcal{A}_{\partial_t^n X} &: \mu \longmapsto \begin{bmatrix} < \mu, \psi_{\partial_t^n X^1} > \\ \vdots \\ < \mu, \psi_{\partial_t^n X^J} > \end{bmatrix}. \end{aligned} \tag{38}$$

3.4 Recursive formulations

Recursive formulations of (29) can be established under the form (Garcia and Bernussou [1998]):

$$\hat{\mu}_t = \hat{\mu}_{t-\Delta t} + K_{t-\Delta t}(v_\mathrm{m} - \mathcal{A}_{\partial_t^n X_\mathrm{m}} \hat{\mu}_{t-\Delta t})_{|[0,t]}; \tag{39}$$

such formulations allow real-time identification (or even the pursuit of μ in case of slowly varying operators $H(t, \partial_t)$).

3.5 In the numerical point of view

Let (v, X) be such that $v = H(\partial_t) X$. We consider a discrete data set $\{v^k_\mathrm{m}, X^k_\mathrm{m}\}_{k=0:K}$ where v^k_m and X^k_m are the respective measurements of v and X at time instant t_k. We also denote v_m and X_m the continuous measured trajectories such that $v_\mathrm{m}(t_k) = v^k_\mathrm{m}$ and $X_\mathrm{m}(t_k) = X^k_\mathrm{m}$.
The time instants t_k are given by:

$$t_0 = 0, \; t_k = t_{k-1} + \Delta t_k, \; k = 1 : K. \tag{40}$$

We consider a discretization $\{\xi^L_l\}_{l=1:L}$ of the variable ξ and the associated L-dimensional approximation μ^L of μ given by (20). At time instant t_k, we consider the following identification model of finite dimension:

$$v^k_m = \varphi^T_m(t_k) \, \mu + \epsilon(t_k, \mu), \tag{41}$$

where :

- $\mu = (\mu^L_1, ..., \mu^L_L)^T$,
- $\varphi^T_m(t_k) = [\Psi^k_1, ..., \Psi^k_L]$ with $\Psi^k_l = \psi_{\partial_t^n X_m}(t_k, \xi^L_l)$,
- and $\epsilon(t_k, \mu)$ is the error equation at time t_k.

The least squares estimator $\hat{\mu}$ is then given by:

$$\hat{\mu} = \arg \min_{\mu \in \mathbb{C}^L} \sum_{k=1}^{K} \epsilon^2(t_k, \mu) \Delta t_k, \tag{42}$$

that is:

$$\hat{\mu} = \left[\sum_{k=1}^{K} \varphi_m(t_k) \, \varphi^T_m(t_k) \Delta t_k \right]^{-1} \sum_{k=1}^{K} \varphi_m(t_k) \, v^k_m \Delta t_k. \tag{43}$$

3.6 On the choice of γ

To choose the contour γ on which depends the problem (28), recall first that γ has to verify the sector condition (7), and then that the transfer function K of $K(\partial_t) = H(\partial_t) \circ \partial_t^{-n}$ has to be analytic in Ω^+_γ (see(17)). So all the singularities of K have to be inside the domain $\overline{\Omega^-_\gamma}$ delimited by γ. However, as the operator $H(\partial_t)$ is unknown, so is the position of the singularities of K. As a consequence the contour γ will be chosen in such a way that the domain $\overline{\Omega^-_\gamma}$ is sufficiently big to contain all the singularities of K. In practice, we often take a contour of sector type:

$$\gamma(\xi) = |\xi| \, e^{\mathrm{i} \, \mathrm{sign}(\xi)(\frac{\pi}{2} + \alpha)}, \tag{44}$$

with $\alpha \in]0, \frac{\pi}{2}]$ a sufficiently "small" angle.
Note that, however, the more γ is close to the axis $\mathrm{i}\mathbb{R}$ (stability limit axis), the finer (and so the more expensive numerically) the discretization in ξ has to be in order to get a good approximation of the model. So in practice, we first identify the system with a small value of angle α and a great value L. Then, if the identification results are good with these values, we can iterate the process with a greater angle α ($\alpha \leqslant \frac{\pi}{2}$) and a smaller L.

