Multivariate numerical differentiation

Samer Riachy, Mamadou Mboup, Jean-Pierre Richard

Multivariate numerical differentiation

Samer Riachy [14], Mamadou Mboup [24], Jean-Pierre Richard [34]

Abstract

We present an innovative method for multivariate numerical differentiation *i.e.* the estimation of partial derivatives of multidimensional noisy signals. Starting from a local model of the signal consisting of a truncated Taylor expansion, we express, through adequate differential algebraic manipulations, the desired partial derivative as a function of iterated integrals of the noisy signal. Iterated integrals provide noise filtering. The presented method leads to a family of estimators for each partial derivative of any order. We present a detailed study of some structural properties given in terms of recurrence relations between elements of a same family. These properties are next used to study the performance of the estimators. We show that some differential algebraic manipulations corresponding to a particular family of estimators leads implicitly to an orthogonal projection of the desired derivative in a Jacobi

[1]Ecole National Supérieure de l'Electronique et de ses Applications, ECS-Lab/EA-3649, 6, avenue du Ponceau, 95014, Cergy-Pontoise, France. **Questions, comments, or corrections to this document may be directed to that email address: Samer.Riachy@ensea.fr**

[2]Université de Reims Champagne Ardenne, CReSTIC-UFR SEN, BP 1039 Moulin de la Housse, F-51687 Reims cedex 2, France. Mamadou.Mboup@univ-reims.fr

[3]Université Lille Nord de France, ECLille, LAGIS, UMR 8146 CNRS, Ecole Centrale de Lille, BP 48, Cité Scientifique, 59651 Villeneuve d'Ascq, France. Jean-Pierre.Richard@ec-lille.fr

[4]Samer Riachy, Mamadou Mboup & Jean-Pierre Richard are with the team Non-A of INRIA Lille-Nord Europe, France.

Preprint submitted to Journal of Computational and Applied MathematicsSeptember 13, 2011

polynomial basis functions, yielding an interpretation in terms of the popular least squares. This interpretation allows one to 1) explain the presence of a spacial delay inherent to the estimators and 2) derive an explicit formula for the delay. We also show how one can devise, by a proper combination of different elementary estimators of a given order derivative, an estimator giving a delay of any prescribed value. The simulation results show that delay-free estimators are sensitive to noise. Robustness with respect to noise can be highly increased by utilizing voluntary-delayed estimators. A numerical implementation scheme is given in the form of finite impulse response digital filters. The effectiveness of our derivative estimators is attested by several numerical simulations.
Keywords: Numerical differentiation, operational calculus, multivariable signals, orthogonal polynomials, inverse problems, least squares, finite impulse response filters.

---

1. INTRODUCTION

Partial derivatives estimation of multivariate signals is a recurrent problem in the fields of engineering and applied mathematics, as for example, in automatic control, signal and image processing. It is known that the differentiation problem, as opposed to integration, is unstable in the presence of noisy data. It has thus the property of ill-posedness.

When the noise level is low, the most common approach is the use of finite differences techniques which present the advantage of low computational cost and easy implementation. They are used in many problems especially in variational methods in image processing for motion estimation [8], [22], [40], image reconstruction and denoising [4] [41] and image segmentation [7], [33] etc. We may also cite other approaches more specific to a particular field. In automatic control, for example, we mention the model based observers [6] and the sliding modes techniques [27]. However, in many practical applications, the noise influence cannot be neglected. It becomes then necessary to consider methods that are more robust to noise. The literature about differentiation is vast, we recall some important approaches in the monovariate case.

An integral operator, known as Lanczos generalized derivative, was proposed in [26]. It is defined by :

$$D_h f(x) = \frac{3}{2h} \int_{-1}^{1} t f(x+t) dt, \tag{1}$$

and approximates $f^{(1)}(x)$ in the sense $f^{(1)}(x) = D_h f(x) + \mathcal{O}(h^2)$. Generalization to higher order derivatives was proposed in [37]:

$$D_h^{(n)} f(x) = \frac{1}{h^n} \int_{-1}^{1} \rho_n(t) f(x+ht) dt, n = 1, 2, \cdots \tag{2}$$

The above formula approximates the n^{th} order derivative $f^{(n)}(x)$ such that $f^{(n)}(x) = D_h^{(n)} f(x) + \mathcal{O}(h^n)$. It was shown that $\rho_n(t)$ is proportional to Legender polynomials of order n. Further studies can be found in [39].

Moreover, differentiation can also be cast into a least squares problem [10], [24], [9], [3]. Robustness with respect to noisy data can be increased by introducing a regularization term which extracts from all possible solutions (approximations) those who, for example, have bounds on the function and/or its derivative. A well known regularization is due to Tikhonov and can be cast as follows. Find g an approximation of $\frac{df}{dt}$ such that

$$\|Ag - f\|^2 + \alpha\|g\| + \beta\|g^{(1)}\|^2 \tag{3}$$

is minimum, where A is an appropriate operator. The regularization parameters α and β, if tuned properly, results in an efficient derivative estimator although tuning is a difficult task. However, the solution cannot be computed in real time.

This paper proposes a different approach. We assume that the structured, information bearing, component of a noisy signal admit a (multivariate) convergent Taylor expansion. In order to estimate the n^{th} order partial derivative, we rewrite the N^{th} (here N and n are multi-indices and $N \geq n$) order truncation of the Taylor expansion in the operational domain using a multidimensional Laplace transform. Adequate differential algebraic operations then allow us to isolate, back in the spacial domain, the desired partial derivative at a given point as a function of multiple iterated integration on the noisy measured signal. Our approach is thus based on pointwise derivative estimation.

This paper constitute an extension of [31] to multidimensional signals, it is a continuation of [5], [35] and [34]. An interesting contribution to multivariate numerical differentiation can be found in [16] and [25]. The matters in this paper are inspired from techniques initiated by M. Fliess *et al.* in 2003 [20] in control theory. Those techniques which are of algebraic flavor are promising in signal processing and estimation [16], [17], [18], [28], [29], [30], [31], [34], control [12], [13], [21] fault detection [19], and finance [14], [15].

To fix the subsequent notations and introduce the basic steps of our approach, we consider the following simple example. Let $I(\boldsymbol{x}) = I(x_1, x_2)$ be a bidimensional signal with two independent variables x_1 and x_2. Its Taylor series expansion of order $N = (1, 1)$ around $(0, 0)$, denoted $\mathbb{I}_N(x_1, x_2)$, writes:

$$\mathbb{I}_N(x_1, x_2) = I_0(\bar{0}) + I_{x_1}(\bar{0})x_1 + I_{x_2}(\bar{0})x_2,$$

where $(\bar{0}) = (0, 0)$, $I_0(\bar{0}) = I(0, 0)$, $I_{x_1}(\bar{0}) = \frac{\partial I}{\partial x_1}(0, 0)$ and $I_{x_2}(\bar{0}) = \frac{\partial I}{\partial x_2}(0, 0)$. In the operational domain one obtains:

$$\hat{\mathbb{I}}_N(s_1, s_2) = \frac{I_0(\bar{0})}{s_1 s_2} + \frac{I_{x_1}(\bar{0})}{s_1^2 s_2} + \frac{I_{x_2}(\bar{0})}{s_1 s_2^2} \tag{4}$$

where $\hat{\mathbb{I}}_N$ is the operational analogue of $\mathbb{I}_N$. Let us isolate $I_{x_1}(\bar{0})$ by multiplying (4) by $s_1 s_2$ and then differentiating once with respect to s_1. The right-hand side of (4) reduces to $\frac{-I_{x_1}(\bar{0})}{s_1^2}$. Applying the same operations to the left-hand side of (4) leads to:

$$s_2 \hat{\mathbb{I}}_N(s_1, s_2) + s_1 s_2 \frac{\partial \hat{\mathbb{I}}_N(s_1, s_2)}{\partial s_1} = -\frac{I_{x_1}(\bar{0})}{s_1^2}. \tag{5}$$

Note that multiplying by s_1 (respectively by s_2) corresponds to differentiation with respect to x_1 (resp. x_2) in the spacial domain. Differentiation is not desirable in the presence of noise. For this reason, we multiply (5) by $\frac{1}{s_1^2 s_2^2}$. Back in the spacial domain, the following form is obtained:

$$\int_0^{X_1} \int_0^{X_2} (X_1 - 2x_1)\mathbb{I}_N(x_1, x_2) dx_1 dx_2 = -I_{x_1}(\bar{0}) \frac{X_1^3}{3!} X_2.$$

Now, if we replace the noise-free Taylor series model $\mathbb{I}_N(x_1, x_2)$ by the actual noisy measurement $J(x_1, x_2)$, we obtain an estimate $\tilde{I}_{x_1}(\bar{0})$ of $I_{x_1}(\bar{0})$

$$\tilde{I}_{x_1}(\bar{0}) = \frac{-6}{X_1} \int_0^1 \int_0^1 (1 - 2x_1) J(X_1 x_1, X_2 x_2) dx_1 dx_2, \tag{6}$$

as a function of the estimation window parameters X_1 and X_2 (here we have used a change of variables to normalize the integrals over $[0, 1] \times [0, 1]$).

Following the terminology introduced in [31], estimators in the form of (6) will be called algebraic partial derivative estimators. In section 2, we recall the multi-index notation, the multivariate Laplace transform and introduce a multivariate version of Jacobi polynomials. In section 3, we introduce the methodology and point out that it provides:

- pointwise estimators,
- a family of estimators to any given order of derivation.

We provide in section 4 a detailed study of some structural properties of our estimators. In section 5 we forge a link with least squares using multivariate Jacobi polynomials with special weighing functions. This link with least squares enables us to show the existence of a spacial delay inherent to a particular family of estimators and provide a formula to quantify the delay. We consider also affine combinations of estimators of a given order derivative, the weights involved in the combination can be parameterized by a single parameter denoted $\boldsymbol{\xi}$. Depending on the choice of $\boldsymbol{\xi}$, we provide:

- delay-free estimators,
- estimators reducing the mis-modeling error induced by the truncation of the Taylor expansion,
- estimators reducing the noise influence in section 6.

Unfortunately, simulations presented in section 7 will show that delay-free estimators are sensitive to noises. Robustness to noises can be highly increased by tolerating a delay through an adequate choice of $\boldsymbol{\xi}$. Thus the parameter

$\boldsymbol{\xi}$ can be seen as an explicit regularization parameter. Unlike, classical least squares where a good choice of regularization parameters is difficult to accomplish, we provide an explicit formula for $\boldsymbol{\xi}$. In section 7, a numerical implementation scheme in the form of a finite impulse response linear filter will be given followed by several numerical simulations. For the clarity of the presentation, all the proofs are deferred to an appendix.

2. Preliminaries

This section recalls the multi-index notation, the multivariate Laplace transform and introduce a multivariate version of Jacobi's polynomials.

2.1. Multi-index notation

Let $\alpha = (\alpha_1, \cdots, \alpha_r)$ be an r-tuple of nonnegative integers $\alpha_m,\ m = 1 \cdots, r;\ m, r \in \mathbb{N}$. We call α a multi-index. We fix some notations. The symbol in bold $\boldsymbol{x}$ denotes a vector in $\mathbb{R}^r$ representing the spacial domain of the multivariate signal. The Laplace (or operational domain) variable is denoted by $\boldsymbol{s} = (s_1, \cdots, s_r)$, where $r \in \mathbb{N}$ stands for the dimension of the multivariate signal. The bold symbol $\boldsymbol{X} \in \mathbb{R}^r$ represents the length of the integration domain. The letters α, κ, μ, l, q, N and n are multi-indices and $m \in \mathbb{N}$ will be used as a pointer varying from 1 to r. The multi-indices α, κ, μ, l, q, N and n affected by the subscript m as for example κ_m, denotes the m^{th} element of κ, *i.e* a nonnegative integer.

For multi-indices $\alpha, \beta \in \mathbb{N}^r$ one defines:

1. Componentwise sum and difference: $\alpha \pm \beta = (\alpha_1 \pm \beta_1, \cdots \alpha_r \pm \beta_r)$.
2. Partial order $\alpha \leq \beta \Leftrightarrow \alpha_m \leq \beta_m, \forall m \in \{1, \cdots, r\}$.

3. Given $\boldsymbol{x} = (x_1, \cdots, x_r) \in \mathbb{R}^r$, we have that $\boldsymbol{x}^\alpha = x_1^{\alpha_1} \cdots x_r^{\alpha_r}$.
4. The total degree of $\boldsymbol{x}^\alpha$ is given by $|\alpha| = \alpha_1 + \cdots + \alpha_r$.
5. Factorial: $\alpha! = \alpha_1! \cdots \alpha_r!$
6. Binomial coefficient:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \cdots \begin{pmatrix} \alpha_r \\ \beta_r \end{pmatrix}. \tag{7}$$

7. $\bar{b} = (b, \cdots, b),\ b \in \mathbb{N},\ \bar{b} \in \mathbb{N}^r$.
8. For $\boldsymbol{x} = (x_1, \cdots, x_r)$ and $\boldsymbol{X} = (X_1, \cdots, X_r) \in \mathbb{R}^r$, $\int_{\bar{0}}^{\boldsymbol{X}} f(\boldsymbol{x}) d\boldsymbol{x} = \underbrace{\int_0^{X_1} \cdots \int_0^{X_r}}_{r} f(x_1, \cdots, x_r) dx_1 \cdots dx_r$.
9. Higher-order partial derivative: $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_r^{\alpha_r}$ where $\partial_m^{\alpha_m} := \frac{\partial^{\alpha_m}}{\partial x_m^{\alpha_m}}$.
10. Denote by $1_m \in \mathbb{N}^r$ the multi-index with zeros for all elements except the m^{th} one *i.e.* $1_m = (0, \cdots, 0, 1, 0, \cdots, 0)$.
11. The tensor product of 2 vectors $\boldsymbol{u},\ \boldsymbol{v} \in \mathbb{R}^r$ is defined by: $\boldsymbol{u} \otimes \boldsymbol{v} = (u_1\boldsymbol{v}, \cdots, u_r\boldsymbol{v})\ \in \mathbb{R}^{r^2}$. $\boldsymbol{u} \otimes \boldsymbol{v} = (u_1v_1, \cdots, u_1v_r, u_2v_1, \cdots, u_2v_r, \cdots, u_rv_r)$.

