



HAL
open science

Exact Arma Lattice Predictors From Autocorrelation Functions

André Monin, Gérard Salut

► **To cite this version:**

André Monin, Gérard Salut. Exact Arma Lattice Predictors From Autocorrelation Functions. IEEE Transactions on Signal Processing, 1994, 42 (4), pp.877 - 886. 10.1109/78.285651 . hal-02916431

HAL Id: hal-02916431

<https://laas.hal.science/hal-02916431>

Submitted on 17 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exact Arma Lattice Predictors

From Autocorrelation Functions.

André MONIN, Gérard SALUT.

Laboratoire d'Automatique et d'Analyse des Systèmes.
Centre National de la Recherche Scientifique.
7 Avenue du Colonel Roche 31077 TOULOUSE Cédex. FRANCE

Abstract: This paper derives an optimal linear L^2 -predictor of ARMA type in lattice form of arbitrarily fixed dimension for a process whose autocorrelation function is known. The algorithm preserves exact optimality at each step, as opposed to asymptotic convergence of more usual algorithms, at the expense of hereditary computation. Only the discrete time case is examined. It is shown how the unnormalized (respectively normalized) lattice form may be reduced to only $4n-2$ parameters (respectively $2n+1$) for a n -th order projection on the past. The normalization algorithm for the forward and backward residuals uses only scalar square root computations. Some examples are given which show the accuracy of this technique compared to those using the classical ARMA form for the predictor.

Keywords: *Linear filtering, linear predictor, reduced filters, lattice form.*

1. Introduction.

1.1. Objectives.

The paper concerns the problem of computing the exact minimum variance ARMA predictor of n -th order, as a time varying lattice, from the autocorrelation function. Approximate procedures have been proposed in the past [1], [2], [3], [4], [5] to achieve the same goal, namely by trivially imbedding the past innovations into a 2-channel standard lattice problem. However, it has been shown by [1], as well as noticed in [5], that this is equivalent to the extended least-squares approximation [6], therefore with the same pitfalls (computed transient coefficients are not the true projection coefficients, and do not guarantee convergence to the optimal, in the general case. For the same reason, the predictor is always biased in the short-term transient.). More recent papers have used stochastic gradient approximations of higher order [7], [8], [9], [10]. Except the fact that they use data records while we use autocorrelations, the problem is indeed similar to ours. However, the aim is to asymptotically determine the stationary optimal predictor parameters, not to obtain the optimal time-varying parameters. The techniques developed are essentially based on extensions of the well-known stochastic gradient search. They are recursive but do not look for the optimum at each step. Moreover, some convergence problem may occur when, for example, the dimension of the unknown system is greater than the model [11]. It is obviously the case when the system is not finite dimensional, which is the case in the low-pass filter treated as an illustration in 5.

To derive the exact ARMA lattice predictor, thus avoiding such pitfalls, is a more difficult task since the feedback parameters of the lattice lead to a non-linear implicit dependence or, equivalently, a growing memory for the computation of the time-varying optimal predictor pa-

rameters. This is equivalent to the exact stochastic gradient (i.e. an infinite order approximation). This difficulty was clearly identified in [1], [2] but was never fully investigated, to the author's knowledge. We examine here this problem for a known autocorrelation function.

Our results, in this respect, are too-fold:

- A general algorithm gives the exact (hereditary) computation of the optimal n-th order ARMA predictor, in time-varying lattice form with zero initial state, for an arbitrary autocorrelation function. The linear dependence in the vector process (y, \tilde{y}) (infinite Hankel rank) is being exploited to obtain the minimum number of parameters.

- In the finite Hankel rank case, the computations are shown to be finitely recursive, and completely solve the so-called realization problem under lattice form.

1.2. Contents.

Let $\{y_t, t \in N\}$, be a discrete-time scalar sequence of $L_p^2(\Omega)$ (mean square random variables) with zero mean. Assume that the only data available is the autocorrelation function of y_t , not necessarily with a finite Hankel rank. One can derive, as in [12], the reduced optimal n-th order non-stationary linear predictor $z_t = \hat{y}_t|_{t-1}$ of y_t (in the minimum variance sense), under its observable canonical state-space form or, equivalently, its input-output ARMA form:

$$z_t = \sum_{i=1}^n a_t^i z_{t-i} + b_t^i y_{t-i} \quad (1.1)$$

The optimal parameters are obtained by minimizing the prediction variance error ($J_t = E[(y_t - z_t)^2]$) with no stationary assumption but explicit state-dimension constraint on the predictor, as shown in [12]. Otherwise, the state variable has to be reconstructed for the whole time horizon, at each step, leading to an infinite-dimensional state-space as in ARMA identification [1].

In such a predictor, computing the values of $\{a_t^i, b_t^i\}_{i=1}^n$ requires the inversion of a $2n \times 2n$ dimensional system at each step. Despite the possibility of recursive inversion, as in [1], the canonical state-space form is generally not tractable for higher dimensions. It is in fact well known that numerical computations are very sensitive with respect to the coefficients of the characteristic polynomial.

Recall that the optimization of J_t leads to the orthogonality equations as:

$$\forall t \geq 0, \forall i = 1 \dots n-1, E[(y_t - z_t)y_{t-i}] = 0 \quad E[(y_t - z_t)z_{t-i}] = 0 \quad (1.2)$$

Clearly, that equations (1.2) represent the projection of y_t on the random variables

$\{y_{t-i}\}_{i=1}^n$ and $\{z_{t-i}\}_{i=1}^n$ with $(x|y) = E[xy]$ as inner product. We write this projection:

$$z_t = P[y_t | y_{t-1}, \dots, y_{t-n}, z_{t-1}, \dots, z_{t-n}] \quad (1.3)$$

The purpose of this paper is to exhibit an orthogonal basis of the linear space spanned by the random variables $\{y_{t-i}, z_{t-i}\}_{i=1}^n$ in order to simplify the projection formula and to improve its numerical accuracy. This orthogonal basis induces a lattice type filter but without the independent modular structure, which is well known to be impossible in the ARMA case. Note that this problem has been studied in [1], [2] in the context of the ARMA identification, but with the extended least squares approximate.

The main contribution here is as follows:

- We first exhibit the optimal n -th order predictor (reduced dimension) under the unnormalized lattice form using $4n-2$ parameters ($2n$ -dimensional state-space) and a complete algorithm is provided which uses hereditary computation on the autocorrelation function $y_t y_{t-\tau}$.

- We then introduce the normalized form using $2n+1$ parameters with its algorithm attached. The problem of extracting square roots of the normalization matrices is solved using only scalar square root computation because of its particular form.

- We also show that such a projection technique (i.e. onto a linear functional of the past with internal memory constraint) yields an associated innovation whose “whiteness” has the same restricted meaning: orthogonality to an n -dimensional sub-space of the past. It yields a linear model of y with “ n -th order white noise”, whose approximation properties are illustrated through applications.

The applications concern the most useful case, and deal with autocorrelation function of infinite Hankel rank (low-pass and fractional order filters), one of which is well known for its sensitivity problems. By converting the role of y and \tilde{y} , it must be noticed that the algorithm gives the exact n -th order best L_p^2 -approximate with minimum phase, an open problem in the literature on optimum systems approximations [13].

2. The orthogonalisation formulas.

Let $X, Y \in (L_p^2)^n$ with inner and outer products written as usual $E[X^T Y]$ and $E[XY^T]$.

The linear projection of a vector X of $(L_p^2)^n$ on the linear product space spanned by vectors Y_1, \dots, Y_p of $(L_p^2)^n$ is obtained by solving the following $n \times p$ dimensional linear system as:

$$\forall i = 1 \dots p, E[XY_i^T] = \sum_{j=1}^p K_j E[Y_j Y_i^T] \quad (2.1)$$

where K_i are $n \times n$ matrices. The projection takes then the following form:

$$P[X|Y_1, \dots, Y_p] = \sum_{i=1}^p K_i Y_i \quad (2.2)$$

The orthogonal projection of a vector X of $(L_p^2)^n$ on a scalar random variable z of L_p^2 leads to:

$$E[(X - P[X|z])z] = 0, \text{ i.e. } P[X|z] = \frac{E[Xz]}{E[z^2]} z. \quad (2.3)$$

In the same way, one may compute the orthogonal projection of a scalar z of L_p^2 on a vector Y of $(L_p^2)^n$. One finds:

$$P[z|Y] = E[zY^T](E[YY^T])^{-1} Y \quad (2.4)$$

We now recall some properties that will be used further to reduce the number of parameters of the predictor in lattice form.

Proposition 1. Let X and Y be vectors of $(L_p^2)^2$. Write $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. If $E[Xy_1] = 0$ and $E[y_1y_2] = 0$, then $P[X|Y] = P[X|y_2]$. \square

Proposition 2. Let Z_1, \dots, Z_p be orthogonal vectors of $(L_p^2)^n$, that is $\forall i, j = 1 \dots n$, $i \neq j \Rightarrow E[Z_i Z_j^T] = 0$. Then the projection of a vector X of $(L_p^2)^n$ on the space spanned by $\{Z_i\}_{i=1}^n$ is: $P[X|Z_1, \dots, Z_p] = P[X|Z_1] + \dots + P[X|Z_p]$. \square

3. Unnormalized lattice form.

Recall that our purpose is to compute the n -th order linear predictor z_t as a projection of the observation signal y_t on the n -th order memory of the past as in (1.3). This will be achieved in two steps. First we exhibit an orthogonal basis of the past $\{y_{t-i}, z_{t-i}\}_{i=1}^n$ using the Levinson formulas. We then reduce the dimension of the involved system using the degeneracy property of the projection space (clearly, $\{y_{t-i}, z_{t-i}\}_{i=1}^n$ is only n -dimensional). We then show how the computation of the predictor's parameters is achieved knowing the autocorrelation function of y_t .