Note that in practice, the available information about the operator $H(\partial_t)$ can help us in the choice of γ.

Some indications about the choice of the discretization points ξ^L_l can be found in Montseny [2005]. Note however that, from the discrete data $\{v^k_\mathrm{m}, X^k_\mathrm{m}\}$, we can identify the impulse response $H(\mathrm{i}\omega)$ only in the frequency band $[\frac{2\pi}{t_K}; \frac{2\pi}{2\Delta t}]$, where $\Delta t = \max(\Delta t_k)$. Consequently, the band $[\xi^L_1; \xi^L_L]$ covered by the discretization in ξ will be chosen in such a way that [5] $[\frac{2\pi}{t_K}; \frac{2\pi}{2\Delta t}] \subset \big|\gamma([\xi^L_1; \xi^L_L])\big|$.

4. IDENTIFICATION OF VOLTERRA MODELS

Now we consider a nonlinear Volterra model of the form:

$$H(\partial_t) X = f(u, X) + v, \tag{45}$$

where $H(\partial_t)$ is a causal convolution operator and f is a nonlinear function. The problem under consideration

[5] The quantity $|\gamma(\xi)|$ is comparable to a frequency. In the case of a contour of sector type we have $|\gamma(\xi)| = |\xi|$.

consists in identifying both operator $H(\partial_t)$ and function f from some measurements $u_{\rm m}$, $v_{\rm m}$ and $X_{\rm m}$ of u, v and X. We will describe two identification methods: with the first one, $H(\partial_t)$ and f are identified simultaneously, whereas with the second one, $H(\partial_t)$ is first identified alone after cancellation of the non linear term f in the model. Only the theoretical principles of these two methods are described; the numerical implementation is similar to the one described in section 3.5.

Suppose $H(\partial_t)$ is γ-diffusive of degree n and let μ be the γ symbol of $K(\partial_t) = H(\partial_t) \circ \partial_t^{-n}$. Then, (45) can be rewritten:

$$\mathcal{A}_{\partial_t^n X}\mu = f(u, X) + v. \tag{46}$$

4.1 Simultaneous identification of $H(\partial_t)$ and f

We consider a topological basis $\{\mathbf{g}^p \otimes \mathbf{k}^q\}_{p,q=1:+\infty}$ of a tensorial product of Hilbert spaces to which f belongs. We so have:

$$f = \sum_{p,q} \mathbf{g}^p \otimes \mathbf{k}^q \, a_{pq}, \text{ with } a := (a_{pq})_{p,q} \in \ell^2. \tag{47}$$

The Volterra model (45) can then equivalently be expressed under the *linear* form:

$$\mathcal{G}_{u,X}(\mu, a) = v, \tag{48}$$

with $\mathcal{G}_{u,X}$ the operator defined by:

$$\mathcal{G}_{u,X} : (\mu, a) \mapsto \mathcal{A}_{\partial_t^n X}\,\mu - \sum_{p,q} \mathbf{g}^p(u)\, \mathbf{k}^q(X)\, a_{pq}. \tag{49}$$

The least squares estimator $(\hat{\mu}, \hat{a})$ of (μ, a) is defined by:

$$(\hat{\mu}, \hat{a}) = \arg \min_{(\mu,a)\in\mathcal{E}\times\ell^2} \|\epsilon(\mu, a)\|^2_{\mathcal{F}}, \tag{50}$$

where $\epsilon(\mu, a) = \mathcal{G}_{u_{\rm m},X_{\rm m}}(\mu, a) - v_{\rm m}$ is the error equation. Formally we so get:

$$(\hat{\mu}, \hat{a}) = \mathcal{G}^{\dagger}_{u_{\rm m},X_{\rm m}} v_{\rm m}. \tag{51}$$

When $n \neq 0$, a prefiltering with a convolution operator (see section 3.2) is necessary so that the noise is not amplified. For complementary details about this method, we will refer to (Casenave and Montseny [2008]).