2.2. Multivariate Laplace transform

Given $\boldsymbol{s} = (s_1, \cdots, s_r) \in \mathbb{C}^r$, $\boldsymbol{x} = (x_1, \cdots, x_r) \in \mathbb{D}^r \subset \mathbb{R}^r$ and a multivariable function $f(\boldsymbol{x}) : \mathbb{D}^r \subset \mathbb{R}^r \to \mathbb{R}$. We recall that the multivariate Laplace transform is given by

$$\mathcal{L}(I(\boldsymbol{x})) = \hat{I}(\boldsymbol{s}) = \int_{\mathbb{R}^{n+}} f(\boldsymbol{x}) \exp^{-\boldsymbol{s}.\boldsymbol{x}^T} d\boldsymbol{x}. \tag{8}$$

Note that the terminology “time domain *vs* frequency domain” is not adequate. As it was noticed in the introduction, we use the terminology “spacial (or spacio-temporal) domain *vs* operational domain”. The multivariate Laplace transform satisfies:

1. Given $\boldsymbol{x} = (x_1, \cdots, x_r)$, $\boldsymbol{s} = (s_1, \cdots, s_r)$ and multi-index $\alpha = (\alpha_1, \cdots, \alpha_r)$, we have the following:

$$\mathcal{L}\left(\frac{\boldsymbol{x}^\alpha}{\alpha!}\right) = \frac{1}{\boldsymbol{s}^{\alpha+\bar{1}}} \tag{9}$$

2. Let $\boldsymbol{x} = (x_1, \cdots, x_r)$, $\boldsymbol{s} = (s_1, \cdots, s_r)$, $\boldsymbol{X} = (X_1, \cdots, X_r)$, and let $\alpha = (\alpha_1, \cdots, \alpha_r)$, and $\beta = (\beta_1, \cdots, \beta_r)$ be two multi-indices. Given a multivariable function $I(\boldsymbol{x})$ and its corresponding Laplace transform $\hat{I}(\boldsymbol{s}) = \mathcal{L}(I(\boldsymbol{x}))$, we have the following:

$$\mathcal{L}^{-1}\frac{1}{\boldsymbol{s}^\alpha}\frac{\partial^\beta \hat{I}(\boldsymbol{s})}{\partial \boldsymbol{s}^\beta} = \frac{1}{(\alpha-\bar{1})!}\int_{\bar{0}}^{X} (\boldsymbol{X}-\boldsymbol{x})^{\alpha-\bar{1}}(-\boldsymbol{x})^\beta I(\boldsymbol{x})d\boldsymbol{x}.$$

2.3. Multivariate orthogonal Jacobi polynomials and least squares

This section introduce a multivariate version of Jacobi's polynomials. They are used to affect a least squares interpretation to a particular class of our estimators. Given multi-indices α, β, n, $p \in \mathbb{N}^r$ and $\boldsymbol{x} = (x_1, \cdots, x_r)$. Let δ_{np} denote a multivariate version of the Kronecker symbol *i.e* $\delta_{np} = 1$ if $(n_1, \cdots, n_r) = (p_1, \cdots, p_r)$ element-wise and $\delta_{np} = 0$ otherwise. A multivariate version of the Jacobi polynomials on the interval $[0,1]^r$ is given by the partial differential equation (Rodriguez formula) which seems new:

$$(\bar{1}-\boldsymbol{x})^\alpha \boldsymbol{x}^\beta \mathcal{P}_n^{\{\alpha,\beta\}}(\boldsymbol{x}) = \frac{(-\bar{1})^n}{n!}\partial^n[(\bar{1}-\boldsymbol{x})^{n+\alpha}\boldsymbol{x}^{n+\beta}]. \tag{10}$$

Those polynomials constitute an orthogonal set, on the interval $[0,1]^r$ with respect to the weight function $\omega(\boldsymbol{x}) = (\bar{1}-\boldsymbol{x})^\alpha \boldsymbol{x}^\beta$, i.e. they satisfy:

$$< \mathcal{P}_n^{\{\alpha,\beta\}}, \mathcal{P}_p^{\{\alpha,\beta\}} > \doteq \int_{\bar{0}}^{\bar{1}} \mathcal{P}_n^{\{\alpha,\beta\}}(\boldsymbol{x})\omega(\boldsymbol{x})\mathcal{P}_p^{\{\alpha,\beta\}}(\boldsymbol{x})d\boldsymbol{x} = \delta_{np}, \tag{11}$$

The norm induced by (11), and denoted $\| \bullet \|$, writes:

$$\|\mathcal{P}_n^{\{\alpha,\beta\}}\|^2 = \int_{\bar{0}}^{\bar{1}} \mathcal{P}_n^{\{\alpha,\beta\}}(\boldsymbol{x})\omega(\boldsymbol{x})\mathcal{P}_n^{\{\alpha,\beta\}}(\boldsymbol{x})d\boldsymbol{x}.$$

Let $P_{n_m}^{\{\alpha_m,\beta_m\}}(x_m)$ denote the standard, one dimensional, Jacobi polynomials, the multivariate polynomials are obtained directly by:

$$\mathcal{P}_n^{\{\alpha,\beta\}}(\boldsymbol{x}) = \prod_{m=1}^{r} P_{n_m}^{\{\alpha_m,\beta_m\}}(x_m).$$

The proof is straightforward upon expanding (10).

3. Partial derivatives estimation

Let us consider a noisy signal $J(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \cdots, x_r)$, and assume that it is constituted of a structured part $I(\boldsymbol{x})$, with an additive noise $\varpi(\boldsymbol{x})$:

$$J(\boldsymbol{x}) = I(\boldsymbol{x}) + \varpi(\boldsymbol{x}). \tag{12}$$

We assume that $I(\boldsymbol{x})$ admits a convergent Taylor series expansion at the origin:

$$I(\boldsymbol{x}) = \sum_{\alpha=\bar{0}}^{N} \frac{\partial^\alpha I(\bar{0})}{\alpha!}\boldsymbol{x}^\alpha + e_R = \sum_{\alpha=\bar{0}}^{N} I_{\boldsymbol{x}^\alpha}(\bar{0})\frac{\boldsymbol{x}^\alpha}{\alpha!} + e_R, \tag{13}$$

where $I_{\boldsymbol{x}^\alpha}(\bar{0}) = \partial^\alpha I(\bar{0})$ are successive partial derivatives at zero and e_R is the truncation error. Let us neglect e_R for a moment and write:

$$\mathbb{I}_N(\boldsymbol{x}) = \sum_{\alpha=\bar{0}}^{N} I_{\boldsymbol{x}^\alpha}(\bar{0})\frac{\boldsymbol{x}^\alpha}{\alpha!} \tag{14}$$

Suppose that there exist a bounded linear operator $\mathcal{O}[\bullet]$ that annihilate from (14) the terms $I_{\boldsymbol{x}^\alpha}(\bar{0})$; $\alpha \neq n$. We obtain thus $I_{\boldsymbol{x}^\alpha}(\bar{0}) = \mathcal{O}(\mathbb{I}_N(\boldsymbol{x}))$ and consequently an estimate $\tilde{I}_{\boldsymbol{x}^n}(\bar{0})$ of $I_{\boldsymbol{x}^n}(\bar{0})$ by $\tilde{I}_{\boldsymbol{x}^n}(\bar{0}) = \mathcal{O}[J(\boldsymbol{x})]$. Moreover, the derivative estimation $\tilde{I}_{\boldsymbol{x}^n}(\bar{0} + \mathfrak{x})$ at another point $\mathfrak{x}$ different from $\bar{0}$ can be given by $\tilde{I}_{\boldsymbol{x}^\alpha}(\bar{0}+\mathfrak{x}) = \mathcal{O}[J(\boldsymbol{x}+\mathfrak{x})]$. If in addition the operator $\mathcal{O}[\bullet]$ is integral, it permits to filter the additive noise $\varpi(\boldsymbol{x})$. Eliminating the undesired terms

$(I_{\boldsymbol{x}^\alpha}(\bar{0}),\ \alpha \neq n)$ from (13) can be done in the Laplace operational domain through straightforward differential algebraic manipulations. For this reason, we apply the multivariate Laplace transform (8) on (14):

$$\hat{\mathbb{I}}_N(\boldsymbol{s}) = \sum_{\alpha=\bar{0}}^{N} \frac{I_{\boldsymbol{x}^\alpha}(\bar{0})}{\boldsymbol{s}^{\alpha+\bar{1}}}, \tag{15}$$

where $\hat{\mathbb{I}}_N(\boldsymbol{s})$ is the operational analog of $\mathbb{I}_N(\boldsymbol{x})$. By examining (15) it can be seen that the terms $I_{\boldsymbol{x}^\alpha}(\bar{0})$ are divided by different powers of $\boldsymbol{s}$. Thus if one can choose adequate multiplication with powers of $\boldsymbol{s}$ and successive higher partial differentiation with respect to $\boldsymbol{s}$ one is able to isolate $I_{\boldsymbol{x}^n}(\bar{0})$. Those successive operations will be called annihilators. There exists many annihilators corresponding to $I_{\boldsymbol{x}^n}$. In this paper we will focus on a special class given by the following proposition:

Proposition 3.1. *Let κ, μ, N, and n be multi-indices in $\mathbb{N}^r$, the differential operator*

$$\Pi^{N,n}_{\kappa,\mu} = \frac{1}{\boldsymbol{s}^{N+\mu+\bar{1}}} \partial^{n+\kappa} \frac{1}{\boldsymbol{s}} \partial^{N-n} \boldsymbol{s}^{N+\bar{1}} \tag{16}$$

annihilate from (15) all the terms $I_{\boldsymbol{x}^\alpha}(\bar{0}),\ \alpha \neq n$ and yields to:

$$\Pi^{N,n}_{\kappa,\mu} \hat{\mathbb{I}}_N(\boldsymbol{s}) = \frac{(-1)^{(n+\kappa)}(n+\kappa)!(N-n)!}{\boldsymbol{s}^{\mu+\kappa+N+n+\bar{2}}} I_{\boldsymbol{x}^n}(\bar{0}). \tag{17}$$

Remark 3.1. *If the truncation error is non zero, then equation (17) no longer holds. Moreover, only noisy observation $J(\boldsymbol{x})$ is available. Replacing $\hat{\mathbb{I}}_N(\boldsymbol{s})$ in (17) by its non truncated and noisy counterpart $\hat{J}(\boldsymbol{s})$ then leads to the operational estimator $\tilde{I}(\bar{0};\kappa,\mu;N)$ of $I_{x^n}(\bar{0})$:*

$$\frac{(-1)^{(n+\kappa)}(n+\kappa)!(N-n)!}{\boldsymbol{s}^{\mu+\kappa+N+n+\bar{2}}} \tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \Pi^{N,n}_{\kappa,\mu} \hat{J}(\boldsymbol{s}), \tag{18}$$

Here, we use the notation $\tilde{I}(\bar{0};\kappa,\mu;N)$ to quote the dependance of the estimator on the parameters κ, μ and N.

Remark 3.2. *If $N = n$, the operational estimator is given by*

$$\frac{(-1)^{(n+\kappa)}(n+\kappa)!}{\boldsymbol{s}^{\mu+\kappa+2n+\bar{2}}}\tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;n) = \frac{1}{\boldsymbol{s}^{n+\mu+\bar{1}}}\partial^{n+\kappa}\boldsymbol{s}^n\hat{J}(s). \tag{19}$$

and it is termed ***minimal*** *because it is based on an n^{th} order Taylor series truncation.*

Remark 3.3. *To $\Pi^{N,n}_{\kappa,\mu}$ in (16), (17) correpond, in the spacial domain, an integral operator $\mathcal{O}^{N,n}_{\kappa,\mu}(\bullet)$ such that $I_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \mathcal{O}^{N,n}_{\kappa,\mu}(\mathbb{I}_N(\boldsymbol{x}))$.*

Recapitulating : By reconsidering the truncation error e_R and the noise influence $\varpi(\boldsymbol{x})$ together with relation (17), the partial derivative $I_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N)$ can be written as :

$$I_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \mathcal{O}^{N,n}_{\kappa,\mu}(\mathbb{I}_N(\boldsymbol{x})) + \mathcal{O}^{N,n}_{\kappa,\mu}(e_R) + \mathcal{O}^{N,n}_{\kappa,\mu}(\varpi(\boldsymbol{x})). \tag{20}$$

We can see here that two kind of errors may degrade the quality of the derivative estimation. The aim of the forthcoming sections is to concentrate on the minimization of these errors. We will see that estimators minimizing the truncation error are not generally the best suited for filtering the noise influence and vice-versa. Then the choice of an estimator for a particular application obey to a compromise.

4. Structural properties and recurrence relations

First it is shown that non-minimal ($N > n$) algebraic estimators based on an N^{th} order Taylor model is an affine combination of minimal ($N = n$) estimators with different κ and μ.

Theorem 4.1. *Let N, n, q, l, κ, and μ be multi-indices in $\mathbb{N}^r$ with $n \leq N$. Then we have:*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l \tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa_l,\mu_l;n), \quad \lambda_l \in \mathbb{Q} \tag{21}$$

where, $q = N - n$, $\kappa_l = \kappa + q - l$, and $\mu_l = \mu + l$. Moreover, if $q \leq n + \kappa$, then the coefficients $\boldsymbol{\lambda}_l$,satisfy

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l = 1.$$

Moreover, by excluding the trivial case where all the $\boldsymbol{\lambda}_l$ are equal to zero except one, we have:

$$\min_l \boldsymbol{\lambda}_l < 0. \tag{22}$$

Now a recurrence relation is given between estimators based on N^{th} and $(N - 1_m)^{th}$ order Taylor model.

Theorem 4.2. *Given multi-indices κ, μ, N and $n \in \mathbb{N}^r$, and an integer $m \in [1, r]$. We have:*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \mathfrak{a}_m \tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu+1_m;N-1_m) + \mathfrak{b}_m \tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa+1_m,\mu;N-1_m) \tag{23}$$

where $\mathfrak{a}_m = \frac{N_m+\kappa_m+1}{N_m-n_m}$, $\mathfrak{b}_m = 1 - \mathfrak{a}_m$ and $N_m > n_m$.

The meaning of the notation 1_m is explained in subsection 2.1.

There exist another recurrence between the estimators based on $(N-\bar{1})^{th}$ and N^{th} order Taylor series expansions. In order to introduce the recurrence relation in a compact form, we state the following lemma.