3.1. Forward and backward residuals formulas.

Define $Z_t = \begin{bmatrix} \tilde{y}_t \\ z_t \end{bmatrix}$, where \tilde{y}_t denotes the innovation $y_t - z_t$ in the sense of (1.1)-(1.3). It

is clear from (2.4) that the projection of y_t on the $2n$ scalars $\{y_{t-i}, z_{t-i}\}_{i=1}^n$ can be equivalently computed as the projection of y_t on the n vectors $\{Z_{t-i}\}_{i=1}^n$. That is:

$$z_t = P[y_t | Z_{t-1}, \dots, Z_{t-n}] \quad (3.1)$$

The aim of this paragraph is to recall an orthogonalisation of the space spanned by $\{Z_{t-i}\}_{i=1}^n$ [1], using the Graham-Shmidt procedure. Let:

$$\begin{cases} X_t^0 = Z_t \\ X_t^1 = Z_{t-1} - P[Z_{t-1} | Z_t] \\ \dots \\ X_t^n = Z_{t-n} - P[Z_{t-n} | Z_{t-n+1}, \dots, Z_t] \end{cases} \quad (3.2)$$

be the backward residuals of the family $\{Z_t, \dots, Z_{t-n+1}\}$ and

$$\left\{ \begin{array}{l} Z_t^0 = Z_t \\ Z_t^1 = Z_t - P[Z_t|Z_{t-1}] \\ \dots \\ Z_t^n = Z_t - P[Z_t|Z_{t-n}, \dots, Z_{t-1}] \end{array} \right. \quad (3.3)$$

be the forward ones.

We have then the following properties:

Property 1. $\{X_t^0, \dots, X_t^{n-1}\}$ is orthogonal, that is $\forall i \neq j, E[X_t^i(X_t^j)^T] = 0$.

Property 2. $\{Z_{t-n+1}^0, \dots, Z_t^{n-1}\}$ is orthogonal.

Property 3. $\forall i = 1 \dots n-1, \{Z_t^i, \dots, Z_{t-i}^i\}_{sp} = \{X_t^0, \dots, X_t^i\}_{sp}$.

Property 4. $\forall i = 1 \dots n-1, \{Z_t^i, \dots, Z_{t-i}^i\}_{sp} = \{Z_{t-i}^0, \dots, Z_t^i\}_{sp}$.

One can then rewrite the relations (3.2) and (3.3) as follows:

$$\left\{ \begin{array}{l} X_t^0 = Z_t \\ X_t^1 = X_{t-1}^0 - P[X_{t-1}^0|Z_t^0] \\ \dots \\ X_t^n = X_{t-1}^{n-1} - P[X_{t-1}^{n-1}|Z_t^{n-1}] \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} Z_t^0 = Z_t \\ Z_t^1 = Z_t^0 - P[Z_t^0|X_{t-1}^0] \\ \dots \\ Z_t^n = Z_t^{n-1} - P[Z_t^{n-1}|X_{t-1}^{n-1}] \end{array} \right. \quad (3.4)$$

According to the last equation of system (3.3), the first component of Z_t^n is equal to \tilde{y}_t while the second one is equal to zero. In fact, a straightforward calculation leads to:

$$\begin{bmatrix} 1 & 0 \end{bmatrix} Z_t^n = \tilde{y}_t - P[\tilde{y}_t|Z_{t-1}, \dots, Z_{t-n}]. \quad (3.5)$$

But, according to (1.2), one has:

$$P[\tilde{y}_t|Z_{t-1}, \dots, Z_{t-n}] = 0. \quad (3.6)$$

On the other hand:

$$\begin{bmatrix} 0 & 1 \end{bmatrix} Z_t^n = z_t - P[z_t|Z_{t-1}, \dots, Z_{t-n}], \quad (3.7)$$

$$\text{since} \quad P[z_t|Z_{t-1}, \dots, Z_{t-n}] = P[y_t|Z_{t-1}, \dots, Z_{t-n}] = z_t. \quad (3.8)$$

From the definition of the projection, the linear predictor becomes:

$$z_t = P[z_t|X_{t-1}^0] + \dots + P[z_t|X_{t-1}^{n-1}] \quad (3.9)$$

The implicit aspect of this formula can be eliminated by using the orthogonality of the innovation \tilde{y}_t with the past $X_{t-1}^0, \dots, X_{t-1}^{n-1}$. It leads to an explicit expression of the predictor, gi-ven the statistics of the observation signal:

$$z_t = P[y_t|X_{t-1}^0] + \dots + P[y_t|X_{t-1}^{n-1}] \quad (3.10)$$

Remark: As was noticed by Benveniste in [2], the vector Z_t is degenerated because it only depends on the scalar sequence y_t . If we examine the relations (3.4), it appears at first that the lattice predictor needs $2n$ matrices 2×2 as parameters, that is $8n$ parameters. In fact, because of the degeneracy of vector Z_t , the system can be reduced to only $4n-2$ parameters as will be shown in the next section.

3.2. Reduction of the number of parameters.

Proposition 3. $\forall i = 0 \dots n-1$, $P[X_{t-1}^i|Z_t^i] = P[X_{t-1}^i|z_t^i]$ where z_t^i is the second component of Z_t^i . \square

Proof: Recall that, according to (3.3), one has: $Z_t^i = Z_t - P[Z_t|Z_{t-1}, \dots, Z_{t-i}]$

Let us examine the first component of these vectors. One has

$$\begin{bmatrix} 1 & 0 \end{bmatrix} Z_t^i = \tilde{y}_t - P[\tilde{y}_t|Z_{t-1}, \dots, Z_{t-i}] = \tilde{y}_t. \quad (3.11)$$

because of the orthogonality relations (1.2). Let us then check the hypothesis of Proposition 1:

$$E[X_{t-1}^i(Z_t^i)^T] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = E[\tilde{y}_t Z_{t-1-i}] - E[\tilde{y}_t P[Z_{t-1-i}|Z_{t-i}, \dots, Z_{t-1}]] \quad (3.12)$$

But according to Property 4, one has:

$$P[Z_{t-1-i}|Z_{t-i}, \dots, Z_{t-1}] = P[Z_{t-1-i}|Z_{t-i}^0, \dots, Z_{t-1}^{i-1}] \quad (3.13)$$

Applying then the Proposition 2, we have:

$$P[Z_{t-1-i}|Z_{t-i}, \dots, Z_{t-1}] = P[Z_{t-1-i}|Z_{t-i}^0] + \dots + P[Z_{t-1-i}|Z_{t-i}^{i-1}]. \quad (3.14)$$

The orthogonality of the family $\{Z_{t-i}^i, Z_{t-i}^{i-1}, \dots, Z_{t-i}^0\}$ implies that Z_t^i is orthogonal to $\{Z_{t-1}^{i-1}, \dots, Z_{t-1}^0\}$, so that $E[\tilde{y}_t Z_{t-1-i}^0] = \dots = E[\tilde{y}_t Z_{t-1-i}^{i-1}] = 0$. As a consequence, it is clear that

$$\forall i = 0 \dots n-1, E[X_{t-1}^i(Z_t^i)^T] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \quad (3.15)$$

On the other hand:

$$\begin{bmatrix} 1 & 0 \end{bmatrix} E[Z_t^i(Z_t^i)^T] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = E[\tilde{y}_t(z_t - P[z_t|Z_{t-1}, \dots, Z_{t-i}])] \quad (3.16)$$

By linear combination of equations (1.2) using the equality (1.1), one can easily derive that $E[\tilde{y}_t z_t] = 0$. But \tilde{y}_t is also orthogonal to $\{Z_{t-i}^j\}_{j=1}^n$, according to (1.2). It leads to:

$$\forall i = 0 \dots n-1, \quad \begin{bmatrix} 1 & 0 \end{bmatrix} E[Z_t^i (Z_t^i)^T] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \quad (3.17)$$

Applying then Proposition 1, the result follows. ■

A useful consequence of this proposition is that the backward and forward residuals of system (3.4) need only the second component of vectors $\{Z_t^i\}_{i=0}^{n-1}$. The residuals may therefore be computed using only $4n-2$ parameters by the following system.

$$\left\{ \begin{array}{l} z_t^0 = z_t \\ z_t^1 = z_t^0 - P[y_t | X_{t-1}^0] \\ \dots \\ z_t^n = z_t^{n-1} - P[y_t | X_{t-1}^{n-1}] \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} X_t^0 = Z_t \\ X_t^1 = X_{t-1}^0 - P[X_{t-1}^0 | z_t^0] \\ \dots \\ X_t^{n-1} = X_{t-1}^{n-2} - P[X_{t-1}^{n-2} | z_t^{n-2}] \end{array} \right. \quad (3.18)$$