4.2 Identification of $H(\partial_t)$ after cancellation of f

For simplicity, we here suppose that $f(u, X) = f(X)$. Let also assume that $n < 2$ (note that in other cases, the identification method presented here-after can not be applied due to the amplification of the measurement noise[6] by operator ∂_t^n: a prefiltering with a convolution operator cannot mitigate the problem because the method is based on the fact that f is a function, which is not the case of $Q(\partial_t) \circ f$).

Let $\varepsilon \in \mathbb{R}^+$ and $(x, y) \in \big(C^0([0, T]; \mathbf{X})\big)^n \times \big(C^0([0, T]; \mathbf{Y})\big)^n$ with $\mathbf{X}, \mathbf{Y}$ two Banach spaces; we denote:

$$\Omega_{x,\varepsilon} := \bigcup\nolimits_{i,j=1:n} \{(i, j)\} \times \Omega^{i,j}_{x,\varepsilon}, \tag{52}$$

where $\Omega^{i,j}_{x,\varepsilon} := \big\{(t, \tau) \in [0, T]^2;\ \big\|x^i(t) - x^j(\tau)\big\| \leqslant \varepsilon\big\}$ (see figure 1 for some examples).

[6] In the case where $n = 1$, we have $\psi_{\partial_t X} = \partial_t \psi_X = \gamma\psi_X + X$; so the measurement noise is not amplified by the computation of $\psi_{\partial_t X}$.

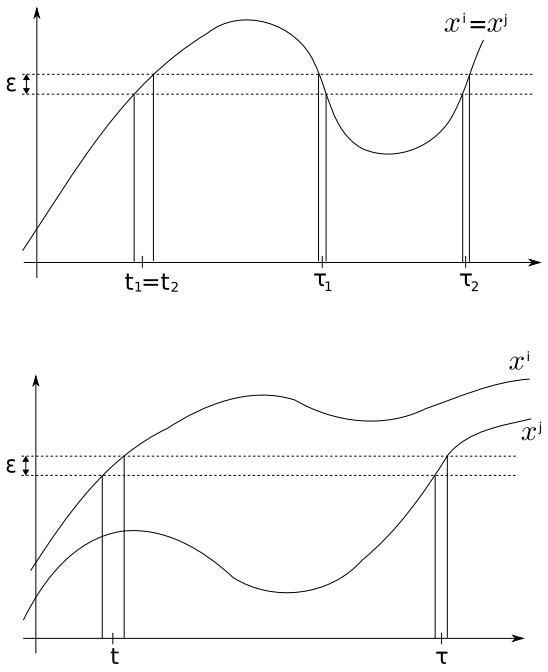


Fig. 1. Examples of (t, τ) in $\Omega^{i,i}_{x,\varepsilon}$ (top) and in $\Omega^{i,j}_{x,\varepsilon}$ (bottom)

Definition 5. The *ε-cancellation operator* $\mathfrak{D}_\varepsilon$ is defined by:

$$(x, y) \mapsto \mathfrak{D}_\varepsilon(x, y) : \begin{array}{ccc} \Omega_{x,\varepsilon} & \to & \mathbf{Y} \\ (i, j, t, \tau) & \mapsto & y^i(t) - y^j(\tau). \end{array} \tag{53}$$

Proposition 4. (1) $\mathfrak{D}_\varepsilon$ is continuous.
(2) For any continuous function $f : \mathbf{X}_0 \subset \mathbf{X} \to \mathbf{Y}$, we have the *cancellation property*:

$$\mathfrak{D}_0(x, f \circ x) = 0. \tag{54}$$

Thanks to proposition 4 and by application of operator $\mathfrak{D}_0(X, .)$ to both members of the equation, model (45) is rewritten:

$$\left\{ \begin{array}{l} \big\langle \mu, \mathfrak{D}_0\big(X, \psi_{\partial_t^n X}\big)\big\rangle = \mathfrak{D}_0\,(X, v) \\ \big\langle \mu, \psi_{\partial_t^n X}(0, .)\big\rangle - f(X(0)) = v(0). \end{array} \right. \tag{55}$$