Let $\mathbf{L}$ be a collection of multi-indices in $\mathbb{N}^r$ such that :

$$\mathbf{L} = \{l; |l| \leq r \text{ AND } l! \leq 1\}. \tag{24}$$

Define a binary relation (denoted $\prec$) on $\mathbf{L}$ such as : given l and $l' \in \mathbf{L}$ then

$$l \prec l' \quad \text{IFF} \quad \sum_{i=1}^{r} 2^{r+1-i} l(i) < \sum_{j=1}^{r} 2^{r+1-j} l'(j), \tag{25}$$

where $l(i)$ is the i^{th} element of l.

Lemma 4.1. *The set* $\mathbf{L}$ *equipped with* $\prec$ *is a totally ordered set.*

Accordingly, the elements of $\mathbf{L}$ can be arranged in an increasing order and indexed such that $\mathbf{L}(1) = \bar{0}$ and $\mathbf{L}(2^r) = \bar{1}$ where $card(\boldsymbol{L}) = 2^r$. On the other hand, let $u_m = (\mathfrak{a}_m, \mathfrak{b}_m);\ m = 1, \cdots, r$ and $\boldsymbol{u} = u_1 \otimes \cdots \otimes u_m \otimes \cdots \otimes u_r \in \mathbb{Q}^{2^r}$. Denote by $\boldsymbol{u}(i)$ the i^{th} element of $\boldsymbol{u}$. We have:

Theorem 4.3. *Given multi-indices* κ, μ, N *and* $n \in \mathbb{N}^r$. *We have:*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; N) = \sum_{i=1}^{2^r} \boldsymbol{u}(i) \tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa + \mathbf{L}(i), \mu + \mathbf{L}(2^r + 1 - i); N - \bar{1}). \tag{26}$$

5. Least squares interpretation and shifted estimators

Consider the subspace of $L_2([0,1])$ spanned by monovariate Jacobi polynomials:

$$\mathcal{H}_{q_m} = span\left\{ P_0^{\{\alpha_m, \beta_m\}}(x_m), \cdots, P_{q_m}^{\{\alpha_m, \beta_m\}}(x_m) \right\},$$

equipped by the inner product :

$$< P_{n_m}, P_{p_m} > \doteq \int_0^1 P_{n_m}(x_m) \omega(x_m) P_{p_m}(x_m) dx_m. \tag{27}$$

$\mathcal{H}_{q_m}$ is a reproducing kernel Hilbert space. Its reproducing kernel is given by :

$$\mathcal{K}_{q_m}(\xi_m, x_m) = \sum_{i=0}^{q_m} \frac{P_i^{\{\alpha_m,\beta_m\}}(\xi_m) P_i^{\{\alpha_m,\beta_m\}}(x_m)}{\|P_i^{\{\alpha_m,\beta_m\}}\|^2}.$$

Let $\boldsymbol{\mathcal{H}}_q$ be a tensor product of r (mono dimensional) subspaces $\mathcal{H}_{q_m}$ of $L_2([0,1])$:

$$\boldsymbol{\mathcal{H}}_q = \mathcal{H}_{q_1} \otimes \cdots \otimes \mathcal{H}_{q_r},$$

It is evident that $\boldsymbol{\mathcal{H}}_q$ is a reproducing kernel Hilbert space. Its reproducing kernel is given by :

$$\boldsymbol{\mathcal{K}}_q(\boldsymbol{\xi}, \boldsymbol{x}) = \mathcal{K}_{q_1}(\xi_1, x_1) \times \cdots \times \mathcal{K}_{q_r}(\xi_r, x_r) = \sum_{l=\bar{0}}^{q} \frac{\mathcal{P}_l^{\{\alpha,\beta\}}(\boldsymbol{\xi}) \mathcal{P}_l^{\{\alpha,\beta\}}(\boldsymbol{x})}{\|\mathcal{P}_l^{\{\alpha,\beta\}}(\boldsymbol{x})\|^2}, \quad (28)$$

where l is a multi-index and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_r) \in [0,1]^r$.

It is now possible to define a q^{th} order least squares approximation of a function $F(\boldsymbol{X}\boldsymbol{x})$, where $\boldsymbol{x} \in [0,1]^r$. It is noted $F_{\text{LS},q}$ and given by :

$$F_{\text{LS},q}(\boldsymbol{X}\boldsymbol{\xi}) \doteq \sum_{l=\bar{0}}^{q} \frac{\langle \mathcal{P}_l^{\{\alpha,\beta\}}(\boldsymbol{x}), F(\boldsymbol{X}\boldsymbol{x}) \rangle}{||\mathcal{P}_l^{\{\alpha,\beta\}}||^2} \mathcal{P}_l^{\{\alpha,\beta\}}(\boldsymbol{\xi}), \quad (29)$$

where $\boldsymbol{\xi} \in [0,1]^r$.

We will now show that the spacial analogue of (18) correspond to a dot product of $J(\boldsymbol{x})$ with the reproducing kernel (28). This projection leads us to detect spacial delay inherent to minimal algebraic estimators (19). By spacial shifting we mean that $\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; n)$ given by (19) although designed from a Taylor expansion around $\bar{0}$ corresponds in fact to a derivative estimation at some point (to be determined) $\bar{0} + \boldsymbol{\xi'}$ with $\boldsymbol{\xi'} = (\xi'_1, \cdots, \xi'_r)$ different from $\bar{0}$.

Proposition 5.1. *Given multi-indices κ, μ, α and $n \in \mathbb{N}^r$. Let $(\partial^n I)_{LS,1}(\boldsymbol{X}\boldsymbol{\xi})$ denote the first order least-squares polynomial approximation of n^{th} order* *derivative on the interval* $[\bar{0}, \boldsymbol{X}]$. *Then the spacial analogue of the minimal* n^{th} *order algebraic derivative estimator (19) is given by:*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; n) = (\partial^n I)_{LS,1}(\boldsymbol{X\xi'}) + \varpi'(\boldsymbol{x}),$$

where $\boldsymbol{\xi'} = (\xi'_1, \cdots, \xi'_r)$ *and*

$$\xi'_m = \frac{\kappa_m + n_m + 1}{\mu_m + \kappa_m + 2(n_m + 1)}, \quad m = 1, \cdots, r \tag{30}$$

are the roots of $\mathcal{P}_{\bar{1}}^{\{\kappa,\mu\}}(\boldsymbol{\xi}) = 0$ *and* $\varpi'(\boldsymbol{x})$ *is the noise contribution. (r roots corresponding each to one of the r Jacobi polynomials :* $\mathcal{P}_{\bar{1}}^{\{\kappa,\mu\}}(\boldsymbol{x})$*).*

We arrive here to a remarkable result, if in some application a combination of partial derivatives is used as in the Laplacian estimation for example, the delays ξ_m (30) has to be adjusted before one aim a high quality estimation.

We show in the following proposition that non minimal algebraic estimators (18) are delay free, *i.e.* $\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; N)$ correspond to the derivative estimation at the point $\bar{0}$.

Proposition 5.2. *Let* κ, μ, q, N, *and* n *be multi-indices* $\in \mathbb{N}^r$. *Let* $(\partial^n I)_{LS,q}(\boldsymbol{X\xi})$ *be the* q^{th} *order least squares approximation. Assume that* $q \leq \kappa + n$ *with* $q = N - n$. *The non-minimal* n^{th} *order algebraic derivative estimator* $\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; N)$ *(18) is given by:*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; N) = \partial^n I_{LS,q}(\bar{0}) + \varpi''(\boldsymbol{x}),$$

where $\varpi''(\boldsymbol{x})$ *is the noise contribution.*

We just showed that an N^{th} order ($N > n$) Taylor expansion lead to a q^{th} ($q = N - n$) order least squares approximation. We also showed, in equation (21), that the n^{th} order estimator based on an N^{th} order Taylor model correspond to an affine combination of minimal n^{th} order estimators where the combination weights $\boldsymbol{\lambda}_l$ are rational. By taking affine combinations where the $\boldsymbol{\lambda}_l$ are real numbers, it is possible to introduce a voluntary delay $\boldsymbol{\xi}_d \in [0,1]^r$. Now, if the delay correspond to one of the zeros of the $(q+\bar{1})^{th}$ order Jacobi polynomial, we achieve, from elementary minimal estimators, a $(q+\bar{1})^{th}$ order least squares approximation. We reduce thus the error induced by truncating the Taylor series expansion.

Proposition 5.3. *Let κ, μ, n, q, and l be multi-indices in $\mathbb{N}^r$. For any $\boldsymbol{\xi} \in [0,1]^r$, there exists a unique set of real numbers $\boldsymbol{\lambda}_l(\boldsymbol{\xi})$, $l = \bar{0}, \cdots, q$ depending on $\boldsymbol{\xi}$ such that*

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l(\boldsymbol{\xi}) \tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa_l, \mu_l; n) = \langle \boldsymbol{\mathcal{K}}_q(\boldsymbol{\xi}, \boldsymbol{x}), \partial^n I(X\boldsymbol{x}) \rangle. \tag{31}$$

Moreover, these coordinates must satisfy

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l(\boldsymbol{\xi}) = 1. \tag{32}$$

In addition, by excluding the trivial case where the $\boldsymbol{\lambda}_l$ are all equal to zero except one, the following hold:

$$\min_l \boldsymbol{\lambda}_l(\boldsymbol{\xi}) < 0. \tag{33}$$

Recall that the traditional approach consist in approximating the signal itself in a set of orthogonal polynomials. Then, the derivative is estimated by differenciating the approximating polynomial. This approach require the estimation of several parameters. Contrarily, our approach leads directly to an expression of the derivative in an orthogonal polynomials set. Thus the derivative estimation reduce to a single parameter identification. Moreover, we note that $\partial^n I(\boldsymbol{X}\boldsymbol{x})$ in equation (31) disappear upon integrating by parts n times as indicated by the following formula:

$$\langle \boldsymbol{\mathcal{K}}_q(\boldsymbol{\xi}, \boldsymbol{x}), \partial^n I(X\boldsymbol{x})\rangle = \langle \frac{\partial^n}{\partial \boldsymbol{x}^n}\boldsymbol{\mathcal{K}}_q(\xi, \boldsymbol{x}), I(\boldsymbol{X}\boldsymbol{x})\rangle.$$

6. Minimizing the noise influence

In this section we are interested by minimizing the noise influence. In fact our estimators can be written as linear time invariant filters with finite impulse response. Consider the minimal estimators formula (19). It can be written in the spacial domain as follows:

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}+\mathfrak{x};\kappa,\mu;n) = \frac{(-1)^n \gamma_{\kappa,\mu,n}}{\boldsymbol{X}^n}\int_{\bar{0}}^{\bar{1}} \partial^n \left[(\bar{1}-\boldsymbol{x})^{\mu+n}\boldsymbol{x}^{\kappa+n}\right] J(\mathfrak{x}+\boldsymbol{X}\boldsymbol{x})d\boldsymbol{x}, \quad (34)$$

where $\gamma_{\kappa,\mu,n} = \frac{(\mu+\kappa+2n+1)!}{(\mu+n)!(\kappa+n)!}$.

Let us denote by $h_{\kappa,\mu}(\boldsymbol{x})$ the following:

$$h_{\kappa,\mu}(\boldsymbol{x}) \doteq \frac{(-1)^n \gamma_{\kappa,\mu,n}}{\boldsymbol{X}^n}\left[\mathrm{H}(\boldsymbol{x}) - \mathrm{H}(\boldsymbol{x}-\bar{1})\right] \partial^n \left[(\bar{1}-\boldsymbol{x})^{\mu+n}\boldsymbol{x}^{\kappa+n}\right], \quad (35)$$

where H is the Heaviside function. We can write the estimators in the form of a linear filter as follows :

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}+\mathfrak{x};\kappa,\mu;n) = \int_{\bar{0}}^{\bar{\infty}} h_{\kappa,\mu}(\boldsymbol{x}) J(\mathfrak{x}+\boldsymbol{X}\boldsymbol{x})d\boldsymbol{x} \quad (36)$$

We consider in this section that the estimation hypercube of length $\boldsymbol{X}$ is small such that the mis-modeling error e_R is small. We suppose that the noise is a wide-sense stationary random process. Given multi-indices κ, μ, n,

q, and l, consider a linear combination of minimal estimators as follow :

$$\tilde{\tilde{I}}_{\boldsymbol{x}^n}(\bar{0}+\mathfrak{x};\kappa,\mu;N)=\int_{\bar{0}}^{\bar{\infty}}\sum_{l=\bar{0}}^{q}\iota_l h_{\kappa_l,\mu_l}(\boldsymbol{x})J(\mathfrak{x}+\boldsymbol{X}\boldsymbol{x})d\boldsymbol{x}. \tag{37}$$

where κ_l and μ_l are multi-indices defined earlier (in theorem 4.1) and $\iota_l \in \mathbb{R}$. Let us consider the set $\mathbf{L}' = \{l; l \leq q\}$. Equipped by the order defined in (25), $(\mathbf{L}', \prec)$ is a totaly ordered set. By arranging the elements of $\mathbf{L}'$ in an ascending order, we construct an index set for the $r \times (|q+\bar{1}|)$ elements ι_l in (37). We can construct thus the vector $\boldsymbol{\iota} = (\iota_{\bar{0}}, \cdots, \iota_l, \cdots, \iota_q) \in \mathbb{R}^{r\times|q+\bar{1}|}$.

With $\boldsymbol{\iota} = (\iota_{\bar{0}}, \cdots, \iota_q)^T \in \mathbb{R}^{r\times(|q+\bar{1}|)}$, c a multi-index and $\boldsymbol{x}'$ an independant variable taking values in $[0,1]^r$, we can then verify that the output noise variance is given by :

$$\begin{aligned} \text{VAR}(e) &= \sum_{l,c}^{q}\iota_l\iota_c\int_{\bar{0}}^{\bar{1}}\int_{\bar{0}}^{\bar{1}}h_{\kappa_l,\mu_l}(\boldsymbol{x})\mathbb{E}[\varpi(\boldsymbol{x})\varpi(\boldsymbol{x}')]h_{\kappa_c,\mu_c}(\boldsymbol{x}')d\boldsymbol{x}d\boldsymbol{x}' \quad (38)\\ &= \boldsymbol{\iota}^T\mathcal{R}\boldsymbol{\iota}, \quad (39)\end{aligned}$$

where $\mathcal{R}$ is a square symmetric matrix of $[r\times(|q+\bar{1}|)]$ lines and $[r\times(|q+\bar{1}|)]$ columns with entries given by :

$$\mathcal{R}_{l,c}=\int_{\bar{0}}^{\bar{1}}\int_{\bar{0}}^{\bar{1}}h_{\kappa_l,\mu_l}(\boldsymbol{x})\mathbb{E}[\varpi(\boldsymbol{x})\varpi(\boldsymbol{x}')]h_{\kappa_c,\mu_c}(\boldsymbol{x}')d\boldsymbol{x}d\boldsymbol{x}' \tag{40}$$

Let $\boldsymbol{z}$ be a $[r\times(|q+\bar{1}|)]-$dimensional vector with one in each entry. The barycentric coordinates $\boldsymbol{\iota}_{\text{MIN}}$ minimizing the variance are given by :

$$\boldsymbol{\iota}_{\text{MIN}}=\frac{\mathcal{R}^{-1}\boldsymbol{z}}{\boldsymbol{z}^t\mathcal{R}^{-1}\boldsymbol{z}}. \tag{41}$$

Finally, if $\varpi(\boldsymbol{x})$ is a white noise then the barycentric coordinates are given by :

$$\boldsymbol{\iota}_{\text{MIN}}=\frac{1}{|q+\bar{1}|}\boldsymbol{z}, \tag{42}$$

showing that the minimum output mean square error is achieved by the centroïd of the points $\tilde{I}_{\boldsymbol{x}^n}(\mathfrak{x};\kappa_l,\mu_l,n),\ l=\bar{0},\cdots,q$.