Recall that according to (3.9), $z_t^n = 0$. It leads to the following lattice structure of computation:

$$\left\{ \begin{array}{l} z_t^0 = z_t \\ z_t^1 = z_t^0 - (k_t^0)^T X_{t-1}^0 \\ \dots \\ z_t^{n-1} = z_t^{n-2} - (k_t^{n-2})^T X_{t-1}^{n-2} \\ 0 = z_t^{n-1} - (k_t^{n-1})^T X_{t-1}^{n-1} \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} X_t^0 = Z_t = \begin{bmatrix} \tilde{y}_t \\ z_t \end{bmatrix} \\ X_t^1 = X_{t-1}^0 - q_t^0 z_t^0 \\ \dots \\ X_t^{n-1} = X_{t-1}^{n-2} - q_t^{n-2} z_t^{n-2} \end{array} \right. \quad (3.19)$$

where the output of the predictor is computed as:

$$z_t = \sum_{i=0}^{n-1} (k_t^i)^T X_{t-1}^i \quad (3.20)$$

According to (4) and (5), the parameters $\{k_t^i\}_{i=0}^{n-1}$ and $\{q_t^i\}_{i=0}^{n-2}$ are determined as follows:

$$\forall i = 0 \dots n-1, \quad \left\{ \begin{array}{l} k_t^i = (E[X_{t-1}^i (X_{t-1}^i)^T])^{-1} E[X_{t-1}^i y_t] \\ q_t^i = \frac{E[X_{t-1}^i y_t]}{E[(z_t^i)^2]} \end{array} \right. \quad (3.21)$$

The figure 1 illustrates the structure obtained.

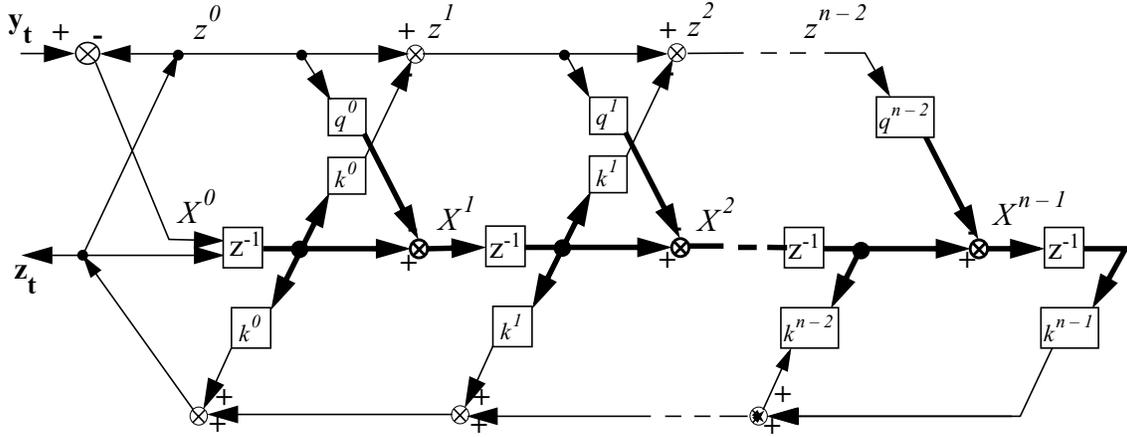


Figure 1 - Unnormalized Lattice Structure.

3.3. Off-line computation of parameters.

We now show how the above parameters are recursively computed, in the general case, (i.e. infinite Hankel rank for the process y) by a growing dimension system (as in [12]). In such a case, the algorithm needs no particular dimensional caution, but only a stationary convergence test. It should be noted that, in case of singular initialization, the first steps call for a particular formula. This will be developed further.

When the process y has a finite recursion model of order m (hankel rank), but $m > n$, it is shown how the corresponding (reduced) predictor's coefficients are themselves generated by finite recursion formulas.

3.3.1 The general algorithm.

First, let us examine the quantities to be computed to derive the lattice parameters defined in (3.21). For simplicity, we write $\bar{X} = E[X]$ the expectation of X . Define $M_t^i = \overline{X_t^i (X_t^i)^T}$. The first system of (3.19) permits us to evaluate z_t^i as:

$$\forall i = 0 \dots n-1, z_t^i = \sum_{j=i}^{n-1} (k_t^j)^T X_{t-1}^j \quad (3.22)$$

As a consequence, according to the Property 1, $X_{t-1}^i z_t^i = M_{t-1}^i k_t^i$. One then simplify the formulas (3.21) as follows:

$$\begin{cases} \forall i = 0 \dots n-1, k_t^i = (M_{t-1}^i)^{-1} \overline{X_{t-1}^i y_t} \\ \forall i = 0 \dots n-1, q_t^i = \overline{X_{t-1}^i y_t} / (z_t^i)^2 \end{cases} \quad (3.23)$$

It appears then that there are two kinds of quantities to be computed: the finitely realizable one, and the hereditary one as $\overline{X_{t-1}^i y_t}$. Note also M_t^i and $(z_t^i)^2$ are finitely realizable.

• Finitely recursive computations.

Recall that, at each step t , M_t^i and $\overline{X_{t-1}^i y_t}$ are assumed to be known, so are $\{k_t^i, q_t^i\}_{i=0}^{n-1}$.

The next value of $\overline{(z_t^i)^2}$ is computed from system (3.21) as follows:

$$\begin{cases} \overline{(z_t^{n-1})^2} = (k_t^{n-1})^T \overline{X_{t-1}^{n-1} (X_{t-1}^{n-1})^T} k_t^{n-1} \\ \overline{(z_t^i)^2} = \overline{(z_t^{i+1})^2} + 2(k_t^i)^T \overline{X_{t-1}^i z_t^i} - (k_t^i)^T \overline{X_{t-1}^i (X_{t-1}^i)^T} k_t^i, \forall i = 0 \dots n-2 \end{cases} \quad (3.24)$$

But according to (3.22) and because of the orthogonality of the family $\{X_{t-1}^0, \dots, X_{t-1}^{n-1}\}$, one has:

$$\overline{X_{t-1}^i z_t^i} = \overline{X_{t-1}^i (X_{t-1}^i)^T} k_t^i \quad (3.25)$$

The system above may then be transform in:

$$\begin{cases} \overline{(z_t^{n-1})^2} = (k_t^{n-1})^T M_{t-1}^{n-1} k_t^{n-1} \\ \overline{(z_t^i)^2} = \overline{(z_t^{i+1})^2} + (k_t^i)^T M_{t-1}^i k_t^i, \forall i = 0 \dots n-2 \end{cases} \quad (3.26)$$

In a similarly way, using system (3.19), the matrices M_t^i may be recursively computed as:

$$\forall i \geq 1, \overline{X_{t-1}^i (X_{t-1}^i)^T} = \overline{X_{t-1}^{i-1} (X_{t-1}^{i-1})^T} - q_t^i \overline{(X_{t-1}^{i-1})^T z_t^{i-1}} - \overline{X_{t-1}^{i-1} z_t^{i-1}} (q_t^{i-1})^T + q_t^{i-1} \overline{(z_t^{i-1})^2} (q_t^{i-1})^T \quad (3.27)$$

Using (3.25) a second time, one has:

$$\begin{cases} M_t^i = M_{t-1}^{i-1} - q_t^{i-1} (k_t^{i-1})^T M_{t-1}^{i-1}, \forall i \geq 1 \\ M_t^0 = \begin{bmatrix} \overline{(y_t)^2} - \overline{(z_t^0)^2} & 0 \\ 0 & \overline{(z_t^0)^2} \end{bmatrix} \end{cases} \quad (3.28)$$

Note that M_t^0 is well defined at step t because $\overline{y_t^2}$ is assumed to be given data in the problem and $\overline{(z_t^0)^2}$ is computed from (3.24).

Hereditary computations (for infinite Hankel rank of y).

We use such an expression in the usual sense of non-finitely recursive computation (as a function, here, of the time horizon of the autocorrelation function y_t, y_τ).

Although $\{X_t^i\}_{i=0}^{n-1}$ are recursive quantities, the expectation $\overline{X_{t-1}^i y_t}$ is not so. In fact, its computation needs all the parameters $\{k_\tau^i, q_\tau^i, \tau \leq t\}_{i=0}^{n-1}$ from the beginning of the algorithm. System (3.18) allows us to derive the following formula:

$$\forall 0 \leq \tau \leq t-1, \left\{ \begin{array}{l} \overline{X_{\tau}^0 y_t} = \begin{bmatrix} \overline{y_{\tau} y_t - z_{\tau}^0 y_t} \\ \overline{z_{\tau}^0 y_t} \end{bmatrix} \\ \overline{X_{\tau}^i y_t} = \overline{X_{\tau-1}^{i-1} y_t} - q_{\tau}^{i-1} \overline{z_{\tau}^{i-1} y_t}, \forall i = 1 \dots n-1 \end{array} \right. \quad (3.29)$$

It appears then that a second hereditary quantity is to be computed, that is $\overline{z_{\tau}^{i-1} y_t}$. A similar computation leads to:

$$\forall 0 \leq \tau \leq t-1, \left\{ \begin{array}{l} \overline{z_{\tau}^{n-1} y_t} = (k_{\tau}^{n-1})^T \overline{X_{\tau-1}^{n-1} y_t} \\ \overline{z_{\tau}^i y_t} = \overline{z_{\tau}^{i+1} y_t} + (k_{\tau}^i)^T \overline{X_{\tau-1}^i y_t}, \forall i = 0 \dots n-2 \end{array} \right. \quad (3.30)$$

where the autocorrelation function of y_t appears as the input of the hereditary recursion (3.29)-(3.30). The figure 2 presents the sequence of computation.