The interest of this new formulation is that, on the one hand, up to the quantity $f_0 := f(X(0))$, the model is *independent* of f, and on the other hand, both μ and f_0 are linearly involved. Formally, these unknowns can then be determined by means of the pseudo-inverse of operator:

$$\mathcal{Y}_0 : (\mu, f_0) \mapsto \left(\begin{array}{c} \big\langle \mu, \mathfrak{D}_0(X_{\rm m}, \psi_{\partial_t^n X_{\rm m}})\big\rangle \\ \big\langle \mu, \psi_{\partial_t^n X_{\rm m}}(0, .)\big\rangle - f_0 \end{array} \right). \tag{56}$$

In general however, on the one hand the set $\Omega_{X_m,0}$ may be too poor to get the strict equivalence of (45) and (55) and, on the other hand, the available data X_m and v_m are noisy. In practice, we so consider the weakened problem:

$$\min_{(\mu,f_0)} \|\epsilon(\mu, f_0)\|^2, \tag{57}$$

where the equation error $\epsilon(\mu, f_0)$ is given by:

$$\epsilon(\mu, f_0) = \mathcal{Y}_\epsilon(\mu, f_0) - \left(\begin{array}{c} \mathfrak{D}_\varepsilon(X_{\rm m}, v_m) \\ v(0) \end{array} \right), \tag{58}$$

with:

$$\mathcal{Y}_\epsilon : (\mu, f_0) \mapsto \left(\begin{array}{c} \big\langle \mu, \mathfrak{D}_\varepsilon(X_{\rm m}, \psi_{\partial_t^n X_{\rm m}})\big\rangle \\ \big\langle \mu, \psi_{\partial_t^n X_{\rm m}}(0, .)\big\rangle - f_0 \end{array} \right). \tag{59}$$

The estimator $(\hat{\mu}, \hat{f}_0)$ of (μ, f_0) is then defined by:

$$(\hat{\mu}, \hat{f}_0) = \mathcal{Y}_\varepsilon^\dagger \begin{pmatrix} \mathfrak{D}_\varepsilon (X_{\rm m}, v_{\rm m}) \\ v_{\rm m}(0) \end{pmatrix}. \tag{60}$$

By means of standard regression methods, the function f can then be easily estimated from its "pseudo graph":

$$P_f = \bigcup_k \left\{ \left(X_{\rm m}^k, \left\langle \hat{\mu}, \psi_{\partial_t^n X_{\rm m}}(t_k, .) \right\rangle - v_{\rm m}^k \right) \right\}. \tag{61}$$

The hypothesis made on f and n are more restrictive than those made for the method of simultaneous identification of f and $H(\partial t)$. Nevertheless, this method presents some interesting advantages. First, as the identification of f is uncoupled from the one of $H(\partial t)$, it is possible to adapt the choice of basis functions to the behavior of function f given by its pseudo-graph. Second, this method can also be used when the function f changes from one data set to the other. This identification method is still under study: more information about it will be found in (Casenave and Montseny [2009a]).

4.3 Iterative method

The identified quantities $\hat{\mu}$ and $\hat{f}$ can be used as initial values for an iterative identification method in order to improve the estimation quality. For example, let's assume that the estimated operator $\hat{H}(\partial_t)$ is exact; we can identify the function f again, this time independently from $H(\partial_t)$. From the so-identified function denoted $\hat{f}_1$, we identify $H(\partial_t)$, and we then reiterate until the quality of the estimation becomes sufficient. Note that the convergence of such a method is still under study.