Remark 6.1. *Formulas similar to (42) can be computed for other types of noises (bandlimited white, pink, brownian...). It suffices to have the auto-correlation function.*

We arrive to a remarkable conclusion : The affine combination of minimal estimators (31) minimizing the truncation error do not in general coincide with the combination (41) minimizing the noise variance. Nevertheless, if $\kappa = \mu$ and $q = \bar{1}$, minimal estimator (which admit a least squares interpretation, see proposition 5.1) coincide with the one minimizing the output white noise variance.

Proposition 6.1. *Consider a white noise and let h^{MV} denote the filter minimizing the output noise variance by :*

$$h^{MV}(\boldsymbol{x},\kappa,\mu,n,q) = \frac{1}{|q+\bar{1}|}\sum_{l=\bar{0}}^{q} h_{\kappa_l,\mu_l}(\boldsymbol{x}) \tag{43}$$

If $q=\bar{1}$ and $\mu=\kappa$ we have that

$$h^{MV}(\boldsymbol{x},\kappa,\mu,n,\bar{1}) = h_{\kappa,\mu}(\boldsymbol{x}), \forall n. \tag{44}$$

where $h_{\kappa,\mu}(\boldsymbol{x})$ defined in (35).

It is known that classical least squares should be regularized in order to increase their robustness to noise. The choice of the regularization terms and the tuning of their parameters is not an easy task. We showed that the filter in (44) admit a least squares interpretation (proposition 5.1) and at the same time minimize the output noise variance.

More generally, we can verify that any minimal algebraic estimator for which $abs(\kappa-\mu)$ (where abs denote the absolute value) is close to the corresponding minimum variance estimator for $q=\bar{1}$. This stems from the identity :

$$h^{MV}(\boldsymbol{x},\kappa,\mu,n,\bar{1}) = \frac{\kappa+\mu+2n+\bar{2}}{2(\mu+n+\bar{1})} h_{\kappa,\mu}(\boldsymbol{x}) + \frac{\mu-\kappa}{2(\kappa+\mu+2n+\bar{2})} h_{\kappa+\bar{1},\mu}(\boldsymbol{x}).$$

Note finally that this result is not valid for $q>\bar{1}$, it suffices to remark that in the monovariate case the barycentric coordinate of (45) are not identical ($\neq \frac{1}{3}$):

$$\begin{aligned} I_{x_m^{n_m}}(0;\kappa_m,\mu_m;n_m) &= \frac{1}{4} I_{x_m^{n_m}}(0;\kappa_m+2,\mu_m;n_m) + \frac{1}{2} I_{x_m^{n_m}}(0;\kappa_m+1,\mu_m+1;n_m) \\ &+ \frac{1}{4} I_{x_m^{n_m}}(0;\kappa_m,\mu_m+2;n_m). \end{aligned} \tag{45}$$

7. Numerical simulations

7.1. Implementation issues

The general form of the algebraic estimators can be written as:

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N) = \int_{\bar{0}}^{\bar{1}} G(\boldsymbol{x}) J(\boldsymbol{X}\boldsymbol{x}) d\boldsymbol{x}. \tag{46}$$

In order to estimate derivatives at a point different from zero, a translation is needed as follows:

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}+\mathfrak{x};\kappa,\mu;N) = \int_{\bar{0}}^{\bar{1}} G(\boldsymbol{x}) J(\boldsymbol{X}\boldsymbol{x}+\mathfrak{x}) d\boldsymbol{x}. \tag{47}$$

A discretization, with a sampling $\boldsymbol{\mathfrak{s}} = (\mathfrak{s}_1, \cdots, \mathfrak{s}_r)$, of the noisy signal $J(\mathfrak{x} + \boldsymbol{X}\boldsymbol{x})$ on the hypercube $[\mathfrak{x}, \mathfrak{x} + \boldsymbol{X}\boldsymbol{x}]^r$, $\boldsymbol{x} \in [0,1]^r$, leads to a hyper-matrix J_d of dimension $\frac{1}{\mathfrak{s}_1} \times \cdots \times \frac{1}{\mathfrak{s}_r}$. Discretize the interval $[0,1]^r$ with the same number of samples and evaluate $G(\boldsymbol{x})$ on the samples leads to a hyper-matrix G_d of the same dimension $\frac{1}{\mathfrak{s}_1} \times \cdots \times \frac{1}{\mathfrak{s}_r}$. Let W_d be a hyper-matrix constituted by the weights of a numerical integration scheme. Let R_d be the hyper-matrix obtained by element-wise multiplication of G_d and W_d. Thus a numerical estimation of (47) is given by:

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0} + \mathfrak{x}; \kappa, \mu; N) = \sum R_d \times J_d, \tag{48}$$

where $\times$ and $\sum$ in (48) denote respectively element-wise multiplication and summation. In the subsequent simulations, Simpson rule for multiple integration is used in W_d [23].

7.2. Numerical simulations

Several first and second order derivative estimators are tested using a noisy bidimensional signal:

$$J(x_1, x_2) = \sin(\frac{1}{2}x_1^2 + \frac{1}{4}x_2^2 + 3)\cos(2x_1 + 1 - e^{x_2}) + \varpi(x_1, x_2). \tag{49}$$

A noise level of $SNR = 25\ dB$ is considered by using the formula

$$SNR = 10\log_{10}\left(\frac{\sum_{i,j}|I(x_{1i}, x_{2j})|^2}{\sum_{i,j}|\varpi(x_{1i}, x_{2j})^2|}\right).$$

A sketch of (49) is shown in figure 1, while figure 2 show a slice of the noisy surface were the derivatives are computed. It is given by $\mathcal{E} = \{(x_1, x_2); x_2 = 0, -1 \leq x_1 \leq 3\}$. In fact, at discrete equidistant points of $\mathcal{E}$ the derivative

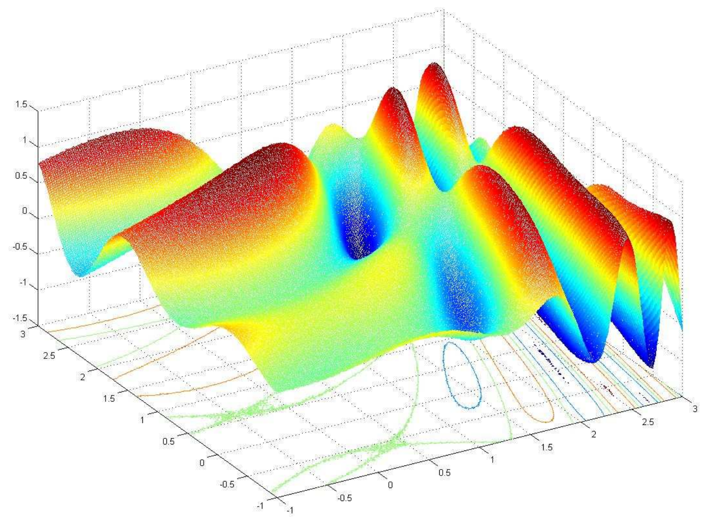

Figure 1: 3-D plot of $I(x_1, x_2) = sin(\frac{1}{2}x_1^2 + \frac{1}{4}x_2^2 + 3)cos(2x_1 + 1 - e^{x_2})$

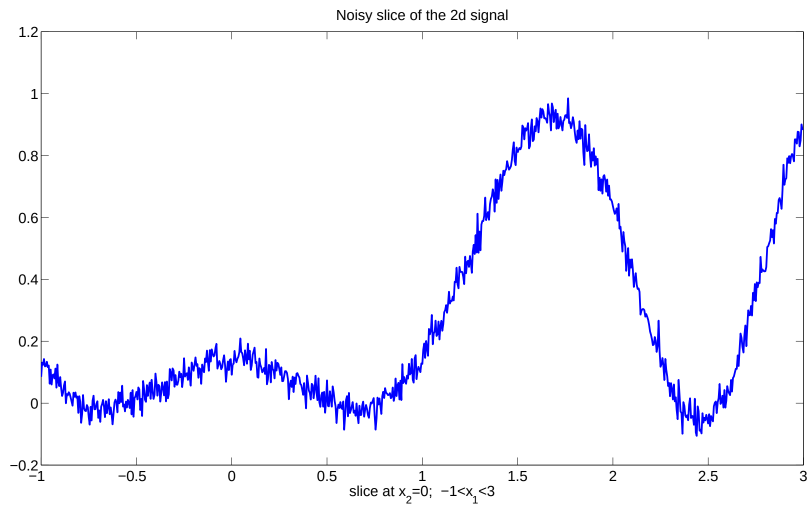

Figure 2: A slice of the noisy surface at $x_2 = 0$ and $-1 < x_1 < 3$, 25 dB

is estimated using a small elementary surface at the point of interest. A sampling step of (0.005×0.005) is used.

A comparison is made with finite differences from [1]. It is important to note that finite differences are not evaluated on the sampling step (0.005×0.005) but on the same elementary surface used to evaluate the algebraic estimators. Using a large surface for finite differences permit to filter the noise.

Three derivatives are evaluated I_{x_1}, $I_{x_1^2}$ and $I_{x_1 x_2}$. For each one four estimators are compared:

1. A minimal estimator from equation (19) *i.e. an estimator based on minimal Taylor series expansion. We use* $\kappa = \mu = \bar{0}$.
2. *A non minimal estimator (18) with* $\kappa = \mu = \bar{0}$ *and* $N = n + \bar{1}$.
3. *An affine combination of minimal estimators (31) with* $\kappa = \mu = \bar{0}$ *where* $\boldsymbol{\xi}$ *is chosen to accomplish an exact estimator for polynomial signals of degree* $n + \bar{1}$. *This class of estimators will be called in the sequel voluntary delayed estimators.*
4. *A finite difference estimator from [1].*

We used the same noise realization as well as the same elementary surface to estimate the derivatives from the four estimators listed above. Some facts, predicted in the theoretical part, can be seen in the simulations especially that:

- the minimal estimators produce delayed estimations,
- the non minimal estimators do not induce a delay,

- the voluntary delayed estimators procure a better representation of the derivative (minimize the truncation error). Because they are exact for polynomial signals of degree $n+\bar{1}$ although based on a Taylor expansion to the order n.

However, we can report some observations which are not studied in the theoretical part and will be investigated in future works:

- non minimal estimators provide poor robustness with respect to noises when compared with minimal ones,
- both minimal and non minimal estimators deform the extremas of the derivative,
- the voluntary delayed estimators are good compromise between robustness to noise and minimization of the truncation error.

The simulations are detailed below.

7.2.1. Estimation of I_{x_1}

The minimal estimator is computed by taking $n = (1,0)$, $N = n$, $\mu = (0,0)$ and $\kappa = (0,0)$. The non minimal one is computed by taking $n = (1,0)$, $N = (2,0)$, $\mu = (0,0)$ and $\kappa = (0,0)$. The voluntary delayed estimator is synthesized using equation (31) to calculate $\boldsymbol{\lambda}_l(\boldsymbol{\xi})$ which gives $\boldsymbol{\lambda}_{(0,0)}(\boldsymbol{\xi}) = -2 + 5\boldsymbol{\xi}$, $\boldsymbol{\lambda}_{(1,0)}(\boldsymbol{\xi}) = 3 - 5\boldsymbol{\xi}$. The value of $\boldsymbol{\xi}$ is found by equating to zero $\mathcal{P}_{(1,0)}^{\{\kappa,\mu\}}(\boldsymbol{\xi}) = 0$ yielding to $\boldsymbol{\xi} = \frac{1}{2} - \frac{1}{2\sqrt{5}}$.

Simulation results are shown in figure 3 a sliding surface consisting of 70×70 elements is used. Amplitude deformation induced by both the minimal and non-minimal estimators is visible on the figure 4 where the signals (of figure 3) are aligned. Notice that the voluntary delayed estimator produce better representation of the derivative.

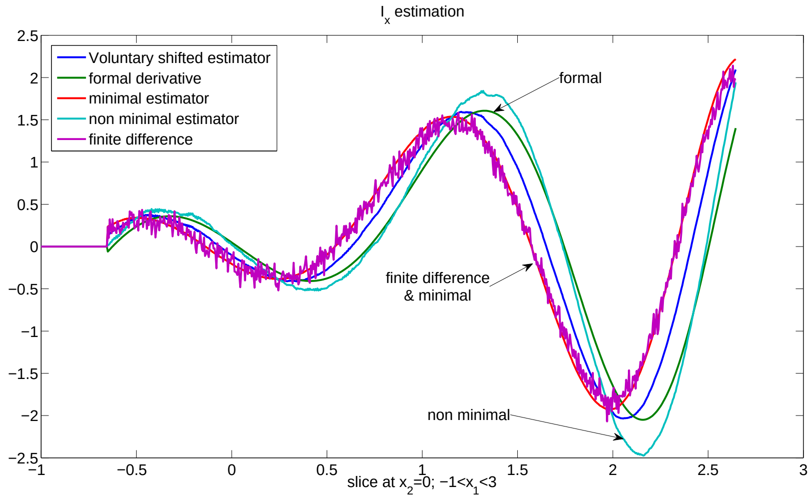

Figure 3: estimation of I_{x_1}

7.2.2. Estimation of $I_{x_1^2}$

The minimal estimator is computed by taking $n = (2, 0)$, $N = n$, $\mu = (0, 0)$ and $\kappa = (0, 0)$. The voluntary delayed estimator is synthesized using equation (31). The corresponding coefficients $\boldsymbol{\lambda}_l(\xi)$, $l \in \{(0, 0), (1, 0)\}$, are given by $\boldsymbol{\lambda}_{(0,0)}(\boldsymbol{\xi}) = -3 + 7\boldsymbol{\xi}$, $\boldsymbol{\lambda}_{(1,0)}(\boldsymbol{\xi}) = 4 - 7\boldsymbol{\xi}$. The value of $\boldsymbol{\xi}$ is the solution of $\mathcal{P}_{(2,0)}^{\{\kappa,\mu\}}(\boldsymbol{\xi}) = 0$ yielding $\boldsymbol{\xi} = \frac{1}{2} - \frac{1}{2\sqrt{7}}$.