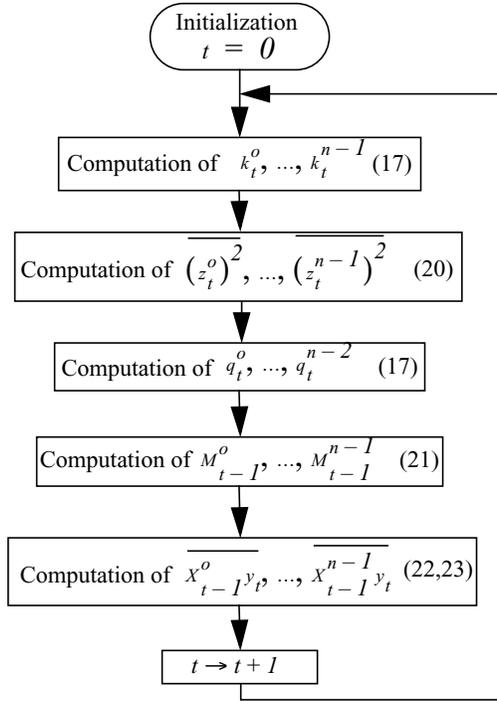


Figure 2 - Unnormalized Lattice Algorithm.

3.3.2 Additional formulas for singular initialization.

Without a priori information on the past, the initial estimator at $t = 0$ is the a priori expectation of the stochastic process to be estimated, so that $z_0 = y_0$. For a zero mean process, this value is zero.

At the first step ($t = 1$), the only relevant data is y_0 since no information on past values of the predictor is available. As a consequence, the projection of y_1 on the past is $z_1 = P[y_1|y_0]$ and only the first components of k_1^0 and q_1^0 are determined, since M_0^0 is singular.

At the second step ($t = 2$), one may then project y_2 on $\{y_1, z_1\}$. This yields the matrix M_1^0 non-singular but not M_1^1 . Therefore, k_2^0 and q_2^0 are well defined as M_1^0 is regular.

At step $t = 3$, k_3^0 and q_3^0 are well defined but the first components of k_3^1 and q_3^1 are de-

terminated. And so on...

Continuing this procedure, one finds that:

Proposition 4. *For all $i \leq n$, after $2i-2$ iterations, one may just define:*

$$z_{2i-2} = P[y_{2i-2} | X_{2i-3}^0, \dots, X_{2i-3}^{i-2}]$$

and $\{k_{2i-2}^j, q_{2i-2}^j\}_{j=0}^{i-2}$ are well determined because the matrices $\{M_{2i-2}^j\}_{j=0}^{i-2}$ are regular. After $2i-1$ iterations, one may define:

$$z_{2i-1} = P[y_{2i-1} | X_{2i-2}^0, \dots, X_{2i-2}^{i-2}, [1 \ 0] X_{2i-2}^{i-1}]$$

and $\{k_{2i-1}^j, q_{2i-1}^j\}_{j=0}^{i-2} \cup \left\{ [1 \ 0] k_{2i-1}^{i-1}, [1 \ 0] q_{2i-1}^{i-1} \right\}$ are well determined since $\{M_{2i-1}^j\}_{j=0}^{i-2}$ are regular but not M_{2i-1}^{i-1} .

From $i = n$ (i.e. $t = 2n$), all parameters k_t, q_t are well determined. \square

3.3.3 Completely finite recursive computation for finite Hankel rank of y .

Suppose that the state-space model of y is a finite-dimensional one. For example:

$$\begin{cases} x_{t+1} = Fx_t + Gw_{t+1} \\ y_t = Hx_t + v_t \end{cases} \quad (3.31)$$

where $x_t \in R^m$, $m \geq n$, w_t and v_t are independent white noises of variances Q and R respectively.

The computations of M_t^i and $\overline{(z_t^i)^2}$ are obviously still finitely recursive and are computed by the same formulas that in 3.3.1. The only difference lies in the computation of $\overline{y_t^2}$. It is achieved through:

$$\begin{cases} \overline{y_t^2} = \overline{Hx_t x_t^T H^T} + R \\ \overline{x_t(x_t)^T} = \overline{Fx_{t-1} x_{t-1}^T F^T} + GQG^T \end{cases} \quad (3.32)$$

Computation of $\overline{X_{t-1}^i y_t}$ is finitely recursive through the $R^{2 \times m}$ matrices $\left\{ \overline{X_{t-1}^i x_t^T} \right\}_{i=0}^{n-1}$

as follows: since $\{X_{t-1}^i\}_{i=0}^{n-1}$ only depends on $\{w_\tau, v_\tau\}_{\tau \leq t-1}$, one has:

$$\overline{X_{t-1}^i x_t^T} = \overline{X_{t-1}^i x_{t-1}^T F^T}, \forall i = 0 \dots n-1 \quad (3.33)$$

$$\left\{ \begin{array}{l} \overline{X_{t-1}^0 x_{t-1}^T} = \begin{bmatrix} \overline{H x_{t-1} x_{t-1}^T - z_{t-1}^0 x_{t-1}^T} \\ \overline{z_{t-1}^0 x_{t-1}^T} \end{bmatrix} \\ \overline{X_{t-1}^i x_{t-1}^T} = \overline{X_{t-2}^{i-1} x_{t-1}^T - q_{t-1}^{i-1} \overline{X_{t-2}^{i-1} x_{t-1}^T}}, \forall i = 1 \dots n-1 \end{array} \right. \quad (3.34)$$

$$\overline{z_{t-1}^i x_{t-1}^T} = \sum_{j=i}^{n-1} (k_{t-1}^j)^T \overline{X_{t-2}^j x_{t-1}^T}, \forall i = 0 \dots n-1 \quad (3.35)$$

Clearly, this recursion requires only the storage at each step of the finite dimensional matrix $\left\{ \overline{X_{t-1}^i x_{t-1}^T} \right\}$. The parameters $\{k_t^i, q_t^i\}_{i=0}^{n-1}$ are obtained with (3.21).

4. Normalized lattice form.

Recall that the variances of the forward residuals are decreasing with the orthogonalization process. As a consequence, the values of the parameters $\{q_t^i\}_{i=0}^{n-2}$ grow and their computation degenerates. Therefore, we now replace these residuals by normalized one.

Normalization of forward residuals is quite easy. One just considers the variables which are colinear to the previous one and have a unit variance. These quantities being scalar, there is a unique way to achieve it, that is:

$$\forall i = 0, \dots, n-1, u_t^i = \overline{z_t^i / (z_t^i)^2} \quad (4.1)$$

Recall that the backward residuals $\{X_t^i\}_{i=0}^{n-1}$ are vectors of $(L_p^2)^2$. So, the normalization procedure is not unique. If one looks for a matrix $\{Q_t^i\}_{i=0}^{n-1}$ of R^2 such that:

$$\forall i = 0, \dots, n-1, V_t^i = Q_t^i X_t^i \quad (4.2)$$

are normalized ($\overline{V_t^i (V_t^i)^T} = I$), the solution is undetermined. In fact, any product of a solution with an orthogonal matrix will succeed. However, in such a special situation, there is a natural way to make this normalization.

4.1. Backward residuals normalization.

$$\text{Define } Q_t^0 = (M_t^0)^{-1/2} = \begin{bmatrix} \overline{(y_t^2 - z_t^2)}^{-1/2} & 0 \\ 0 & \overline{(z_t^2)}^{-1/2} \end{bmatrix}.$$

It is clear that $V_t^0 = Q_t^0 X_t^0$ is then normalized. Let us examine the recursive equation of V_t^i for $i > 0$. According to system (3.19), one has:

$$V_t^i = Q_t^i (X_{t-1}^i - q_t^{i-1} u_t^{i-1} \|z_t^{i-1}\|) \quad (4.3)$$

where $\|z\| = (\overline{z^2})^{1/2}$ denotes the norm of z . But according to (3.22),

$$\begin{aligned} q_t^{i-1} \|z_t^{i-1}\| &= \overline{X_{t-1}^{i-1} y_t} / \|z_t^{i-1}\| \\ &= ((Q_{t-1}^{i-1})^{-1} \overline{V_{t-1}^{i-1} y_t}) / \|z_{t-1}^{i-1}\| \end{aligned} \quad (4.4)$$

Define then

$$p_t^i = \overline{V_{t-1}^i u_t^i} = \overline{V_{t-1}^i y_t} / \|z_t^i\|. \quad (4.5)$$

One may rewrite (4.3) as follows:

$$V_t^i = Q_t^i (Q_{t-1}^{i-1})^{-1} (V_{t-1}^{i-1} - p_t^{i-1} u_t^{i-1}) \quad (4.6)$$

We have the following proposition:

Proposition 5. *If $\forall i = 1, \dots, n-1$, $Q_t^i = (I - p_t^{i-1} (p_t^{i-1})^T)^{-1/2} Q_{t-1}^{i-1}$, then the family $\{V_t^i = Q_t^i X_t^i\}_{i=0}^{n-1}$ is orthonormalized. \square*

