Lattice algorithm for adaptive stable identification and robust reconstruction of non stationary AR processes with missing observations

Rawad Zgheib, Gilles Fleury, Elisabeth Lahalle

Lattice algorithm for adaptive stable identification and robust reconstruction of non stationary AR processes with missing observations

Rawad Zgheib, Gilles Fleury, and Elisabeth Lahalle

***Abstract*—This paper deals with the problem of adaptive reconstruction and identification of non stationary AR processes with randomly missing observations. Existent methods use a direct realization of the filter. Therefore, the estimated parameters may not correspond to a stable all-pole filter. In addition, when the probability of missing a sample is high, existent methods may converge slowly or even fail to converge. We propose, at our knowledge, the first algorithm based on the lattice structure for online processing of signals with missing samples. It is an extension of the RLSL algorithm to the case of missing observations, using a Kalman filter for the prediction of missing samples. The estimated parameters guarantee the stability of the corresponding all-pole filter. In addition it is robust to high probabilities of missing a sample. It offers a fast parameter tracking even for high probabilities of missing a sample. It is compared to the Kalman pseudo linear RLS algorithm, an already proposed algorithm using a direct realization of the filter. The proposed algorithm shows better performance in reconstruction of audio signals.**

***Index Terms*—missing observations, lattice, identification, reconstruction, stability, robustness**

I. Introduction

In many practical situations, periodically sampled signals with missing observations may be encountered. This is the case, for example, of errors in transmission, or of temporary unavailability of measurement. It is also the case, in coding of audio signals or images, for compression purposes.

In many applications, such as digital communications or systems tracking, on line processing is necessary. We are interested here in on line reconstruction and identification of signals that can be modeled by an AR process with randomly missing observations. The loss of samples process follows a Bernoulli law independent of the signal.

Several methods, such as in [1], [2], [8], [12], [16], [18], have already been developed for on line processing of AR signals with missing data. They use at each time only the past available data. In [2], Bondon proposed an expression of the optimal predictor in the least mean square sense of AR processes with incomplete past. The predictor is adaptive, however, it is not recursive. In [1], the problem of recursive estimation of the output in missing-data situations is addressed. In [8], Isaksson derives a recursive EM algorithm for the identification in missing data situations. It is based on the off line version described in [7]. However, at each time, the inversion of a matrix is required to update the parameters. Therefore, it suffers from a high computational complexity. In [12], an LMS-like algorithm for simultaneous reconstruction and identification is developed. In [16], the pseudo-linear RLS algorithm, an adaptation of the RLS algorithm to the case of signals with missing data, is derived. However these two algorithms converge toward biased estimation of the parameters. In [18], we proposed an alternative to the LMS-like algorithm based on the incomplete past predictor [2]. It offers simultaneous optimal identification and reconstruction. However, this algorithm is quite time consuming. Recently [19], we proposed to use a Kalman filter for the prediction with the pseudo linear RLS algorithm. This algorithm is fast. In addition, it offers optimal reconstruction and identification in the least mean square sense.

All previously presented methods consider a direct realization of the linear infinite-impulse response (IIR) filter. Thus the estimated parameters may not correspond to a stable all-pole filter unless the poles of the filter are at each time constrained to be inside the unit circle (ex: Schur-Cohn factorization...). This might be quite time consuming. In addition, when the probability of missing a sample is large, existent methods may converge slowly or even fail to converge. Therefore we propose to identify the signal using the lattice structure of the filter. A lattice structure is characterized by the reflection coefficients k_i. The reflection coefficients estimated from data using methods such as proposed in [4], [13] guarantee the stability of the corresponding AR filter. The values of the AR parameters are uniquely determined by the reflection coefficients through the Durbin-Levinson algorithm [5], [11].