Simulation results are shown in figure 5, a sliding surface consisting of 100×100 elements is used. The amplitude deformation induced by the minimal estimator can be seen on the figure 6 where the signal (of figure 5) are aligned.

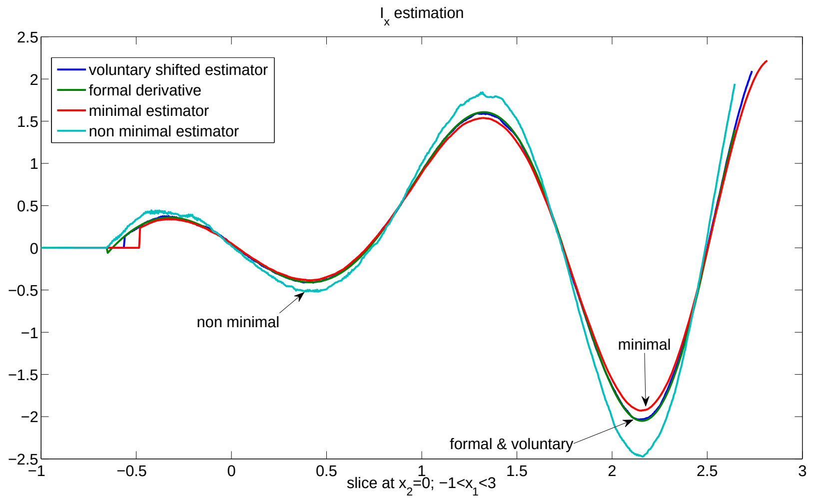

Figure 4: Alignment of the signals shown in figure 3.

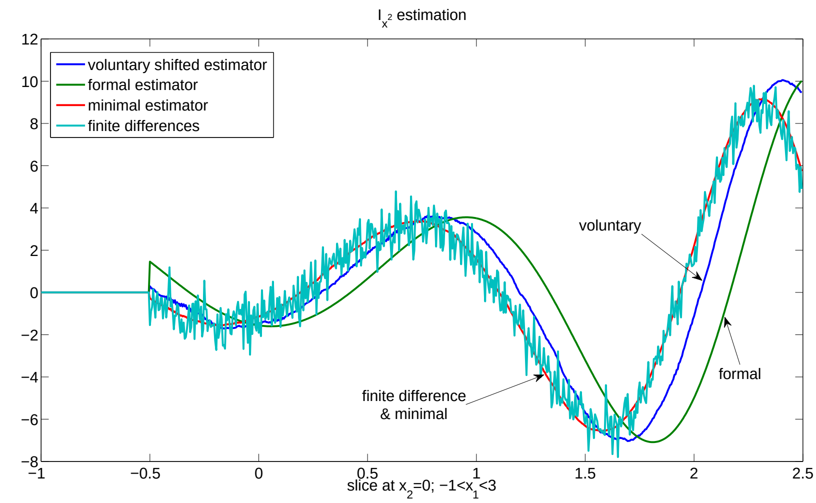

Figure 5: estimation of $I_{x_1^2}$

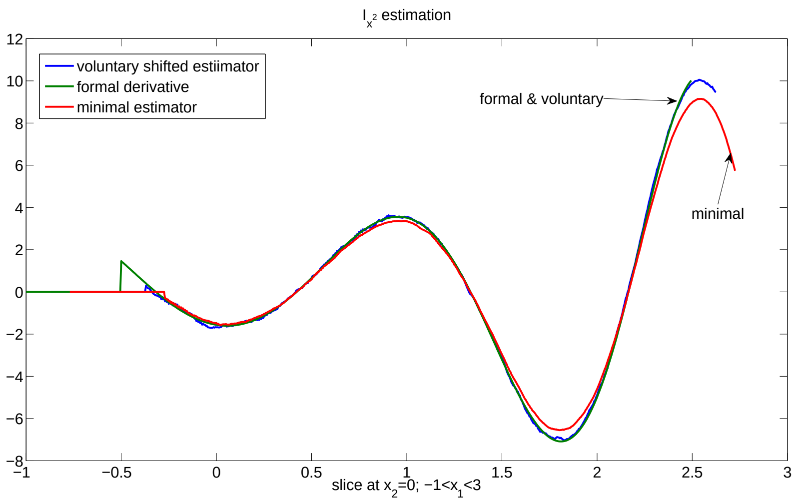

Figure 6: Alignment of the signals depicted in figure 5

7.2.3. Estimation of $I_{x_1x_2}$

The minimal estimator is computed by taking $n = (1,1)$, $N = n$, $\mu = (0,0)$ and $\kappa = (0,0)$. The non minimal one is computed by taking $n = (1,1)$, $N = (2,2)$, $\mu = (0,0)$ and $\kappa = (0,0)$. The voluntary delayed estimator is synthesized using equation (31) to calculate $\boldsymbol{\lambda}_l(\xi)$ with $l \in \{(0,0),(0,1),(1,0),(1,1)\}$. We obtain $\boldsymbol{\lambda}_{(0,0)}(\boldsymbol{\xi}) = (-2+5\xi_1)(-2+5\xi_2)$, $\boldsymbol{\lambda}_{(1,0)}(\boldsymbol{\xi}) = (-2+5\xi_1)(3-5\xi_2)$, $\boldsymbol{\lambda}_{(0,1)}(\boldsymbol{\xi}) = (3-5\xi_1)(-2+5\xi_2)$, $\boldsymbol{\lambda}_{(1,1)}(\boldsymbol{\xi}) = (3-5\xi_1)(3-5\xi_2)$. The values of the delays $\boldsymbol{\xi} = (\xi_1, \xi_2)$ are given by the solutions of $\mathcal{P}^{\{\kappa,\mu\}}_{(1,0)}(\boldsymbol{\xi}) = 0$ and $\mathcal{P}^{\{\kappa,\mu\}}_{(0,1)}(\boldsymbol{\xi}) = 0$. Due to the symmetry of the cross derivative $I_{x_1x_2}$ the delays are equal $\xi_1 = \xi_2 = \frac{1}{2} - \frac{1}{2\sqrt{5}}$. This is not the case if one wants to estimate $I_{x_1^2x_2}$ for example, the delays ξ_1 and ξ_2 are different for this case.

Simulation results are shown in figure 7, a sliding surface consisting of 100×100 elements is used. The curves in the figure 7 are aligned and displayed in figure 8.

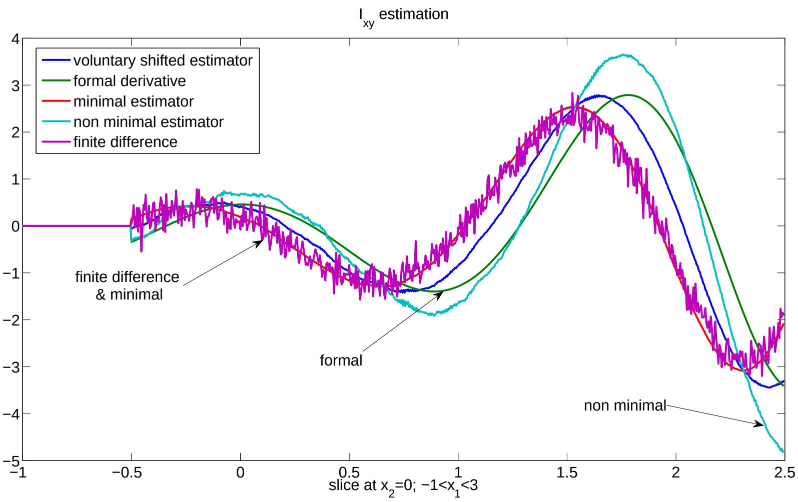

Figure 7: estimation of $I_{x_1 x_2}$

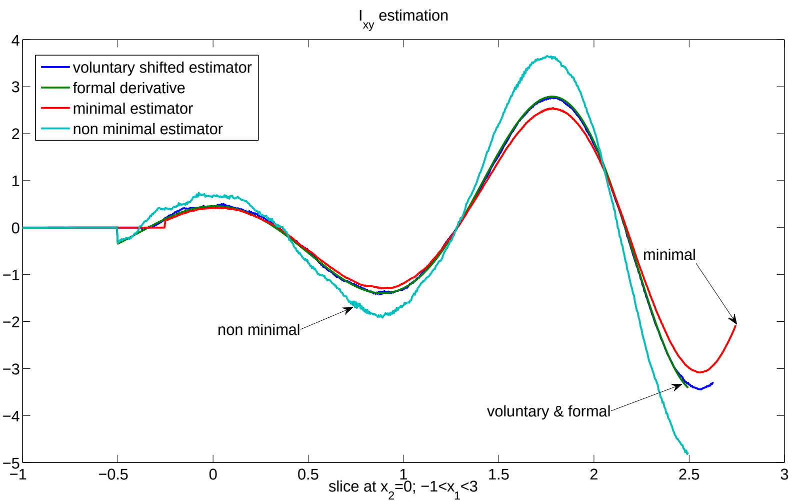

Figure 8: Alignment of the signals shown in figure 7

7.2.4. Minimum variance vs minimum e_R estimators

The same non minimal voluntary delayed estimator of $I_{x_1^2}$ ($q = (1,0)$, $\lambda_{(0,0)}(\boldsymbol{\xi}) = -3 + 7\boldsymbol{\xi}$, $\lambda_{(1,0)}(\boldsymbol{\xi}) = 4 - 7\boldsymbol{\xi}$, $\boldsymbol{\xi} = \frac{1}{2} - \frac{1}{2\sqrt{7}}$) is compared with one minimizing the noise variance *i.e.* $\iota_l = \frac{1}{2}$ (which coincide with the minimal estimator, $n = (2,0)$, $N = n$, $\mu = (0,0)$, $\kappa = (0,0)$). A sliding surface whose size (60×60 samples) is smaller than the one used previously (100×100) in the estimation of $I_{x_1^2}$ is used for the computations. Recall that reducing the sliding surface reduces the truncation error and accentuate the noise influence. Simulation results are shown in figure 9. They clearly show that the minimum variance estimator produce a smoother estimation.

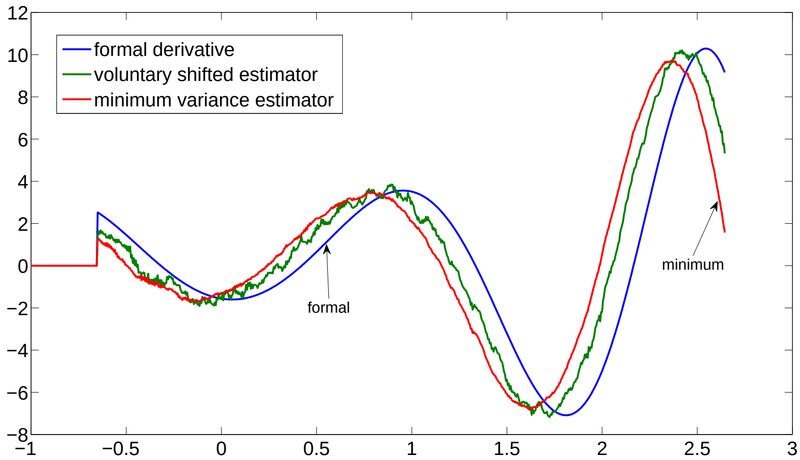

Figure 9: minimum variance and non minimal voluntary delayed estimators of $I_{x_1^2}$

7.2.5. Signal with varying frequency

Finally, another slice of the signal is considered at $x_1 = 2$ and $-1 < x_2 < 3$. The particularity of this slice is that its (pseudo) frequency increases with

x_2. The slice is shown in figure 10.

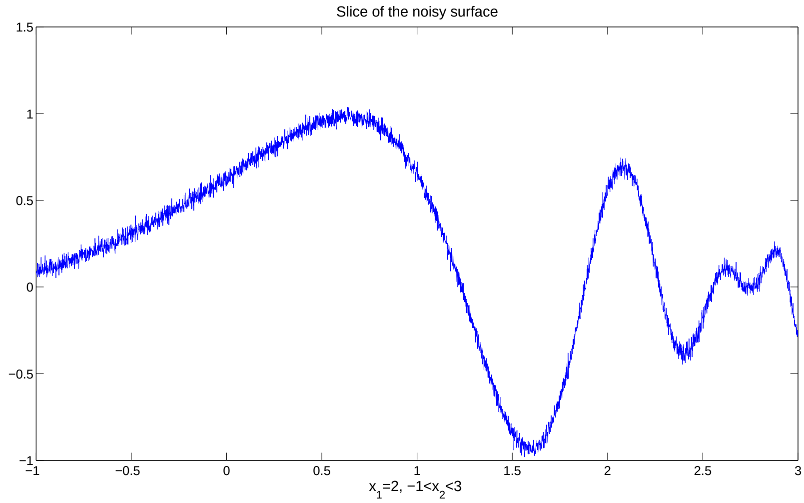

Figure 10: A slice of the noisy surface at $x_1 = 2$ and $-1 < x_2 < 3$, 25 dB

The estimation of $I_{x_1x_2}$ is shown in 11. Note that the result of the minimal estimator ($n = (1, 1)$, $N = n$, $\mu = (0, 0)$ and $\kappa = (0, 0)$) degrades when the frequency increase. If one decreases the sliding surface size, better results are obtained with the minimal estimator at higher frequencies but degrades at low frequencies. On the contrary, very good estimations are obtained at both high and low frequencies with the voluntary shifted estimator ($\boldsymbol{\lambda}_{(0,0)}(\boldsymbol{\xi}) = (-2+5\xi_1)(-2+5\xi_2)$, $\boldsymbol{\lambda}_{(1,0)}(\boldsymbol{\xi}) = (-2+5\xi_1)(3-5\xi_2)$, $\boldsymbol{\lambda}_{(0,1)}(\boldsymbol{\xi}) = (3-5\xi_1)(-2+5\xi_2)$, $\boldsymbol{\lambda}_{(1,1)}(\boldsymbol{\xi}) = (3-5\xi_1)(3-5\xi_2)$). The results of both estimators can be better seen on the figure 12 where the curves of figure 11 are aligned by adjusting the delays.