Proof: Recursively, the proposition is true for $i = 0$. Let us compute the covariance matrix of V_t^i . According to (4.3), one has:

$$\begin{aligned} \overline{V_t^i (V_t^i)^T} &= Q_t^i (Q_{t-1}^{i-1})^{-1} \left(\overline{V_{t-1}^{i-1} (V_{t-1}^{i-1})^T} + p_t^{i-1} \overline{(u_t^{i-1})^2} (p_t^{i-1})^T \right. \\ &\quad \left. - \overline{V_{t-1}^{i-1} u_t^{i-1}} (p_t^{i-1})^T - p_t^{i-1} \overline{u_t^{i-1} (V_{t-1}^{i-1})^T} \right) (Q_{t-1}^{i-1})^{-T} (Q_t^i)^T \end{aligned} \quad (4.7)$$

But if V_{t-1}^{i-1} and u_t^{i-1} are normalized, according to the definition of p_t^i , one has:

$$\overline{V_{t-1}^{i-1} (V_{t-1}^{i-1})^T} = Q_{t-1}^i (Q_{t-1}^{i-1})^{-1} (I - p_t^{i-1} (p_t^{i-1})^T) (Q_{t-1}^{i-1})^{-T} (Q_{t-1}^i)^T \quad (4.8)$$

Substituting above the value of Q_t^i , the result follows. \blacksquare

4.2. Forward residuals normalization.

One has to compute the variances of each residual. System (3.19) leads to the expression of the normalized forward residuals as:

$$u_t^i = (u_t^{i-1} \|z_t^{i-1}\| - (k_t^{i-1})^T (Q_{t-1}^{i-1})^{-1} V_{t-1}^{i-1}) / \|z_t^i\| \quad (4.9)$$

But according to (3.21) and the definition of p_t^i , one has $k_t^i = (Q_{t-1}^i)^T p_t^i \|z_t^i\|$. Equation (4.9) takes then the following form:

$$u_t^i = \frac{\|z_t^{i-1}\|}{\|z_t^i\|} (u_t^{i-1} - (p_t^{i-1})^T V_{t-1}^{i-1}) \quad (4.10)$$

Equation (3.22) leads to $\|z_t^i\|^2 = \sum_{j=i}^{n-1} (k_t^j)^T M_{t-1}^j k_t^j$. As a consequence, using the fact that

$$M_t^i = (Q_t^i)^{-1} (Q_t^i)^{-T}, \text{ one may derive: } \frac{\|z_t^{i-1}\|}{\|z_t^i\|} = (1 - (p_t^{i-1})^T p_t^{i-1})^{-1/2}$$

Proposition 6. *The normalized lattice form is therefore:*

$$\forall i = 1, \dots, n-1, \begin{cases} u_t^i = (1 - (p_t^{i-1})^T p_t^{i-1})^{-1/2} (u_t^{i-1} - (p_t^{i-1})^T V_{t-1}^{i-1}) \\ V_t^i = (I - p_t^{i-1} (p_t^{i-1})^T)^{-1/2} (V_{t-1}^{i-1} - p_t^{i-1} u_t^{i-1}) \end{cases}$$

with

$$\begin{cases} u_t^0 = z_t / \|z_t\| \\ V_t^0 = \begin{bmatrix} (\overline{y_t^2} - \overline{z_t^2})^{-1/2} & 0 \\ 0 & \|z_t\|^{-1} \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ z_t \end{bmatrix} \cdot \square \end{cases}$$

The figure 3 shows the lattice structure generated. The coefficients q^i and Q^i denote the normalization factors.

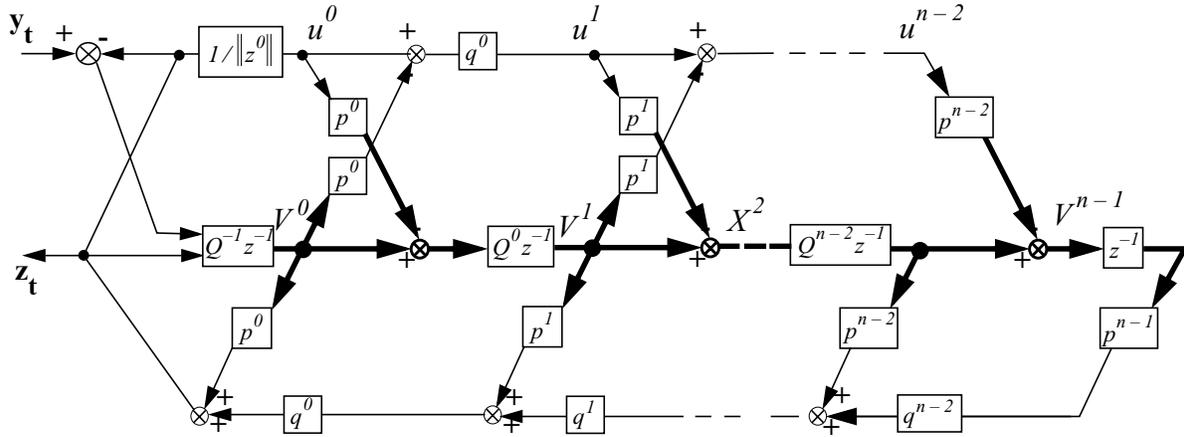


Figure 3 - Normalized Lattice Structure.

Remarks:

- Note that, in this case, only $2n+1$ parameters are used to derive the lattice form predictor. That is $\{p_t^i\}_{i=0}^{n-1}$ and $\|z_t^0\|$. It results from the fact that the forward and backward residuals are both computed from the same set of parameters.

- Square roots of matrices may be computed here from scalar square roots, using the following algebraic identity: $(I - p p^T)^{-1/2} = I + p p^T \left(\frac{(I - p^T p)^{-1/2} - I}{p^T p} \right)$, which derives from a straightforward calculation.

4.3. Computation of parameters.

4.3.1 The general algorithm.

It is clear that:

(4.11)

According to definition (4.5) and the first remark above, the computation of parameters requires only the quantities $\left\{ \overline{V_{t-1}^i y_t} \right\}_{i=0}^{n-t}$. As it was noticed in the unnormalized case, these quantities are also hereditary ones and call for $\left\{ \overline{V_{\tau}^i y_t}, \overline{u_{\tau}^i y_t}, \tau \leq t-1 \right\}_{i=0}^{n-t}$. In fact, suppose that $\{p_{\tau}^i, \tau \leq t-1\}_{i=0}^{n-t}$ and $\{\|z_{\tau}^0\|, \tau \leq t-1\}$ are known, according to (4.11), their computation is achieved by the following algorithm, with τ going from 1 to $t-1$:

- First, assuming that $\left\{ \overline{V_{\tau-1}^i y_t} \right\}_{i=0}^{n-t}$ is well defined, one computes:

$$\begin{cases} \overline{u_{\tau}^{n-1} y_t} = (p_{\tau}^{n-1})^T \overline{V_{\tau-1}^{n-1} y_t} \\ \overline{u_{\tau}^i y_t} = (p_{\tau}^i)^T \overline{V_{\tau-1}^i y_t} + \sqrt{1 - (p_{\tau}^i)^T p_{\tau}^i} \overline{u_{\tau}^{i+1} y_t}, \forall i = 0 \dots n-2 \end{cases} \quad (4.12)$$

- Then, one computes:

$$\begin{cases} \overline{V_{\tau-1}^o y_t} = \begin{bmatrix} (\overline{y_t y_{\tau}} - \overline{u_{\tau}^o y_t} \|z_{\tau}^0\|) / (\sqrt{\overline{y_t y_{\tau}} - \|z_{\tau}^0\|^2}) \\ \overline{u_{\tau}^o y_t} \end{bmatrix} \\ \overline{V_{\tau}^i y_t} = (I - p_{\tau}^{i-1} (p_{\tau}^{i-1})^T)^{-1/2} (\overline{V_{\tau-1}^{i-1} y_t} - p_{\tau}^{i-1} \overline{u_{\tau}^{i-1} y_t}), \forall i = 1 \dots n-1 \end{cases} \quad (4.13)$$

4.3.2 Additional formulas for singular initialization.

As in 3.3.2, the first $2n$ steps of computation of the parameters lead to a singular problem if one chooses to initialize the predictor in the absence of a priori information on the past. For a zero mean process, one may consider that at $t = 0$, one has:

$$\left\{ \begin{array}{l} u_0^i = 0, \forall i = 0 \dots n-1 \\ V_0^i = 0, \forall i = 1 \dots n-1 \\ V_0^0 = \begin{bmatrix} y_0 / \|y_0\| \\ 0 \end{bmatrix} \end{array} \right. \quad (4.14)$$

At the first step ($t = 1$), one can only project the observation on the first component of V_0^0 . As a consequence, according to Proposition 6, one has $u_1^0 = (p_1^0)^T V_0^0$ and the second component of p_1^0 is zero. But V_0^0 and u_1^0 being normalized, one has $\|p_1^0\| = 1$ and so $p_1^0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Therefore, $\forall 1 \leq i \leq n-1, V_1^i = 0$.

At the second step ($t = 2$), p_2^0 is defined and $\forall i = 1 \dots n-1, p_2^i = 0$. Remark that at this step, $\|p_2^0\| = 1$.