Many adaptive algorithms for the identification of lattice filters, such as the recursive least squares (RLSL) algorithms, have already been proposed [6], [10]. In [6], Friedlander presents a tutorial review of lattice structures and their use for adaptive prediction of time series. In [14], Makhoul *et al.* present a general method for adaptive updating of lattice coefficients in the linear predictive analysis of non stationary signals. The absolute value of the estimated reflection coefficients is always less than one thus guaranteeing the stability of the corresponding all-pole filter. In addition, it is simple and is adapted to non stationary signals. We are interested here in the extension of this method to the identification of signals with missing observations using a Kalman filter for the prediction of missing samples. At our knowledge, it is the first algorithm that permits on line processing of signals with missing samples using a lattice structure of the filter. The parameters thus

Department of Signal Processing and Electronic Systems, Supelec, Gif-sur-Yvette, France.
E-mail : firstname.lastname@supelec.fr

estimated guarantee the stability of the corresponding all-pole filter. In addition, the proposed algorithm is robust to large probabilities of losing a sample.

In the following, we begin by some basics about the lattice structure of the filter. A RLSL algorithm for adaptive stable identification of non stationary processes is presented. It is based on the reflection coefficients calculated in [14]. In section 3, the Kalman filter presented in [19] for the prediction of an AR process subject to missing samples is described. In section 4, an extension of the RLSL algorithm to the identification of signals with missing samples is presented. It uses for the prediction the Kalman filter presented in section 3. Finally, examples illustrate the performances of the new algorithm. It is compared to the Kalman pseudo-linear RLS algorithm [19]. Both algorithms are applied to the identification and reconstruction of non stationary AR processes and to the reconstruction of audio signals with random missing samples.

II. Adaptive identification of lattice filters

A. Lattice structure

It is preferred to identify a signal using a lattice structure of the filter because the stability and the robustness of the identified filter are thus guaranteed.

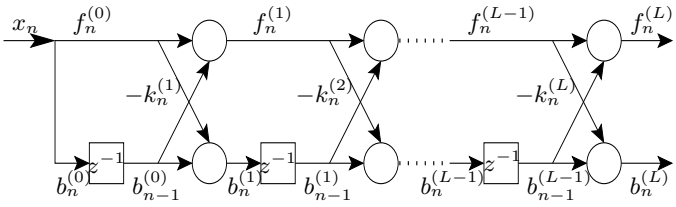


Fig. 1. Basic all-zero lattice filter.

The lattice structure is directly induced from the Levinson recursions [11]. Let x_n be the input signal, $\{a_i^{(L)}\}$ the optimal forward AR(L) predictor coefficients, $\{\cancel{a}_i^{(L)}\}$ the optimal backward predictor AR(L) coefficients. For an AR(L) process, the backward predictor coefficients are the same as the forward predictor ones but their order is reversed, e.g. $\cancel{a}_i^{(L)} = a_{L-i}^{(L)}$. At time n, for the predictor of order L, the forward prediction error is $f_n^{(L)} = x_n - \sum_{i=1}^{L} a_i^{(L)} x_{n-i}$, and the backward prediction error is $b_n^{(L)} = x_{n-L} - \sum_{i=1}^{L} a_i^{(L)} x_{n-L+i}$. From Fig. 1, at a stage l corresponding to the predictor of order l, the following recursions hold:

$$f_n^{(0)} = b_n^{(0)} = x_n, \tag{1a}$$

$$f_n^{(l)} = f_n^{(l-1)} - k_n^{(l)} b_{n-1}^{(l-1)}, \tag{1b}$$

$$b_n^{(l)} = b_{n-1}^{(l-1)} - k_n^{(l)} f_n^{(l-1)}. \tag{1c}$$

where $k_n^{(l)}$ is the reflection coefficient at stage l and time n. For an L^{th} predictor, $0 \leq l \leq L$, where $k_n^{(0)} = 1$. The L^{th} predictor coefficients at time n, $\{a_{i,n}^{(L)}\}$, are uniquely computed from the reflection coefficients using the following recursions:

- For $l = 0$ to L

$$a_{l,n}^{(l)} = k_n^{(l)}, \tag{2}$$

  - For $i = 1$ to $l-1$

$$a_{i,n}^{(l)} = a_{i,n}^{(l-1)} + k_n^{(l)} a_{l-i,n}^{(l-1)} \tag{3}$$

  - end
- end.