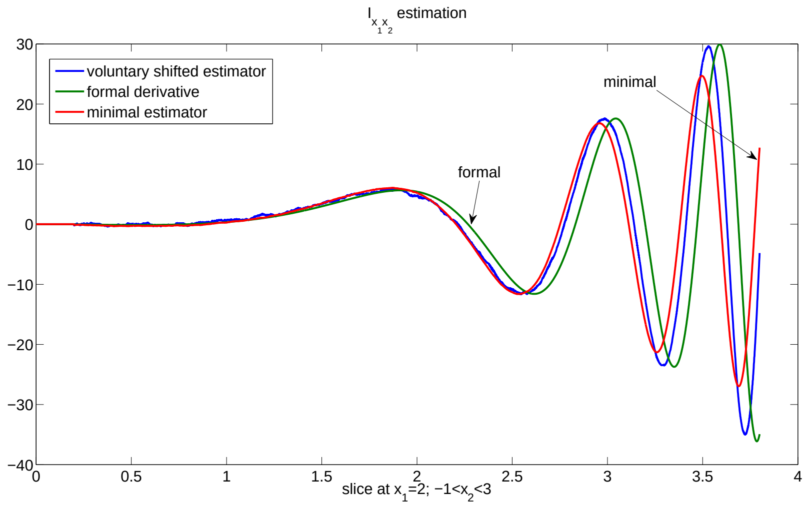

Figure 11: estimation of $I_{x_1x_2}$

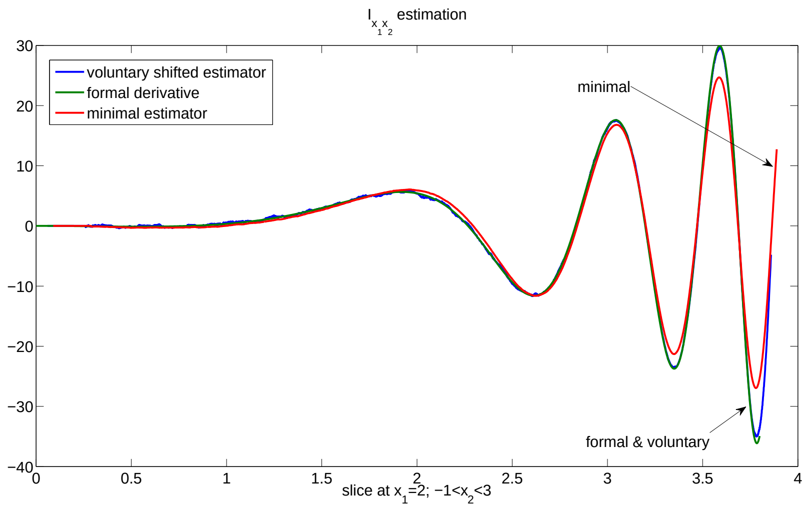

Figure 12: Alignment of the signals shown in figure 11

8. CONCLUSION

In this paper, we presented a partial derivatives estimation method for multidimensional signals. On a small interval the signal is represented by a truncated Taylor expansion. Then the application of multivariate Laplace transform together with adequate algebraic manipulations enabled us to express the desired partial derivative of any order as a function of iterated integrals of the noisy signal. Several recurrence relations and structural properties were provided. An interpretation of the estimators as least square minimization is also done by expressing the estimators in an orthogonal basis constituted by Jacobi polynomials. This projection enabled us not only to show a spacial shifting inherent to a specific class of estimators but also to synthesize a new class of estimators minimizing the truncation remainder of the Taylor local model. We provided also another class of estimators minimizing the noise influence. Finally we provided a numerical implementation scheme in the form a finite impulse digital filters. Our estimators are very performant in practical applications especially in image and video processing, our first results in edge detection in images and motion detection in image sequences are conclusive. They will be published in future papers.

9. proofs

proof 9.1 (of proposition 3.1). *Multiply (15) by* $s^{N+\bar{1}}$ *to obtain*

$$s^{N+\bar{1}}\hat{I}(s) = \sum_{\alpha=\bar{0}}^{N} s^{N-\alpha} I_{\boldsymbol{x}^{\alpha}}(\bar{0}).$$

For $\alpha > n$*, we have* $N - n > N - \alpha$*. Consequently,* $\partial^{N-n} s^{N-\alpha} = 0$ *and* $\partial^{N-n} s^{N-n} = (N-n)!$*. This means that* $\partial^{N-n} s^{N-\alpha}$ *annihilates all the coeffi-*

cients of $I_{\boldsymbol{x}^\alpha}(\bar{0})$ with $\alpha > n$ in the Taylor expansion (15). To isolate $I_{\boldsymbol{x}^n}(\bar{0})$, it remains to annihilate the terms with $\alpha < n$. It can be verified that applying $\partial^n \frac{1}{s}$ does the job. By further (partial) differentiating κ times followed by a multiplication by $\frac{1}{\boldsymbol{s}^{N+\mu+1}}$ we have the relation (17).

proof 9.2 (of theorem 4.1). *Set $p = \kappa + n$ and write (16) in the form*

$$\Pi_{\kappa,\mu}^{N,n} = \frac{1}{\boldsymbol{s}^{N+\mu+\bar{1}}}\partial^p \frac{1}{\boldsymbol{s}}\partial^q \boldsymbol{s}^{q+\bar{1}} s^n.$$

It can be rewritten in the form

$$\begin{aligned}\Pi_{\kappa,\mu}^{N,n} &= \frac{1}{s_1^{N_1+\mu_1+1}}\frac{\partial^{p_1}}{\partial s_1^{p_1}}\frac{1}{s_1}\frac{\partial^{q_1}}{\partial s_1^{q_1}} s_1^{q_1+1} s_1^{n_1} \times \cdots \times \frac{1}{s_m^{N_m+\mu_m+1}}\frac{\partial^{p_m}}{\partial s_m^{p_m}}\frac{1}{s_m}\frac{\partial^{q_m}}{\partial s_m^{q_m}} s_m^{q_m+1} s_m^{n_m} \\ &\times \cdots \times \frac{1}{s_r^{N_r+\mu_r+1}}\frac{\partial^{p_r}}{\partial s_r^{p_r}}\frac{1}{s_r}\frac{\partial^{q_r}}{\partial s_r^{q_r}} s_r^{q_r+1} s_r^{n_r}\end{aligned}$$

Let $\mathcal{W}_m = \frac{1}{s_m^{N_m+\mu_m+1}}\frac{\partial^{p_m}}{\partial s_m^{p_m}}\frac{1}{s_m}\frac{\partial^{q_m}}{\partial s_m^{q_m}} s_m^{q_m+1} s_m^{n_m}$ and following the lines of [31], we start first by evaluating $\frac{\partial^{q_m}}{\partial s_m^{q_m}}(s_m^{q_m+1} s_m^{n_m})$ then we evaluate $\frac{\partial^{p_m}}{\partial s_m^{p_m}}\left[\frac{1}{s_m}\frac{\partial^{q_m}}{\partial s_m^{q_m}}(s_m^{q_m+1} s_m^{n_m})\right]$:

$$\begin{aligned}\mathcal{W}_m &= \sum_{i=0}^{q_m}\sum_{j=0}^{\min(p_m,q_m-i)} \begin{pmatrix} q_m \\ i \end{pmatrix}\begin{pmatrix} p_m \\ j \end{pmatrix} \times \\ &\times \frac{(q_m+1)!}{(q_m+1-i)(q_m-i-j)!}\frac{1}{s_m^{\mu_m+n_m+1+i+j}}\partial^{p_m+q_m-i-j} s_m^{n_m}. \quad (50)\end{aligned}$$

Let $l_m = i + j$, the above expression yields

$$\begin{aligned}\mathcal{W}_m &= \sum_{i=0}^{q_m}\sum_{l_m=i}^{\min(p_m+i,q_m)} \begin{pmatrix} q_m \\ i \end{pmatrix}\begin{pmatrix} p_m \\ l_m - i \end{pmatrix} \times \\ &\times \frac{(q_m+1)!}{(q_m+1-i)(q_m-l_m)!}\frac{1}{s_m^{\mu_m+n_m+1+l_m}}\partial^{p_m+q_m-l_m} s_m^{n_m}.\end{aligned}$$

By permuting the sums, one gets

$$\begin{aligned}\mathcal{W}_m &= \sum_{l_m=0}^{q_m}\sum_{i=\max(0,l_m-p_m)}^{l_m}\begin{pmatrix} q_m \\ i \end{pmatrix}\begin{pmatrix} p_m \\ l_m-i \end{pmatrix}\times \\ &\times \frac{(q_m+1)!}{(q_m+1-i)(q_m-l_m)!}\frac{1}{s_m^{\mu_m+n_m+1+l_m}}\partial^{p_m+q_m-l_m}s_m^{n_m}.\end{aligned}$$

Now let $\lambda'^{m}_{l_m}$ *(where the subscript* l_m *and the superscript* m *are integers) be given by :*

$$\lambda'^{m}_{l_m} = \sum_{i=\max(0,l_m-p_m)}^{l_m}\begin{pmatrix} q_m \\ i \end{pmatrix}\begin{pmatrix} p_m \\ l_m-i \end{pmatrix}\frac{(q_m+1)!}{(q_m+1-i)(q_m-l_m)!}.$$

Then $\mathcal{W}_m$ *is given by :*

$$\mathcal{W}_m = \sum_{l_m=0}^{q_m}\lambda'^{m}_{l_m}\frac{1}{s_m^{\mu_m+n_m+1+l_m}}\partial^{p_m+q_m-l_m}s_m^{n_m}.$$

Let $\lambda^m_{l_m} = (-1)^{q_m-l_m}\frac{(n_m+\kappa_m+q_m-l_m)!}{p_m!q_m!}\lambda'^{m}_{l_m}$ *and note that* $\lambda^m_{l_m}$ *are rational numbers. By using Vandermonde identity, we obtain for* $q_m \leq p_m$

$$\begin{aligned}\lambda^m_{l_m} &= (-1)^{q_m-l_m}\begin{pmatrix} p_m+q_m-l_m \\ p_m \end{pmatrix}\sum_{i=0}^{l_m}\begin{pmatrix} p_m \\ l_m-i \end{pmatrix}\begin{pmatrix} q_m+1 \\ i \end{pmatrix} \qquad (51)\\ &= (-1)^{q_m-l_m}\begin{pmatrix} p_m+q_m-l_m \\ p_m \end{pmatrix}\begin{pmatrix} p_m+q_m+1 \\ l_m \end{pmatrix}.\end{aligned}$$

Finally define $C(q_m) = \sum_{l_m=0}^{q_m}\lambda^m_{l_m}$*, and using the identity* $\begin{pmatrix} a+1 \\ b \end{pmatrix} =$

$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a \\ b-1 \end{pmatrix}$ *we have the following recurrence relation (see [31])*

$$\begin{aligned} C(q_m+1) &= C(q_m) + \sum_{l_m=0}^{q_m+1} (-1)^{l_m} \begin{pmatrix} p_m+l_m \\ l_m \end{pmatrix} \begin{pmatrix} p_m+q_m+1 \\ p_m+l_m \end{pmatrix} \\ &= C(q_m) + \begin{pmatrix} p_m+q_m+1 \\ p_m \end{pmatrix} \sum_{l_m=0}^{q_m+1} (-1)^{l_m} \begin{pmatrix} q_m+1 \\ l_m \end{pmatrix} = C(q_m). \end{aligned}$$

If $q_m = 1$ *we have that* $C(1) = 1$*, and by using the above relation we have* $C(q_m) = \sum_{l_m=0}^{q_m} \lambda^m_{l_m} = 1, \forall q_m$.

Back to $\Pi^{N,n}_{\kappa,\mu}$*, it can now be rewritten as:*

$$\Pi^{N,n}_{\kappa,\mu} = \prod_{m=0}^{r} \mathcal{W}_m = \prod_{m=0}^{r} \sum_{l_m=0}^{q_m} \lambda'^m_{l_m} \frac{1}{s_m^{\mu_m+n_m+1+l_m}} \partial^{p_m+q_m-l_m} s_m^{n_m}$$

Now let $u_1 = (\lambda'^1_0 \cdots, \lambda'^1_{l_1}, \cdots, \lambda'^1_{q_1}), \cdots,$ $u_m = (\lambda'^2_0 \cdots, \lambda'^2_{l_m}, \cdots, \lambda'^2_{q_m}), \cdots,$ $u_r = (\lambda'^r_0 \cdots, \lambda'^r_{l_r}, \cdots, \lambda'^r_{q_r})$*, and* $\lambda' = u_1 \otimes \cdots \otimes u_m \otimes \cdots \otimes u_r$ *(all elements belong to* $\mathbb{Q}$*).*

On the other hand, let $\mathbf{L}' = \{l; l \leq q\}$ *be a collection of multi-indices. Define on* $\mathbf{L}'$ *the order relation* $\prec$ *by:*

$$l \prec l' \quad \text{IFF} \quad \sum_{i=1}^{r} 2^{r+1-i} l(i) < \sum_{j=1}^{r} 2^{r+1-j} l'(j).$$

It is clear that the set $\mathbf{L}'$ *equipped with* $\prec$*is a totaly ordered set. Its elements can be thus arranged in an ascending order. Since* $\dim(\boldsymbol{\lambda}') = card(\mathbf{L}')$*, the set* $\mathbf{L}'$ *can be used as an index set for* $\boldsymbol{\lambda}'$ *i.e* $\boldsymbol{\lambda}'_l$ *is the* l^{th} *element of* $\boldsymbol{\lambda}'$*.*

Thus $\Pi^{N,n}_{\kappa,\mu}$ *can be written as :*

$$\Pi^{N,n}_{\kappa,\mu} = \sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}'_l \frac{1}{\boldsymbol{s}^{\mu+n+l+\bar{1}}} \partial^{p+q-l} \boldsymbol{s}^n.$$

Using (17), we have $\boldsymbol{\lambda}_l = (-1)^{q-l}\frac{(n+\kappa+q-l)!}{p!q!}\boldsymbol{\lambda}'_l$.