At third step ($t = 3$), only p_3^0 is well defined but $\|p_2^0\| = 1$ implies that the matrix $I - p_2^0(p_2^0)^T$ is singular. As a consequence, because $\|p_3^1\| = 1$, one has $p_3^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

And so on... Continuing this procedure, one find that:

Proposition 7. *For all $i \leq n$, computation of the predictor's parameters using definition (27) proceeds as follows:*

- after $2i-1$ iterations, $\{p_{2i-1}^j\}_{j=0}^{i-2}$ are well determined, $p_{2i-1}^{i-1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and

$\forall j = i \dots n-1, p_{2i-1}^j = 0$.

- after $2i$ iterations, $\{p_{2i}^j\}_{j=0}^{i-1}$ are well determined, $\|p_{2i}^{i-1}\| = 1$, and

$\forall j = i+1 \dots n-1, p_{2i}^j = 0$.

- after $2n$ iterations, all parameters are well determined.q

Note that after these first singular steps, p^{n-1} stands normalized. The figure 4 illustrates the algorithm obtained:

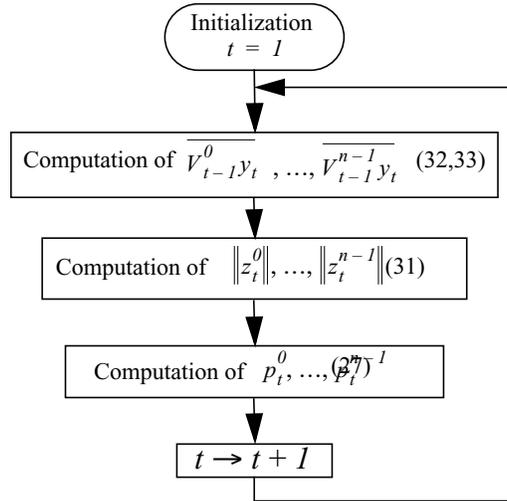


Figure 4 - Normalized Lattice Algorithm.

Obviously, the finite Hankel rank case leads to a finitely computation of parameters with the normalized lattice form, as in 3.3.3 for the unnormalized lattice form.

5. Examples of applications.

As is well known the optimal linear predictor of a finite Hankel rank signal process may be viewed as a particular ARMA model (with zero inside the unit circle), the innovation process \tilde{y} being a white noise input since the model and its whitening filter are inverse of one another. Otherwise, the linear predictor in reduced dimension makes the innovation process a “pseudo-white noise” in the sense that its orthogonalization property is limited to n delays. Its variance is defined here by the asymptotic value (in the stationary case) of $\overline{\tilde{y}^2} = \overline{y^2} - \overline{z_t^2}$. In this sense, the ARMA predictor yields the best n -th order approximate model of the signal in the L_p^2 metric. This will be exploited to illustrate the accuracy of the method presented here because one can compare easily the approximation of the model with the model itself.

The algorithm being particularly designed to the prediction of a signal which can only be represented by an infinite dimensional system, we deal here with the two following cases:

- The low-pass filter.
- A fractional system.

5.1. Test on finite dimensional linear systems.

An obvious preliminary step before any application is to test the algorithm in a situation analog to Kalman filtering (n -dimensional linear realization of the autocorrelation). The compared impulse responses gave identical values up to the tenth significant number.

The figures 5 and 6 illustrate the convergence accuracy of the algorithm for a 4-dimensional linear system. The module and phase of the model thus identified are plotted for the 20 first steps of computation.

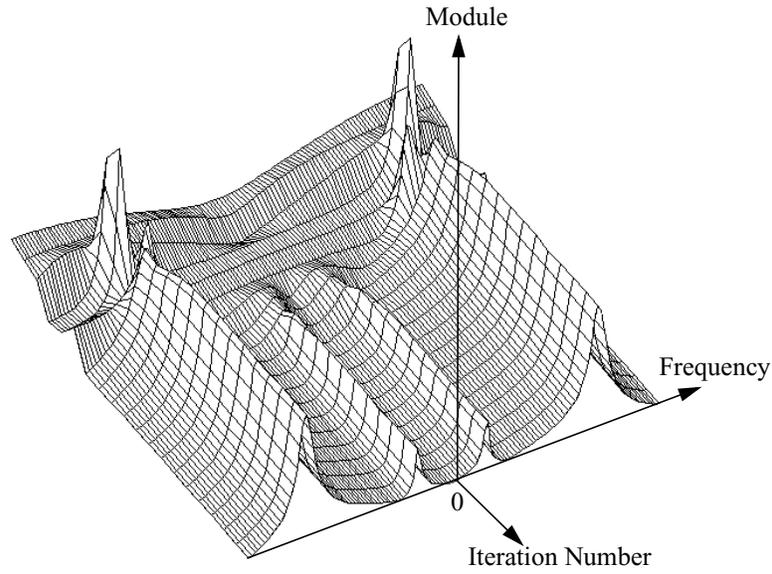


Figure 5 - Module of a 4-Dimensional System Identified.

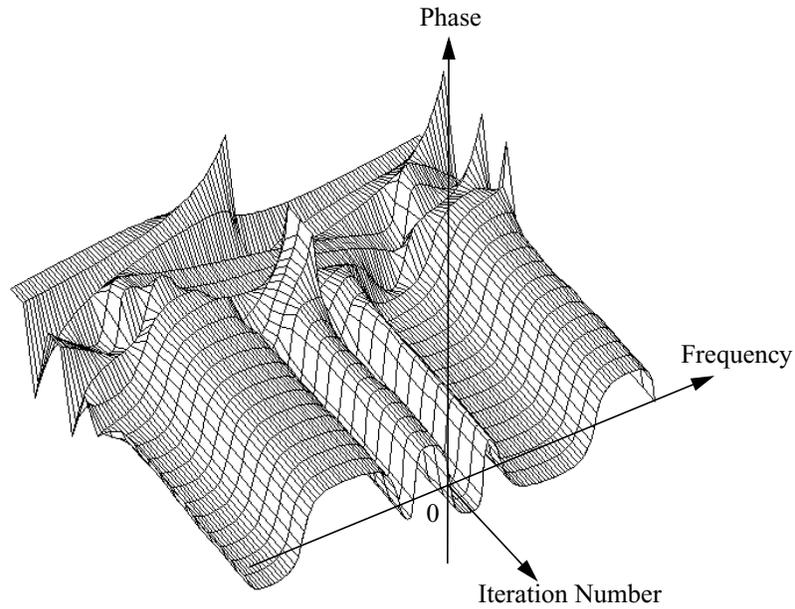


Figure 6 - Phase of a 4-Dimensional System Identified.

5.2. Low-pass filter.

Consider y_t a signal resulting of the sum of a random frequency uniformly distributed in the interval $[0, f]$ and a white noise of variance R . The autocorrelation of this signal is not realizable by a finite dimension system. Its value is obtained by the inverse Fourier transform of the function $I_{[-f, f]}$ i.e.:

$$\overline{y_t y_{t+\tau}} = \frac{\sin(4\pi f\tau)}{4\pi f\tau} + R\delta_0(\tau) \quad (5.1)$$

where $\delta_0(\tau)$ is the Kronecker discrete function.

Taking successively different dimensions n of the predictor ($n = 2, 4$ and 12 here), we obtained the optimal linear reduced predictor with its gain and phase characteristics of the figures 7 and 8 (with $f = 0.1 \times$ sampling-frequency, and $R = 0.1$, ratio signal-noise).

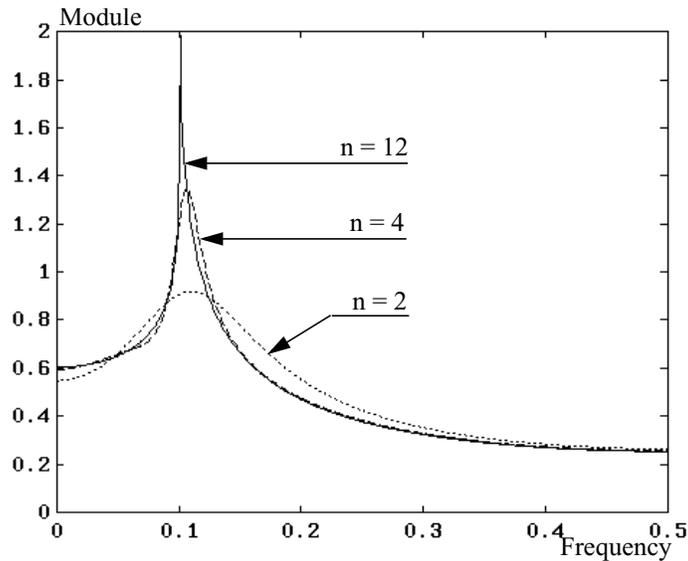


Figure 7 - Module of the Optimal Low-pass Predictor in Reduced Dimension

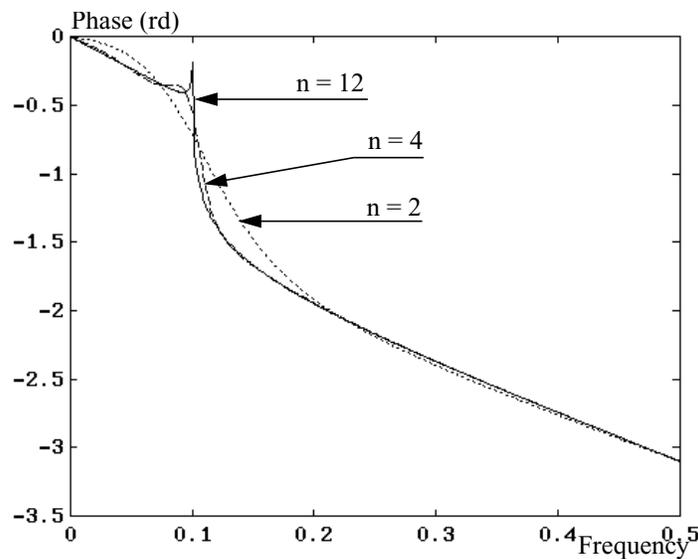


Figure 8 - Phase of the Optimal Low-pass Predictor in Reduced Dimension

By inverting the input and roles of y , \tilde{y} , the resulting model of the signal (best n -th order L_p^2 approximation has characteristics represented in the figures 9 and 10).