4.4 Identified model

The physical system under consideration can be described by the model $\hat{H}(\partial_t)X = \hat{f}(u, X) + v$ where $\hat{H}$ and $\hat{f}$ are deduced from the identified parameters. Furthermore, by assuming that $\hat{H}(\partial_t)^{-1}$ is γ-diffusive of degree n', we deduce from the equivalent expression $X = \hat{H}(\partial_t)^{-1} \left(\hat{f}(u, X) + v \right)$ the following input-output state realization :

$$\begin{cases} \partial_t \psi = \gamma\, \psi + \hat{f}\left(u, \left\langle (\hat{\mu})^{-1} \sharp \delta^{n+n'}, \psi \right\rangle \right) + v,\ \psi(0, .) = 0, \\ X = \left\langle (\hat{\mu})^{-1} \sharp \delta^{n+n'}, \partial_t^{n'} \psi \right\rangle, \end{cases} \tag{62}$$

which is available up to the computation of $(\hat{\mu})^{-1} \sharp \delta^{n+n'}$ which can be numerically performed as shown in (Casenave [2009]).

5. APPLICATION TO A MODEL OF FLAME

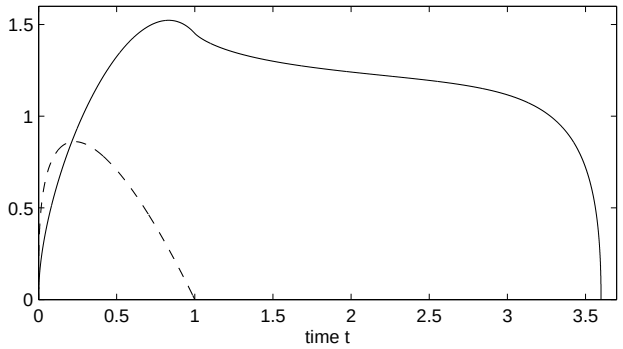
In this section, we illustrate the identification methods presented above by implementing them on data elaborated from numerical simulations of a complex dynamic phenomenon (Joulin [1985]).

5.1 The model under consideration

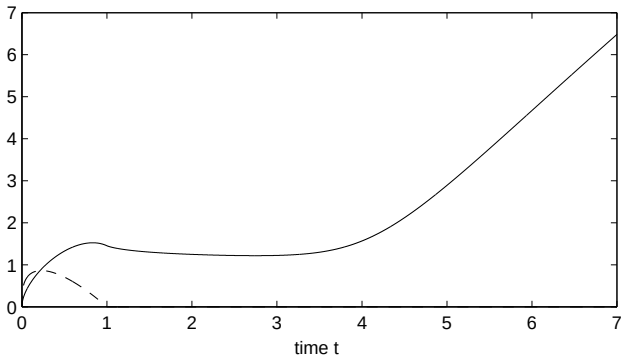
In (Joulin [1985]), Joulin elaborated a Volterra model to describe, in suitable thermodynamic conditions, the evolution of a spherical flame initiated by a source at point 0 in a mixture of reactive species. Under some reasonable physical hypothesis, such a phenomenon can be described by a system of two partial differential equations relating to the temperature and the mass density of the mixture. By considering the reactive zone as a thin sheet located on a sphere with radius $X(t)$, Joulin has established that when the flame is developing in free space, X is solution of the following nonlinear singular Abel-Volterra equation [7] ($e(t)$ designates the source strength at time t):

$$X(t)\partial_t^{\frac{1}{2}} = 2\, X(t) \ln X(t) + 2\, e(t) \quad \forall t > 0, \tag{63}$$