Consider now the following sum

$$\sum_{l_1=0}^{q_1}\lambda^1_{l_1}\times\cdots\times\sum_{l_m=0}^{q_m}\lambda^m_{l_m}\times\cdots\times\sum_{l_r=0}^{q_r}\lambda^r_{l_r}=\sum_{l=\bar{0}}^{q}\boldsymbol{\lambda}_l$$

which from the monovariate case [31] we have that $\sum_{l_m=0}^{q_m}\lambda^m_{l_m}=1$ *thus :*

$$\sum_{l=\bar{0}}^{q}\boldsymbol{\lambda}_l = 1.$$

Finally, based on the monovariate case [31], it is straightforward to see that:

$$\min_l \boldsymbol{\lambda}_l < 0. \tag{52}$$

proof 9.3 (of theorem 4.2). *Set* $q = N - n$ *and* $\nu = N + \mu + \bar{1}$, *equation (17) can be rewritten as :*

$$\frac{(-1)^{n+\kappa}(n+\kappa)!q!}{s^{\mu+\kappa+N+n+\bar{2}}}I_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N)=\frac{1}{s^{\nu}}\partial^{n+\kappa}\frac{1}{s}\partial^q s^{N+\bar{1}}\hat{I}(\boldsymbol{s}).$$

It can be written in the spacial domain as

$$\frac{(-1)^{n+\kappa}(n+\kappa)!q!\boldsymbol{X}^{\mu+\kappa+N+n+\bar{1}}}{(\mu+\kappa+N+n+\bar{1})!}I_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;N)=$$
$$\int_{\bar{0}}^{\bar{1}}(-1)^{n+\kappa}\frac{(\bar{1}-\boldsymbol{x})^{\nu-\bar{1}}}{(\nu-\bar{1})!}\boldsymbol{x}^{n+\kappa}\int_{\bar{0}}^{\boldsymbol{x}}(-1)^q\xi^q\partial^{N+\bar{1}}I(\boldsymbol{X}\xi)d\xi d\boldsymbol{x} \tag{53}$$

$$=\int_{\bar{0}}^{\bar{1}}\Bigg\{(-1)^{n+\kappa-(n_m+\kappa_m)1_m}(-1)^{n_m+\kappa_m}\frac{(\bar{1}-\boldsymbol{x})^{\nu-\bar{1}-(\nu_m-1)1_m}}{(\nu-\bar{1}-(\nu_m-1)1_m)!}$$
$$\frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!}\boldsymbol{x}^{n+\kappa-(n_m+\kappa_m)1_m}x_m^{n_m+\kappa_m}$$
$$\int_{\bar{0}}^{\boldsymbol{x}}(-1)^{q-q_m1_m}(-1)^{q_m}\xi^{q-q_m1_m}\xi_m^{q_m}\partial^{N+\bar{1}}I(\boldsymbol{X}\xi)d\xi\Bigg\}d\boldsymbol{x}$$

$$=\int_{\bar{0}}^{\bar{1}-1_m}\Bigg\{(-1)^{n+\kappa-(n_m+\kappa_m)1_m}\frac{(\bar{1}-\boldsymbol{x})^{\nu-\bar{1}-(\nu_m-1)1_m}}{(\nu-\bar{1}-(\nu_m-1)1_m)!}\boldsymbol{x}^{n+\kappa-(n_m+\kappa_m)1_m}$$
$$\int_{\bar{0}}^{\boldsymbol{x}-x_m}\Big[(-1)^{q-q_m1_m}\xi^{q-q_m1_m}$$
$$\int_0^1\left((-1)^{n_m+\kappa_m}\frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!}x_m^{n_m+\kappa_m}\int_0^{x_m}(-1)^{q_m}\xi_m^{q_m}\partial^{N+\bar{1}}I(\boldsymbol{X}\xi)d\xi_m\right)dx_m\Big]\frac{d\xi}{d\xi_m}\Bigg\}\frac{d\boldsymbol{x}}{dx_m}.$$

with $\frac{d\xi}{d\xi_m} = d\xi_1 \times \cdots \times d\xi_{m-1} \times d\xi_{m+1} \times \cdots \times d\xi_m$ *and* $\frac{d\boldsymbol{x}}{dx_m} = dx_1 \times \cdots \times dx_{m-1} \times dx_{m+1} \times \cdots \times dx_m$. *(Note that the first 2 integrals in the above relation are with multi-indices while the 2 others are classical integrals).*

Let

$$\Omega = \int_0^1 (-1)^{n_m+\kappa_m} \frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!} x_m^{n_m+\kappa_m} \int_0^{x_m} (-1)^{q_m} \xi_m^{q_m} \partial^{N+\bar{1}} I(\boldsymbol{X}\xi) d\xi_m dx_m. \tag{54}$$

By integrating by parts with respect to the second integral in (54) we have

$$\Omega_1 = \int_0^1 (-1)^{n_m+\kappa_m} \frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!} x_m^{n_m+\kappa_m} \int_0^{x_m} (-1)^{q_m+1} q_m \xi_m^{q_m-1} \partial^{(N+\bar{1}-1_m)} I(\boldsymbol{X}\xi) d\xi_m dx_m$$
$$+ \int_0^1 (-1)^{N_m+\kappa_m} \frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!} x_m^{N_m+\kappa_m} \partial^{(N+\bar{1}-1_m)} I(\boldsymbol{X}\boldsymbol{x}) dx_m$$

Now, integrate by parts with respect to the first integral in (54):

$$\begin{aligned}
\Omega_2 &= \frac{-1}{n_m+\kappa_m+1} \left\{ \int_0^1 (-1)^{n_m+\kappa_m+1} x_m^{n_m+\kappa_m+1} \frac{(1-x_m)^{\nu_m-2}}{(\nu_m-2)!} \int_0^{x_m} (-1)^{q_m} \xi_m^{q_m} \partial^{N+\bar{1}} I(\boldsymbol{X}\xi) \right. \\
&\quad \left. d\xi_m dx_m - \int_0^1 (-1)^{N+\kappa_m+1} x_m^{N+\kappa_m+1} \frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!} \partial^{N+\bar{1}} I(\boldsymbol{X}\boldsymbol{x}) dx_m \right\} \\
&= \frac{-1}{n_m+\kappa_m+1} (A-B)
\end{aligned}$$

where A *(resp.* B*) represent the first (resp. the second) term in parenthesis. Integrating by parts* A *and* B *lead to* A_1 *and* B_1 *respectively which are given by:*

$$A_1 = \int_0^1 (-1)^{N_m+\kappa_m+1} x_m^{N_m+\kappa_m+1} \frac{(1-x_m)^{\nu_m-2}}{(\nu_m-2)!} \partial^{(N+\bar{1}-1_m)} I(\boldsymbol{X}\boldsymbol{x}) dx_m$$
$$- \int_0^1 (-1)^{N+\kappa_m+1} x_m^{N_m+\kappa_m+1} \frac{(1-x_m)^{\nu_m-2}}{(\nu_m-2)!} \int_0^{x_m} q_m \xi_m^{q_m-1} \partial^{(N+\bar{1}-1_m)} I(\boldsymbol{X}\xi) d\xi_m dx_m$$

and

$$\begin{aligned}
B_1 &= \int_0^1 (-1)^{N_m+\kappa_m+1} [(N_m+\kappa_m+1) x_m^{N_m+\kappa_m} \frac{(1-x_m)^{\nu_m-1}}{(\nu_m-1)!} + x_m^{N_m+\kappa_m+1} \frac{(1-x_m)^{\nu_m-2}}{(\nu_m-2)!}] \\
&\quad \partial^{(N+\bar{1}-1_i)} I(\boldsymbol{X}\boldsymbol{x}) dx_m.
\end{aligned}$$

Then Ω_2 *will have the form:*

$$\Omega_2 = \frac{-1}{n_m+\kappa_m+1} (A_1 - B_1).$$

Recall that $\Omega = \Omega_1 = \Omega_2$ then it is possible to write:

$$\Omega = \mathfrak{a}_m \Omega_1 + (1 - \mathfrak{a}_m)\Omega_2. \tag{55}$$

Finally plugging (55) in (53) and arranging terms lead to the formula (23).

proof 9.4 (of lemma 4.1). *Antisymmetry, transitivity and totality can be easily shown.*

proof 9.5 (of theorem 4.3). *The proof is straightforward upon applying successively (23) for $m = 1, \cdots, r$.*

proof 9.6 (of proposition 5.1). *Recall the first order least squares approximation as defined in (29). Note that the minimal relation (19) can be written in the spacial domain as follows :*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; n) = \frac{(\mu + \kappa + 2n)!(-\bar{1})^{n+\kappa}}{(n+\kappa)! \boldsymbol{X}^n} \int_{\bar{0}}^{\bar{1}} (\bar{1} - \boldsymbol{x})^{\mu+n} \boldsymbol{x}^{\kappa+n} \partial^n J(\boldsymbol{X}\boldsymbol{x}) d\boldsymbol{x}. \tag{56}$$

which correspond to a projection of $\tilde{I}_{\boldsymbol{x}^n}(\bar{0}, \kappa, \mu; n)$ on $\mathcal{P}_{\bar{0}}^{\{\kappa,\mu\}}(\boldsymbol{\xi}) = 1$. This is equivalent to say that equation (56) is also satisfied on the zeroes of $\mathcal{P}_{\bar{1}}^{\{\kappa,\mu\}}(\boldsymbol{\xi})$ given by (30). Note that the partial derivative ∂^n in formula (56) disappear upon integrating n-times by parts.

proof 9.7 (of theorem 5.2). *Recall first the mono variable case for $r = 1$, (i.e. $\boldsymbol{x} = x_1$, $\boldsymbol{X} = X_1$, $N = N_1$, $\kappa = \kappa_1$, $\mu = \mu_1$, $n = n_1$) equation (17) gives:*

$$\frac{(-1)^{(n_1+\kappa_1)}(n_1+\kappa_1)!(N_1-n_1)!}{s_1^{\mu_1+\kappa_1+N_1+n_1+2}} I_{x_1^{n_1}} = \frac{1}{s_1^{N_1+\mu_1+1}} \frac{d^{n_1+\kappa_1}}{ds_1^{n_1+\kappa_1}} \frac{1}{s_1} \frac{d^{N_1-n_1}}{ds_1^{N_1-n_1}} s_1^{N_1+1} \hat{I}(s_1).$$

Back in the spacial domain this equation gives:

$$I_{x_1^{n_1}}(0;\kappa_1,\mu_1;N_1) = \frac{(-1)^{n_1+\kappa_1}(\mu_1+\kappa_1+N_1+n_1+1)!}{(N_1-n_1)!(n_1+\kappa_1)!X_1^{\mu_1+\kappa_1+N_1+n_1+1}} \int_0^{X_1} \Upsilon_1(x_1)I(x_1)dx_1 \tag{57}$$

with

$$\begin{aligned}\Upsilon_1(x_1) &= \sum_{i=0}^{N_1-n_1} \begin{pmatrix} N_1-n_1 \\ i \end{pmatrix} \frac{(N_1+1)!}{(n_1+i+1)!} \sum_{j=0}^{N_1+\kappa_1} \begin{pmatrix} n_1+\kappa_1 \\ j \end{pmatrix} \frac{(n_1+1)!}{(1+j-\kappa_1)!} \times \\ &\times \frac{(X_1-x_1)^{\nu_1+\kappa_1-j-2}(-x_1)^{i+j}}{(\mu_1+\kappa_1-j-2)!}.\end{aligned}$$

By integrating by parts n_1 times equation (57) one obtain:

$$I_{x_1^{n_1}}(0,\kappa_1,\mu_1;N_1) = \int_0^{X_1} \Omega_1(x_1)\frac{d^{n_1}I(x_1)}{dx_1^{n_1}}dx_1$$

with

$$\Omega_1(x_1) = \underbrace{\int_0^{X_1}\cdots\int_0^{X_1}}_{n_1} \frac{(-1)^{n_1+\kappa_1}(\mu_1+\kappa_1+N_1+n_1+1)!}{(N_1-n_1)!(n_1+\kappa_1)!X_1^{\mu_1+\kappa_1+N_1+n_1+1}}\Upsilon_1(x_1)dx_1.$$

Let $q_1 = N_1 - n_1$. It was shown in [31] that

$$\Omega_1(x_1) = \mathcal{K}_{q_1}(0,x_1) = \sum_{i=0}^{q_1} \frac{P_i^{\{\kappa_1,\mu_1\}}(0)P_i^{\{\kappa_1,\mu_1\}}(x_1)}{\|P_i^{\{\kappa_1,\mu_1\}}\|^2}. \tag{58}$$

On the other hand take equation (16), (17) and rewrite it as :

$$\begin{aligned}\frac{(-1)^{(n_1+\kappa_1)}(n_1+\kappa_1)!(N_1-n_1)!}{\boldsymbol{s}^{\mu_1+\kappa_1+N_1+n_1+2}} I_{\boldsymbol{x}^n}(\bar{0},\kappa,\mu;N) &= \frac{1}{s_1^{N_1+\mu_1+1}}\partial^{n_1+\kappa_1}\frac{1}{s_1}\partial^{N_1-n_1}s_1^{N_1+1}\times\cdots \\ &\times \frac{1}{s_m^{N_m+\mu_m+1}}\partial^{n_m+\kappa_m}\frac{1}{s_m}\partial^{N_m-n_m}s_m^{N_m+1}\cdots\hat{I}(\boldsymbol{s}).\end{aligned}$$

Going back to the spacial domain one obtain :

$$I_{\boldsymbol{x}^n}(\bar{0}) = \int_{\bar{0}}^{\boldsymbol{X}} \prod_{m=1}^{r} \left(\Omega_m(x_m)\right)\partial^n I(\boldsymbol{x})d\boldsymbol{x}.$$

By using formula (28) and (58) we deduce that :

$$\prod_{m=1}^{r} (\Omega_m(x_m)) = \boldsymbol{\mathcal{K}}_q(\bar{0}, \boldsymbol{x}). \tag{59}$$