Comparisons with the well-known Chebychev filters are made in the figures 11, 12, 13 and 14 with $f = 0.25 \times \text{sampling-frequency}$, and $R = 0.01$ (ratio signal-noise). The module and phases of the identified spectrum have been computed for $n = 6$ and $n = 12$.

The figures 15 and 16 show the module and phase identified during the first 20 step of the algorithm for the 6-dimensional low-pass filter.

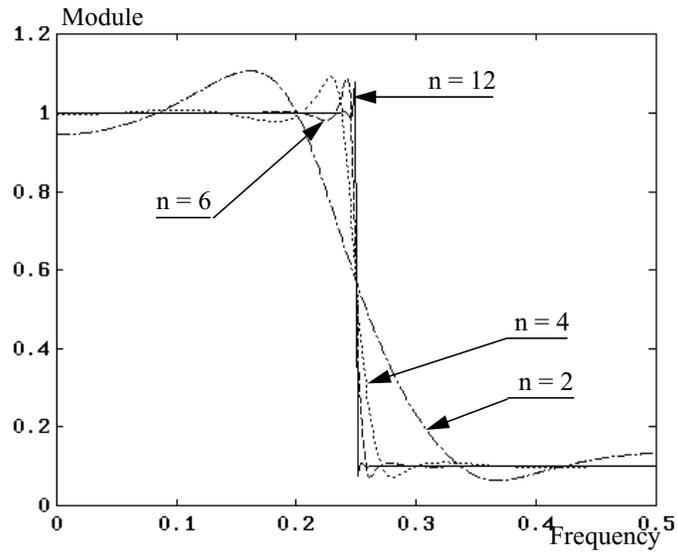


Figure 9 - Module of the Optimal Low-pass Filter in Reduced Dimension

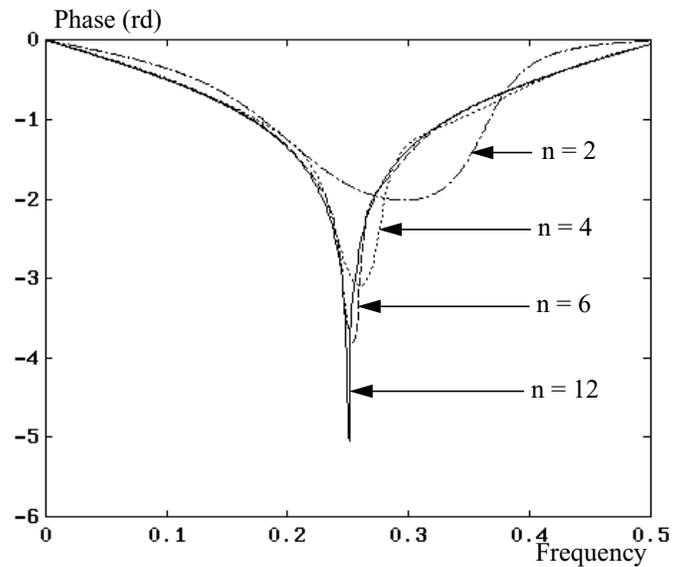


Figure 10 - Phase of the Optimal Low-pass Filter in Reduced Dimension

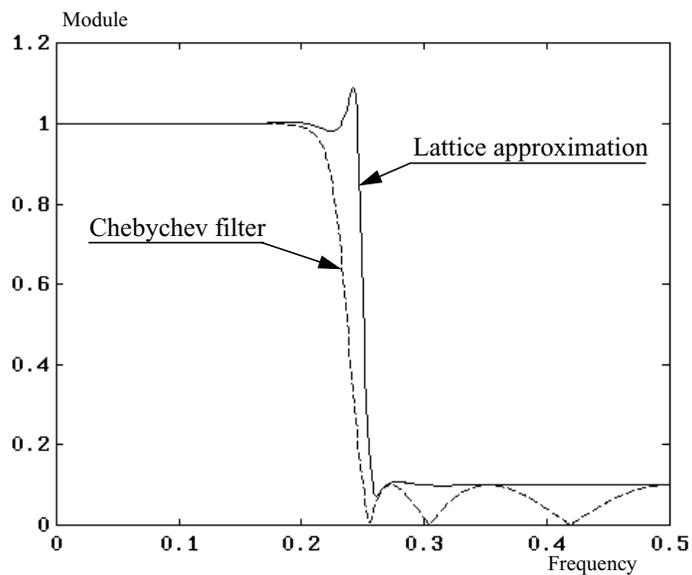


Figure 11 - Comparison with Chebyshev Filter. $n = 6$.

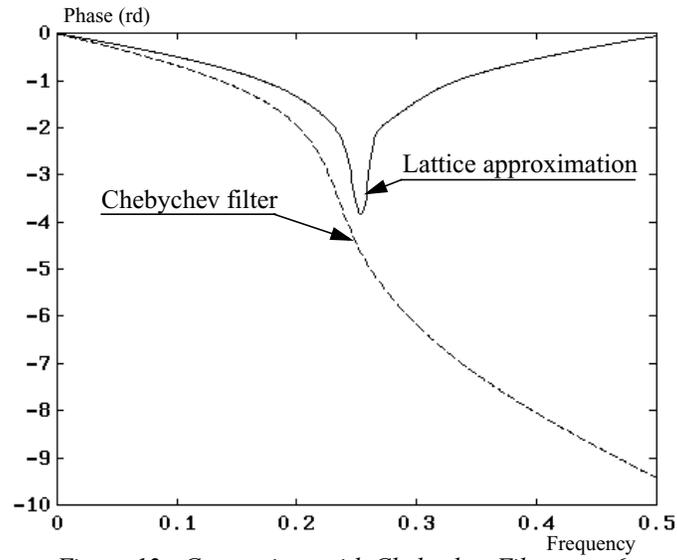


Figure 12 - Comparison with Chebychev Filter. $n = 6$.

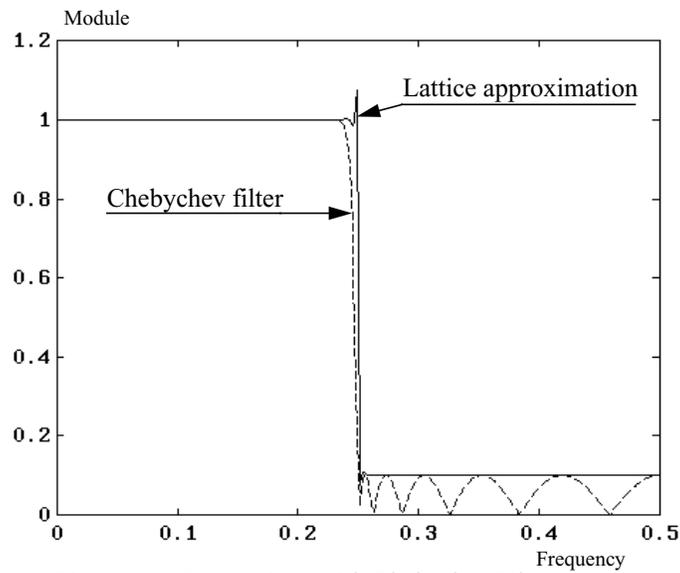


Figure 13 - Comparison with Chebychev Filter. $n = 12$.

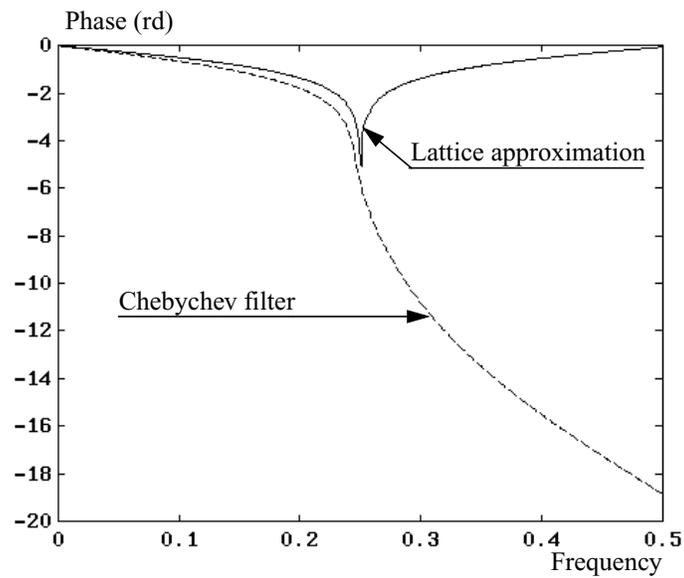


Figure 14 - Comparison with Chebychev Filter. $n = 12$.