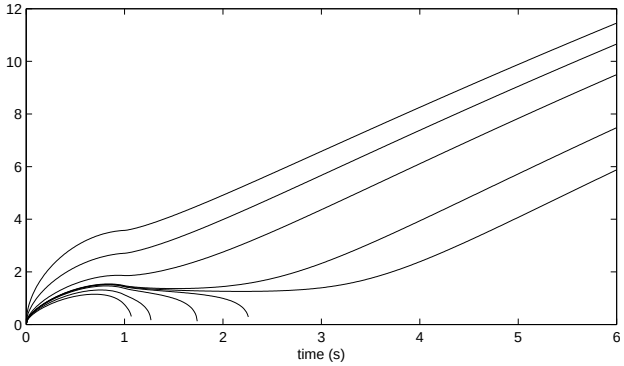
with the additional conditions: $X(0) = 0$, $e \geqslant 0$, $X \geqslant 0$ (whose physical interpretation is obvious). By denoting $H(\partial_t) = \frac{1}{2}\partial_t^{\frac{1}{2}}$, $f(u, X) := \ln X = f(X)$ and $u = \frac{e}{X}$, (63) can be formally rewritten under the form (45). We have:



(a) Source (- - -) and radius of the flame (—) with $E = 1.7390$.



(b) Source (- - -) and radius of the flame (—) with $E = 1.7393$.



(c) Trajectories X used for identification.

Fig. 2.

[7] Here adimensional for simplicity.

- $\frac{H(p)}{p} = \frac{1}{2}p^{-\frac{1}{2}}$ holomorphic in $\mathbb{C} \setminus \mathbb{R}^-$,
- and $\frac{H(p)}{p} \to 0$ when $|p| \to 0$ in $\mathbb{C} \setminus \mathbb{R}^-$;

so $H(\partial_t)$ is γ-diffusive of degree 1 for any γ of the sector form (44).

As shown in (Audounet et al. [1998]), the evolution problem (63) is well-posed, that is the solution X exists, is unique and depends continuously on e. It has also been shown that there exists a threshold relating to the power of the source e, beyond which the flame is developing whereas a quenching occurs below, as it is highlighted in Fig 2a and 2b, the source function used for the simulation being given as in (Audounet et al. [1998] by:

$$e(t) = E\, t^{0.3}(1-t)\, \mathbf{1}_{[0,1]}(t). \tag{64}$$

In real conditions, various perturbations are involved in the evolution of X (due for example to the loss of spatial symmetries), and both the convolution operator $H(\partial_t)$ and the function f will be more or less far from the ideal ones. So, an identification process can be justified if accuracy of the model is required.

5.2 Numerical identification results and comments

We consider in the sequel the problem of identification of $H(\partial_t)$ and/or f from data $(e_{\rm m}, X_{\rm m})$ obtained from highly accurate numerical simulations of (63) (Montseny [2010]).

The data are composed of 9 discrete simulated trajectories $(e_{\rm m}^{j,k}, X_{\rm m}^{j,k})_{k=0:K}$, $j = 1:9$ (see figure 2c), associated with 9 different sources e^j of the form [8] (64) with $E =$ 1.5, 1.6, 1.7, 1.73, 1.74, 1.75, 2.0, 3.0 and 5.0 respectively. The time step Δt is constant and equals to 10^{-5} and the final time is given by $t_K = 6$ ($K = 6 \times 10^5$).
The inputs are supposed to be known:

$$e_{\rm m}^{j,k} = e^j(t_k), \tag{65}$$

and the outputs are corrupted by additive measurement noise:

$$X_{\rm m}^{j,k} = X^j(t_k) + v^{j,k}, \tag{66}$$

where $\{v^{j,k}\}$ is a numerical white noise with zero mean and of standard deviation $s = 10^{-1}$.

To identify $H(\partial_t)$, we use $L = 80$ values ξ_l^L geometrically spaced to cover 5 decades from 10^0 to 10^5, and we consider a contour γ of the form (44) with $\alpha = \frac{\pi}{4}$. To identify f, we use 30 basis functions $\mathbf{k}^q, q = 1:30$, of hat type and one function basis $\mathbf{g}^1 = 1$.