A lattice structure corresponds to a stable filter if all the reflection coefficients for $1 \leq l \leq L$ verify:

$$|k_n^{(l)}| < 1, \quad 1 \leq l \leq L \tag{4}$$

In addition to the advantages given above, the lattice has an important orthogonalization property: the 'decoupling' of consecutive stages of the lattice. Therefore, the L^{th} order least squares lattice prediction filter contains in it prediction filters of all lower orders. More precisely, the first l sections of the L^{th} order lattice predictor (Fig. 1) form the l^{th} order prediction filter. Due to this property, the global minimization at the lattice output is replaced by a sequence of local minimization problems, one at each stage of the lattice.

Many methods exist for the calculation of the reflection coefficients. They consider the minimization of the mean quadratic forward prediction error, the mean quadratic backward prediction error or a function of both errors. The minimization of the mean quadratic forward and backward prediction errors, called method forward and method backward respectively [13], leads to two reflection coefficients. In the stationary scalar case, both coefficients are equal. The main drawback of these methods is that the stability of the corresponding filter is not guaranteed. Itakura and Saito [9], proposed to calculate the reflection coefficients as the geometric mean of the coefficients obtained with the forward and backward methods. The reflection coefficient thus calculated verifies (4). Burg [4], proposed to calculate the reflection coefficient as the harmonic mean of the coefficients obtained with the forward and backward methods. The coefficient calculated by Burg guarantees the stability of the corresponding all-pole filter. Moreover, the coefficient thus calculated corresponds to the minimization of the sum of the quadratic forward and backward prediction errors [13]. Since we are interested in on line identification, we introduce in the following the recursive least square lattice algorithms.

B. RLSL algorithm

The RLSL algorithms are recursive both in time and order. They are numerically efficient requiring only $O(L)$ operations per time update, compared to $O(L^2)$ for the RLS algorithm using a direct realization of the filter. Besides the advantages of lattice structures, they exhibit excellent convergence behavior and fast parameter tracking capability. Therefore, they are adapted to the identification of non stationary processes [6], [10].

In [10], a geometrical approach is used to establish the time update recursions of the RLSL estimation algorithms. A square root normalized least squares lattice form algorithm is presented. The unnormalized RLSL algorithm [6], [10] minimizes simultaneously both the mean quadratic forward prediction error and backward prediction error. This algorithm leads to two reflection coefficients that does not guarantee the

stability of the corresponding filter. In the normalized form, all the variables are normalized to unit variance. A normalized lattice structure of the filter enjoys some advantages such as the stability of the identified model. Only one reflection coefficient is estimated for each stage l at each time n and it verifies (4).

In [14], Makhoul *et al.* proposed a general method for adaptive updating of reflection coefficients for each stage l and at each time n in the linear predictive analysis of non stationary signals. They determine the reflection coefficients by minimizing the weighted sum of the quadratic forward and backward prediction errors. To calculate $k_n^{(l)}$, they proposed to minimize:

$$E_n^{(l)} = \sum_{i=1}^{n} w_{n-i} \left(f_n^{(l)2} + b_n^{(l)2} \right). \tag{5}$$

where w_{n-i} is the weighting sequence. Since in a time varying situation, we are mainly interested in the most recent "history" of the signal, it is convenient to weight the errors with an exponential forgetting factor, e.g. $w_{n-i} = \lambda^{n-i}$ for $0 < \lambda \leq 1$.

Substituting (1) in (5) and minimizing with respect to $k_n^{(l)}$ leads to:

$$k_n^{(l)} = -\frac{2\sum_{i=1}^{n} \lambda^{n-i} f_n^{(l-1)} b_{n-1}^{(l-1)}}{\sum_{i=1}^{n} \lambda^{n-i} \left[f_n^{(l)2} + b_n^{(l)2} \right]} = -\frac{C_n^{(l)}}{D_n^{(l)}}. \tag{6}$$