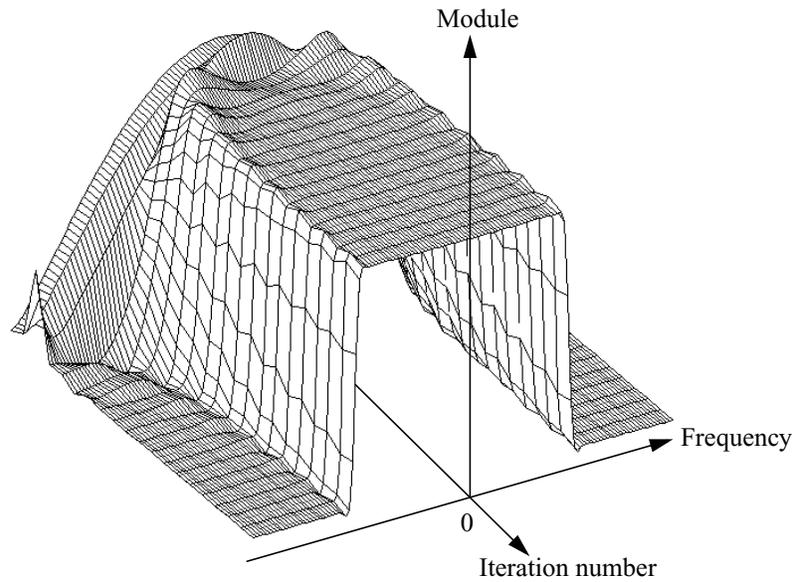


Figure 15 - Evolution of the module of Optimal Filter.

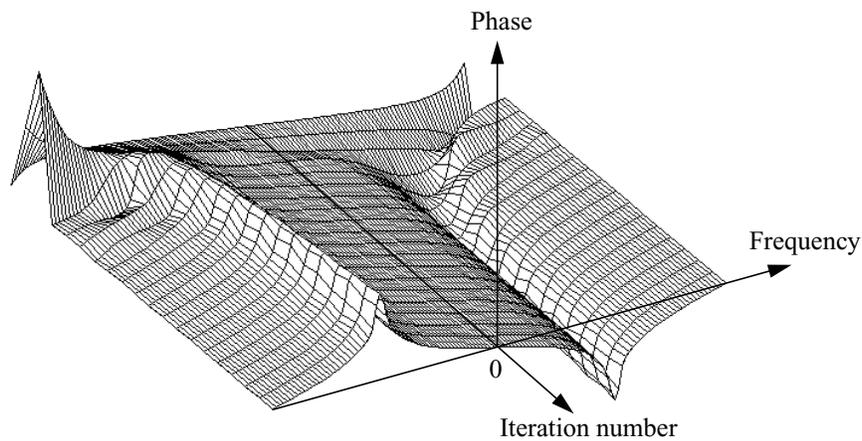


Figure 16 - Evolution of the phase of Optimal Filter.

5.3. Fractional system.

It can be shown that the use of fractional systems in control is of great interest with regard to the robustness of the control involved [14].

Consider an irrational impulse response of the form $F(z) = \frac{z^{-1}}{\sqrt{1-az^{-1}}}$. It is clear that

this system can't be realized by a finite system. The autocorrelation of the signal is the inverse z-transform of this spectrum. A series development of the square root leads to the following

autocorrelation: $\overline{y_t y_{t+\tau}} = \sum_{\sigma \geq 0} b_{\tau+\sigma} b_{\sigma}$ with $b_{\sigma} = \frac{1.3 \dots (2\sigma-1)}{2^{\sigma}(\sigma!)} a^{\sigma+1}$.

The figure 17 illustrates the approximation modelling yielded with $\tau = 0.999$ (near the continuous time situation). The power spectrum is plotted in the "loglog" coordinates to show the singular attenuation of 3dB by octave (or 10dB by decade) which characterizes this fractional system.

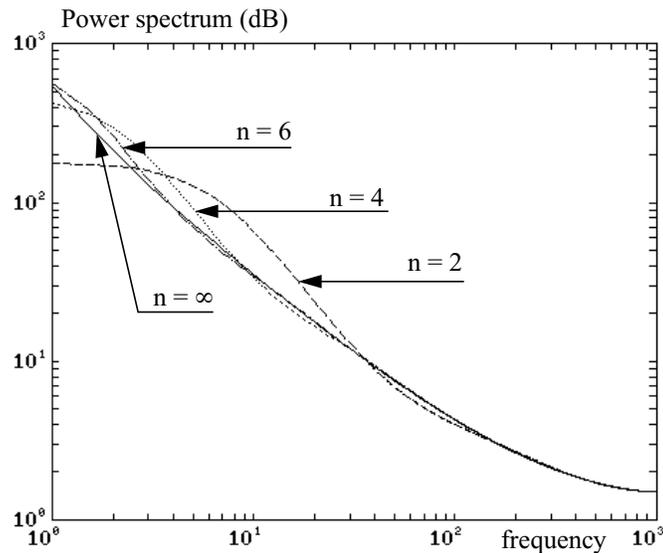


Figure 17 - Power Spectrum of the Optimal Approximation of a Fractional System.

6. Conclusion.

Let us recall as a conclusion that our problem deals with the computation of the exact time-varying ARMA predictor of order n for any square-integrable signal y_t , in the basis of its correlation function $C(\tau)$. It is well known that such a general problem has no finite dimensional solution, in term of a dynamical system with $C(\tau)$ as the input. It is however an essential objective to design the optimal predictor in a recursive fashion which:

- preserves exact optimality at each step (minimum variance of predictor error)
- uses a minimum amount of new computations when the number of points of the autocorrelation function increases.

Our paper derives an algorithm which fulfills these goals, at the expense of the hereditary computation of the coefficients, in terms of the autocorrelation function. The hereditary aspect of the algorithm makes it essentially different from other algorithms. It is crucial throughout the paper and accounts for:

- exact optimality, for each t , as a time-varying predictor of fixed order.
- convergence behavior, as a (unique) minimum-phase synthesis of L_p^2 approximations.

The usefulness of the algorithm derived is that it yields the exact solution of best L_p^2 -approximation in infinite-dimensional situations such as ideal filters, distributed parameters, etc. The lattice form involved by the orthogonalization process is of interest concerning numerical calculus and allows to deal with greater dimensions than with classical ARMA forms.

A final point concerns the extension of these ideas when applied to raw data. This is a quite different problem, since one has to replace theoretical minimum variance in L_p^2 by effective least square residuals which comes a maximum likelihood formulation (with a necessarily gaussian assumption for the problem to make sense). Contrarily to the kind of separation principle often invoked, the optimal result does not consist in replacing the autocorrelation terms by the experimental average of corresponding correlation products, since this is suboptimal in the adaptive transient. This is the object of a forthcoming paper which extends to real-time applications the approach introduced here.

7. References.

- [1] **A. Benveniste, C. Chaure**, "AR and ARMA identification algorithm of Levinson type : an innovations approach", IEEE Transactions on Automatic Control, vol. AC-26, n°6, 1981.
- [2] **A. Benveniste**, *Algorithme en treillis pour les modèles stationnaires.*, Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal, vol. 2, Editions du Centre National de la Recherche Scientifique, Paris, 1982.
- [3] **M. Morf, D. T. Lee, J. R. Nickolls, A. Vieira**, "A classification of algorithms for ARMA models and ladder realizations", Proceedings IEEE Int. Conf. Acoust. Speech, Signal Processing, Hartford, 1977.
- [4] **M. Morf, D. Lee**, "Recursive square root ladder estimation algorithm", Proceedings IEEE Int. Conf. Acoust. Speech, Signal Processing, Denver, 1980.
- [5] **D. T. Lee, B. Friedlander, M. Morf**, "Recursive ladder algorithms for ARMA modelling", IEEE Transactions on Automatic Control, vol. AC-27, pp. 753-764, 1992.
- [6] **T. Soderstrom, L. Ljung, I. Gustavson**, "A theoretical analysis of recursive identification methods", Automatica, vol. 14, pp. 231-244, 1978.
- [7] **G. Glentis, N. Kalouptsidis**, "Fast adaptive algorithm for multichannel filtering and system identification", IEEE Trans. on Acoust. Speech, Signal Processing, 1992.
- [8] **John J. Shynk**, "Adaptive IRR Filtering", IEEE Acoust. Speech, Signal Processing Magazine, vol. 6, pp. 4-21, 1989.
- [9] **F. Michaut**, *Méthodes adaptatives pour le signal.*, Edition Hermès, Paris, 1992.
- [10] **A. Benveniste, M. Metivier, P. Priouret**, *Algorithmes adaptatifs et approximations stochastiques*, Edition Masson, Paris, 1987.
- [11] **K.J. Asttröm, T. Söderström**, "Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model", IEEE Trans. on Automatic Control, vol AC-19, n°6, pp. 769-774, 1974.
- [12] **A. Monin, G. Salut**, "The optimal linear predictor in reduced dimension", Proceedings European Control Conference, Grenoble, France, 1991.
- [13] **M. Olivi, S. Steer**, " Approximation en norme L^2 des systèmes dynamiques", A.P.I.I. Journal, n°24, pp. 481-510, 1990.
- [14] **A. Oustaloup**, *Systèmes asservis linéaires d'ordre fractionnaire*, Edition Masson, Paris, 1983.