The frequency response of the identified operator $\hat{H}(\partial_t)$ and the curve of the identified function $\hat{f}$ are given in figure 3 when f is supposed to be known (see section 3), when f and $H(\partial_t)$ are simultaneously identified (see section 4.1) and when $H(\partial_t)$ is first identified alone after cancellation of f (see section 4.2). For the 3 methods, the identification of $H(\partial_t)$ is good in the frequency band $[10^0, 10^2] \subset [\frac{2\,\pi}{K\,\Delta t}, \frac{2\,\pi}{2\,\Delta t}] = [1.05, 3.14\,10^5]$ (in the free noise case, this frequency band covers 4 decades). The identification of f is also good, even around the singularity of the function f at $X = 0$. Note however that the greater the value of X is, the larger the difference between $f(X)$ and its estimate $\hat{f}(X)$ is; this is partly due to the fact that there are more data $X_{\rm m}^{j,k}$ with value near from zero than others.

[8] Note that the source function is physically realistic but rather poor from the point of view of information, which strengthen the difficulty of the problem.

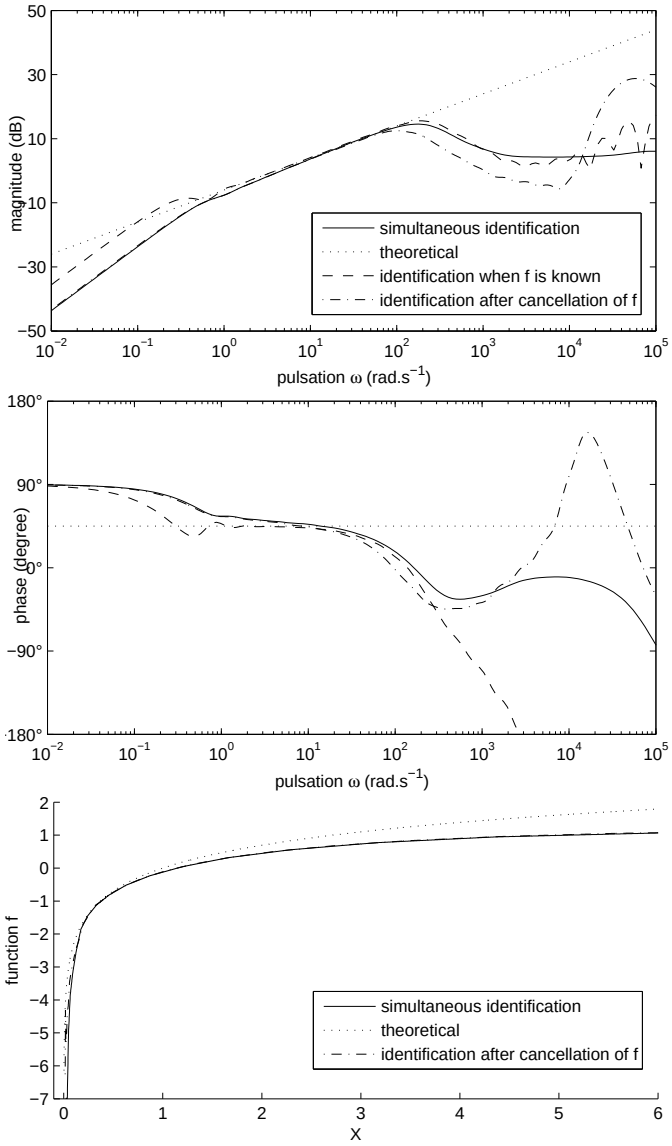


Fig. 3. Identification results.

6. CONCLUSION

In this paper, we have shown how to use the diffusive representation in order to identify the dynamic part of a system. Thanks of the linearity of the γ-symbol in the state realization of the operator $H(\partial_t)$, we can identify it by means of classical least squares methods. Several questions must yet be studied in order to improve the results. For example, among the most significant, the involved hilbertian norms should be judiciously chosen and adapted to the specific properties of the class of models under consideration; indeed, this choice is crucial in terms of sensitivity with respect to perturbations of any nature. The estimation bias could also be reduced. All these questions are currently under study.