where $f_n^{(l)}$ and $b_n^{(l)}$ are the forward and backward prediction errors for the stage l at time n. The reflection coefficient given by equation (6) is always guaranteed to verify the stability condition given by equation (4) [15]. It must be noted here that for $\lambda = 1$, the coefficient thus calculated is equal to Burg's coefficient [4]. This method is then an adaptation of Burg's method to the adaptive non stationary case. Makhoul *et al.* [14], [15] proposed an adaptive estimation of $k_n^{(l)}$ through the recursive computation on time of $C_n^{(l)}$ and $D_n^{(l)}$. An unnormalized RLSL algorithm is thus obtained by combining the adaptive estimation on time of $k_n^{(l)}$ and the recursive equations on order (1). The recursive equations, at time n, of the RLSL algorithm based on the reflection coefficient calculated in [14] are:

- Initialize for $l = 0$

$$f_n^{(0)} = b_n^{(0)} = x_n, \tag{7a}$$
$$k_n^{(0)} = 1, \tag{7b}$$

- For $l = 1$ to $\min(L, n)$

$$C_n^{(l)} = \lambda C_{n-1}^{(l)} + 2 f_n^{(l-1)} b_{n-1}^{(l-1)}, \tag{8a}$$
$$D_n^{(l)} = \lambda D_{n-1}^{(l)} + f_n^{(l-1)2} + b_{n-1}^{(l-1)2}, \tag{8b}$$
$$k_n^{(l)} = -\frac{C_n^{(l)}}{D_n^{(l)}}, \tag{8c}$$
$$f_n^{(l)} = f_n^{(l-1)} - k_n^{(l)} b_{n-1}^{(l-1)}, \tag{8d}$$
$$b_n^{(l)} = b_{n-1}^{(l-1)} - k_n^{(l)} f_n^{(l-1)}, \tag{8e}$$

- end.

where L is the filter order and $1 \leq n \leq N$, N being the length of the signal. At each time n, a recursion of order L is required for the update of the different reflection coefficients knowing all the prediction errors up to time $n-1$ and the current signal value. This algorithm exhibits similar performance as the normalized RLSL algorithm [10] and is less computationally expensive. We propose then to extend it to the identification of signals with missing samples, combined with the Kalman filter presented in [19] for the prediction.

III. Prediction using Kalman filter

In this section the Kalman filter used for the prediction of AR processes subject to missing samples is described.

A. State space representation.

Let $\{x_n\}$ be an AR process of order L with parameters $(a_i^{(L)})_{1 \leq i \leq L}$. It satisfies the following difference equation:

$$x_n = a_1^{(L)} x_{n-1} + \ldots + a_L^{(L)} x_{n-L} + \epsilon_n. \tag{9}$$

Where $\{\epsilon_n\}$ is the innovation process, a white noise of variance σ_ϵ^2. The loss process is modeled by an i.i.d binary variable $\{c_n\}$, $c_n = 1$ if x_n is available, otherwise $c_n = 0$. The probability to measure x_n is $P\{c_n = 1\} = 1 - q$. Let the observation process $\{y_n\}$ be defined as:

$$y_n = c_n x_n = \begin{cases} x_n & \text{if } x_n \text{ is available,} \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

Thus, y_n can be regarded as the measurement of x_n subject to missing data.

The AR process $\{x_n\}$ with missing observations admits the following state space representation [19]:

$$\begin{cases} \boldsymbol{x}_{n+1} &= A\, \boldsymbol{x}_n + \epsilon_n \left[1\, 0 \ldots 0\right]^\top \\ y_{n+1} &= \boldsymbol{c}_{n+1}^\top \boldsymbol{x}_{n+1} \end{cases} \tag{11}$$

where $A = \left[\begin{array}{ccc|c} a_1^{(L)} & \ldots & \ldots & a_L^{(L)} \\ \hline 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{array}\right]$ is a LxL matrix and $\boldsymbol{x}_n = \begin{bmatrix} x_n \\ \vdots \\ x_{n-L+1} \end{bmatrix}$, $\boldsymbol{c}_n = \begin{bmatrix} c_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ are L-vectors.

The sample to predict is at each time the first element of the state $\boldsymbol{x}_{n+1}$.