The formula (59) show that the non minimal estimator of order q can be expressed as an orthogonal projection in a multivariate Jacobi basis of order q given by :

$$I_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; N) = \sum_{l=\bar{0}}^{q} \frac{\langle \mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{x}), \partial^{(n)} I(\boldsymbol{X}\boldsymbol{x})\rangle}{\|\mathcal{P}_l^{\{\kappa,\mu\}}\|^2} \mathcal{P}_l^{\{\kappa,\mu\}}(\bar{0}) \tag{60}$$

proof 9.8 (of proposition 5.3). *Recall first the multivariate Bernstein polynomials of degree q (where q and i are multi-indices in $\mathbb{N}^r$) on the interval $[0,1]^r$ given by*

$$\begin{aligned} \mathcal{B}_i^q(\boldsymbol{x}) &= \begin{pmatrix} q \\ i \end{pmatrix} (\bar{1} - \boldsymbol{x})^{q-i} \boldsymbol{x}^i \\ &= \begin{pmatrix} q_1 \\ i_1 \end{pmatrix} (1 - x_1)^{q_1 - i_1} x_1^{i_1} \times \cdots \times \begin{pmatrix} q_r \\ i_r \end{pmatrix} (1 - x_r)^{q_r - i_r} x_r^{i_r} \\ &= \prod_{m=1}^{r} B_{i_m}^{q_m}(x_m), \end{aligned} \tag{61}$$

where $B_{i_m}^{q_m}(x_m)$ are the monovariate Bernstein polynomials. Note from [36] and references therein that it is possible to write each monovariate Jacobi polynomial of degree l_1 where $l_1 \leq q_1$ in the Bernstein basis to the order q_1 using the following formula:

$$P_{l_1}^{\{\kappa_1,\mu_1\}}(x_1) = \sum_{i_1=0}^{q_1} N_{l_1,i_1} B_{i_1}^{q_1}(x_1)$$

where

$$N_{l_1,i_1} = \frac{1}{\begin{pmatrix} q_1 \\ i_1 \end{pmatrix}} \sum_{j=\max(0,i_1+l_1-q_1)}^{\min(i_1,l_1)} (-1)^{l_1-j} \begin{pmatrix} q_1 - l_1 \\ i_1 - j \end{pmatrix} \begin{pmatrix} l_1 + \kappa_1 \\ j \end{pmatrix} \begin{pmatrix} l_1 + \mu_1 \\ l_1 - j \end{pmatrix}$$

Accordingly, in the multivariate case, we have:

$$\mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{x}) = \sum_{i=0}^{q} M_{l,i} \mathcal{B}_i^q(\boldsymbol{x}),$$

where $M_{l,i}$ *is a constant depending on* $N_{l_m,i_m},\ m \in [0, r]$.

Furthermore, let us consider the minimal algebraic estimator formula (19), it can be written in the time domain as follows

$$I_{\boldsymbol{x}^n}(\bar{0}; \kappa, \mu; n) = \frac{(\mu + \kappa + 2n + 1)!}{(\mu + n)!(\kappa + n)!} \int_{\bar{0}}^{\bar{1}} (\bar{1} - \boldsymbol{x})^{\mu+n} \boldsymbol{x}^{\kappa+n} \partial^n I(\boldsymbol{X}\boldsymbol{x}) d\boldsymbol{x} \qquad (62)$$

Note that the differentiation ($\partial^n I(\boldsymbol{X}\boldsymbol{x})$) *under the integral sign is only formal, it disappear upon intgrating by parts* n *times. Consider now the sum:*

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l I_{\boldsymbol{x}^n}(\bar{0}; \kappa_l, \mu_l; n), \qquad (63)$$

where the $\boldsymbol{\lambda}_l \in \mathbb{R}, l \in [\bar{0}, q]$ *are to be determined and* $\kappa_l = \kappa + q - l$ *and* $\mu_l = \mu + l$ *multi-indices in* $\mathbb{N}^r$.

Let $\gamma_{\kappa,\mu,n} = \frac{(\mu+\kappa+2n+1)!}{(\mu+n)!(\kappa+n)!}$ *the above sum can be written in the form*

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l I_{\boldsymbol{x}^n}(\bar{0}; \kappa_l, \mu_l; n) = \int_{\bar{0}}^{\bar{1}} \mathcal{D}(\boldsymbol{x})(\bar{1} - \boldsymbol{x})^{\mu+n} \boldsymbol{x}^{\kappa+n}\ \partial^n I(\boldsymbol{X}\boldsymbol{x}) d\boldsymbol{x} \qquad (64)$$

with $\mathcal{D}(\boldsymbol{x})$ *given by:*

$$\mathcal{D}(\boldsymbol{x}) = \sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l \gamma_{\kappa_l,\mu_l,n} \boldsymbol{x}^{q-l} (\bar{1} - \boldsymbol{x})^l. \qquad (65)$$

On the other hand, recall the multivariate reproducing kernel property (28) we can write

$$\boldsymbol{\mathcal{K}}_q(\boldsymbol{\xi},\boldsymbol{x}) = \sum_{l=\bar{0}}^{q} \frac{\mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{\xi})\mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{x})}{\|\mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{x})\|^2} = \sum_{l=\bar{0}}^{q} \chi_l(\boldsymbol{\xi})\mathcal{P}_l^{\{\kappa,\mu\}}(\boldsymbol{x}) \tag{66}$$

Next express the Jacobi polynomials in the above relation in a q^{th} order Bernstein polynomials, we obtain:

$$\boldsymbol{\mathcal{K}}_q(\xi,\boldsymbol{x}) = \sum_{l=\bar{0}}^{q} \chi_l(\boldsymbol{\xi}) \sum_{i=\bar{0}}^{q} M_{l,i}\mathcal{B}_i^q(\boldsymbol{x}). \tag{67}$$

Finally by equating

$$\boldsymbol{\mathcal{K}}_q(\xi,\boldsymbol{x}) = \mathcal{D}(\boldsymbol{x}) \tag{68}$$

One can see that for any $\boldsymbol{\xi} = (\xi_1,\cdots,\xi_r) \in [0,1]^r$, there exist a unique set of $\boldsymbol{\lambda}_l$ noted from now on $\boldsymbol{\lambda}_l(\boldsymbol{\xi})$ satisfying relation (68). This set of $\boldsymbol{\lambda}_l(\boldsymbol{\xi})$ is determined upon identifying corresponding powers of $\boldsymbol{x}^{q-l}(\bar{1}-\boldsymbol{x})^l$, and this set is unique because $\boldsymbol{\lambda}_l(\boldsymbol{\xi})$ appear linearly in $\mathcal{D}(\boldsymbol{x})$ in (68). From the mono variable case [31] it is easy to see that

$$\sum_{l=\bar{0}}^{q} \boldsymbol{\lambda}_l(\boldsymbol{\xi}) = 1. \tag{69}$$

proof 9.9 (of proposition 6.1). *In fact we can verify the following relation :*

$$\tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa,\mu;n) = \sum_{i=1}^{2^r} \frac{1}{2^r} \tilde{I}_{\boldsymbol{x}^n}(\bar{0};\kappa + \mathbf{L}(i), \mu + \mathbf{L}(2^r+1-i);n),$$

where $\mathbf{L}$ is given by (24).

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover, 1965.

[2] M.A. Al-Alaoui. *A class of second order integrators and low-pass differentiators*, IEEE Trans. Circuits Syst. I, vol. 42, No. 4, pp. 220-223, 1995.

[3] R. S. Anderssen and P. Bloomfield, "Numerical Differentiation Procedures for Non-Exact Data," Numer. Math. vol. 22, pp. 157-182 Springer-Verlag 1974.

[4] G. Aubert and L. Vese, "A variational method in image recovery". *SIAM Journal on Numerical Analysis*, vol. 34, No. 5, pp. 1948-1979, 1997.

[5] Y. Bachalany, S. Riachy, M. Mboup and J.-P. Richard. "Différentiation numérique multivariable II : projection orthogonale et filtrage RIF". 6^{ieme} *Conférence internationale francophone d'automatique*, 2010.

[6] J.-P. Barbot, D. Boutat, and T. Floquet. *An observation algorithm for nonlinear systems with unknown inputs*, Automatica, Vol. 45, No. 8, pp. 1970-1974, 2009.

[7] A. Brook, R. Kimmel and N.A. Sochen, "Variational Segmentation for Color images". *International Journal of Computer Vision*, vol. 18, pp. 247-268, 2003.

[8] A. Bruhn, J. Weickert and C. Schnörr, "Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods". *International Journal of Computer Vision*, vol. 61, pp. 211-231, 2005.

[9] J. Cullum, "Numerical differenciation and regularization," SIAM Journal on numerical analysis Vol. 8, No. 2, pp. 254-265, 1971.

[10] T.E. Duncan, P. Mandl et B. Pasik-Duncan "Numerical differentiation and parameter estimation in higher-order linear stochastic systems". *IEEE Trans. Automat. Contr.*, vol. 41, pp. 522-532, 1996.

[11] C.K. Chen and J.H. Lee,"Design of high-order digital differentiators using L_1 error criteria." *IEEE Trans. Circuits Syst. II*, vol. 42, No. 4, pp. 287-291, 1995.

[12] M. Fliess and C. Join, "Intelligent PID controllers" 16^{th} *Mediterranean conference on control and automation*, june 25-27, 2008. (available at http://hal.inria.fr).

[13] M. Fliess and C. Join, "Model free control and Intelligent PID controllers: Towards a possible trivialization of nonlinear control?" 15^{th} *IFAC Symposium on system identification*, 2009. (available at http://hal.inria.fr).

[14] M. Fliess and C. Join, "Towards new technical indicators for trading systems and risk management" 15^{th} *IFAC symposium on system identification*, 2009. (available at http://hal.inria.fr).

[15] M. Fliess and C. Join, "A mathematical proof of the existence of trends in financial time series" *Systems Theory: Modelling, Analysis and Control*, 2009. (available at http://hal.inria.fr).

[16] M. Fliess, C. Join, M. Mboup and A. Sedoglavic, "Estimation des dérivées d'un signal multidimensionnel avec application aux images et aux videos", *Actes* 20^e *Col. GRETSI*, Louvain-la-neuve 2005 (available at http://hal.inria.fr/inria-00001116).

[17] M. Fliess, C. Join, M. Mboup and H. Sira-Ramirez, "Compression différentielle de transitoires bruités", *C.R. Acad. Sci. Paris* Ser. I, vol.339, pp. 821-826, 2004.

[18] M. Fliess, C. Join, M. Mboup and H. Sira-Ramirez, "Analyse et representation de signaux transitoires: application à la compression, au débruitage et à la detection de ruptures", *Actes* 20^e *Coll. GRETSI*, Louvain-la-neuve, 2005 (available at http://hal.inria.fr/inria-00001115).

[19] M. Fliess, C. Join et H. Sira-Ramrez, "Non-linear estimation is easy ", *Int. J. Modelling Identification and Control 3,* vol. 4, No. 1, pp. 12-27, 2008. Available from http://hal.inria.fr/inria-00158855/fr/.

[20] M. Fliess and H. Sira-Ramirez, "An algebraic framework for linear identification" *ESAIM. COCV*, vol. 9, pp. 151-168, 2003.

[21] P.-A. Gédouin, C. Join, E. Delaleau, J.-M. Bourgeot, S. Arbab and S. Calloch"Model free control of shape memory alloys and antagonistic actuators" 17^{th} *IFAC world congress Seoul Korea*, 2008. (available at http://hal.inria.fr).

[22] B. Horn and B. Schunck, "Determining optical flow". *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.

[23] A. Horwitz, "A generalization of Simpsons Rule" *Approximation Theory and its Applications*, vol. 9, pp. 71-80, 1993.

[24] S. Ibrir et S. Diop, "A numerical procedure for filtering and efficient high-order signal differenciation". *Int. J. Appl. Math. Compt. Sci.*, vol. 14, pp. 201-208, 2004.

[25] C. Join and S. Tabbone, “Robust curvature extrema detection based on new numerical derivation” *Advanced Concepts for Intelligent Vision Systems, ACIVS*, 2008.

[26] C. Lanczos, “Applied Analysis”, *Prentice-Hall, Englewood Cliffs,* NJ, 1956.

[27] A. Levant, “Higher-order sliding modes, differentiation and output-feedback control” *International Journal of Control* vol. 76, pp. 924-941, 2003.

[28] D. Liu, O. Gibaru, W. Perruquetti, “Differentiation by integration with Jacobi polynomials” *Journal of Computational and Applied Mathematics*, vol. 235, No. 9, pp. 3015-3032, 2010.

[29] M. Mboup, “Parameter estimation for signals described by differential equations”, *Applicable Analysis*, Vol. 88, No. 1, pp. 29-52, 2009.

[30] M. Mboup, C. Join and M. Fliess. “A revised look at numerical differentiation with an application to nonlinear feedback control”. 15^{th} *Mediterranean conference on control and automation*, june 27-29, 2007.

[31] M. Mboup, C. Join and M. Fliess. “Numerical Differentiation with annihilators in noisy environment”. *Numerical Algorithms*, vol. 5, No. 4, pp. 439-467, 2009.

[32] J. Mikusinski, *Operational calculus*, vol. 1 and 2 Varsovie and Oxford University Press, Oxford 1983.

[33] J.M. Morel, S. Solimini and A. Brook, "Segmentation of images by variational methods: A constructive approach". *Revista Matematica de la Universidad Complutense de Madrid*, vol. 1, pp. 169-182, 1998.

[34] S. Riachy, Y. Bachalany, M. Mboup and J.-P. Richard. "An algebraic method for multi-dimensional derivative estimation". 16^{th} *Mediterranean conference on control and automation*, 2008.

[35] S. Riachy, Y. Bachalany, M. Mboup and J.-P. Richard. "Différenciation numérique multivariable I : estimateurs algébriques et structure". 6^{ieme} *Confrence internationale francophone d'automatique*, 2010.

[36] A. Rababah. "Jacobi-Bernstein basis transformation." *Computational methods in applied mathematics*, Vol. 4, No. 2, pp. 206-214, 2004.

[37] S.K. Rangarajana, S.P. Purushothaman, "Lanczos generalized derivative for higher orders," *Journal of Computational and Applied Mathematics*, vol. 177, pp. 461-465, 2005.

[38] R.A. Roberts et C.T. Mullis. "Digital signal processing". *Addison-Wesley*, 1987.

[39] Z. Wang, R. Wen, "Numerical differentiation for high orders by an integration method," *Journal of Computational and Applied Mathematics* vol. 234, pp. 941-948, 2010.

[40] J. Weickert and C. Schnörr, "Variational optic flow computation with a spatio-temporal smoothness constraint". *Journal of Mathematical Imaging and Vision*, vol. 14, pp. 245-255, 2001.

[41] J. Weickert, "Anisotropic diffusion in image processing," *ECMI Series, Teubner-Verlag, Stuttgart, Germany*, 1998.