REFERENCES

Audounet, J., Giovangigli, V., and Roquejoffre, J. (1998). A threshold phenomenon in the propagation of a point source initiated flame. *Physica D: Nonlinear Phenomena*, 121(3-4), 295–316.

Ben-Israel, A. and Greville, T. (2003). *Generalized inverses: Theory and applications*. Springer Verlag, New York, USA.

Carmona, P. and Coutin, L. (1998). Fractional brownian motion and the markov property. *Electronic Communications in Probability*, 3, 95–107.

Casenave, C. (2009). *Représentation diffusive et inversion opératorielle pour l'analyse et la résolution de problèmes dynamiques non locaux*. Ph.D. thesis, Université Paul Sabatier, Toulouse.

Casenave, C. and Montseny, G. (2008). Identification of Nonlinear Volterra Models by means of Diffusive Representation. In *17th IFAC World Congress*, 4024–4029. Seoul (Korea).

Casenave, C. and Montseny, G. (2009a). Diffusive Identification of Volterra Models by Cancellation of the Nonlinear Term. In *15th IFAC Symposium on System Identification, SYSID 2009*. Saint-Malo (France).

Casenave, C. and Montseny, G. (2009b). Optimal Identification of Delay-Diffusive Operators and Application to the Acoustic Impedance of Absorbent Materials. In *Topics in Time Delay Systems: Analysis, Algorithms and Control*, volume 388, 315. Springer Verlag, Lecture Notes in Control and Information Science (LNCIS).

Casenave, C. and Montseny, G. (2010). Identification and State Realization of Non-Rational Convolutive Models by Means of Diffusive Representation. *Submitted to IET Control Theory & Applications*.

Degerli, Y., Lavernhe, F., Magnan, P., and Farre, J. (Ap. 1999). Bandlimited $1/f^{\alpha}$ noise source. *Electronics Letters*, 35(7), 521–522.

Garcia, G. and Bernussou, J. (1998). Identification of the dynamics of a lead acid battery by a diffusive model. *European Series on Applied and Industrial Mathematics (ESAIM): Proceedings*, 5, 87–98.

Joulin, G. (1985). Point-source initiation of lean spherical flames of light reactants: an asymptotic theory. *Combustion science and technology*, 43(1-2), 99–113.

Lenczner, M. and Montseny, G. (2005). Diffusive realization of operator solutions of certain operational partial differential equations. *CRAS - Mathématiques*, 341(12), 737–740.

Levadoux, D. and Montseny, G. (2003). Diffusive formulation of the impedance operator on circular boundary for 2D wave equation. In *The Sixth International Conference on Mathematical and Numerical Aspects of Wave Propagation*. Jyväskylä (Finland).

Monin, A. and Salut, G. (1996). ARMA lattice identification: a new hereditary algorithm. *IEEE Transactions on Signal Processing*, 44(2), 360–370.

Montseny, E. (2010). Desingularization of non local dynamic models by means of operatorial transformations and application to a spherical flame model. *Submitted to 4th IFAC Symposium on System, Structure and Control*.

Montseny, G. (2005). *Représentation diffusive*. Hermes Science Publ.

Montseny, G. (2007). Diffusive representation for operators involving delays. In *Applications of Time-delay Systems*, 217–232. Springer Verlag, Lecture Notes in Control and Information Science (LNCIS).

Montseny, G. (Nov. 2004). Simple approach to approximation and dynamical realization of pseudodifferential time-operators such as fractional ones. *IEEE Trans on Circuits and System II*, 51(11), 613–618.

Mouyon, P. and Imbert, N. (2002). Identification of a 2d turbulent wind spectrum. In *Aerospace Science and Technology*, volume 6, 3599–605. Kansas City (USA).

Rumeau, A., Bidan, P., Lebey, T., Marchin, L., Barbier, B., and Guillemet, S. (2006). Behavior modeling of a $cacu_3ti_4o_{12}$ ceramic for capacitor applications. In *IEEE Conference on Electrical Insulation and Dielectric Phenomena*. Kansas City (Missouri USA).