B. Kalman filter.

The predicted and filtered estimates of the state are denoted respectively $\hat{\boldsymbol{x}}_{n+1|n}$ and $\hat{\boldsymbol{x}}_{n+1|n+1}$. The corresponding estimation errors covariance matrices are noted respectively $P_{n+1|n}$ and $P_{n+1|n+1}$. K_{n+1} is the Kalman filter gain.

The Kalman filter equations are described in many books such as [3]. In our case, there is no disturbance in the observation. So, the Kalman filter equations reduce to:

$$P_{n+1|n} = A P_{n|n} A^\top + R_\epsilon, \tag{12}$$

where, $R_\epsilon = \sigma_\epsilon^2 \left[1\ 0 \dots 0\right]^\top \left[1\ 0 \dots 0\right]$.

$$K_{n+1} = \begin{cases} \frac{P_{n+1|n}\boldsymbol{c}_{n+1}}{\boldsymbol{c}_{n+1}^\top P_{n+1|n}\boldsymbol{c}_{n+1}} & \text{if } x_{n+1} \text{ is available} \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

$$P_{n+1|n+1} = (I_d - K_{n+1}\boldsymbol{c}_{n+1}^\top)P_{n+1|n}, \tag{14}$$

$$\hat{\boldsymbol{x}}_{n+1|n+1} = \hat{\boldsymbol{x}}_{n+1|n} + K_{n+1}(y_{n+1} - \hat{y}_{n+1|n}), \tag{15}$$

where,

$$\hat{\boldsymbol{x}}_{n+1|n} = A\hat{\boldsymbol{x}}_{n|n}, \tag{16}$$

and

$$\hat{y}_{n+1|n} = \boldsymbol{c}_{n+1}^\top \hat{\boldsymbol{x}}_{n+1|n}. \tag{17}$$

Due to the state space representation chosen, the previous Kalman filter can be simplified. Indeed, if, at time $n+1$, x_{n+1} is available, i.e., $c_{n+1} = 1$, the term $P_{n+1|n}\boldsymbol{c}_{n+1}$ is equal to the first column of $P_{n+1|n}$ and $(\boldsymbol{c}_{n+1}^\top P_{n+1|n}\boldsymbol{c}_{n+1})$ is a scalar equal to the first term of the matrix $P_{n+1|n}$. Consequently the Kalman gain K_{n+1} is calculated by dividing each element of the first column of $P_{n+1|n}$ by the first element of $P_{n+1|n}$. This requires L multiplications instead of $2L(L+1)$ multiplications as in the general case of a Kalman filter. In addition, we deduce that the first element of K_{n+1}, $K_{n+1}(1)$, is always equal to 1 if the data is available. Recalling the filter's equation system (15), the first equation of it gives that when a sample is available:

$$\begin{aligned} \hat{x}_{n+1|n+1} &= \hat{x}_{n+1|n} + K_{n+1}(1)\left(y_{n+1} - \hat{y}_{n+1|n}\right) \\ &= y_{n+1} = x_{n+1}. \end{aligned}$$

An observed sample is thus unchanged by the Kalman filter. Otherwise, if x_{n+1} is missing, i.e., $c_{n+1} = 0$, the predicted state $\hat{\boldsymbol{x}}_{n+1|n}$ is not filtered, $\hat{\boldsymbol{x}}_{n+1|n+1} = \hat{\boldsymbol{x}}_{n+1|n}$. It must be noted here that if the sample x_{n+1} is missing, its estimation updated during the next $L-1$ steps, is different from its prediction provided by the Kalman filter at time $n+1$, i.e. $\hat{x}_{n+1|n+t} \neq \hat{x}_{n+1|n+1}$ for $1 \leq t \leq L-1$. Indeed, the state is constituted of the last L samples or their predictions if they are missing. When a sample is available, the state is filtered through the equation (15). As a result, some terms are added to the predictions of the missing samples contained in the state. Hence, the prediction of a missing sample, updated during the $L-1$ next steps, is formulated as :

$$\hat{x}_{n+1|n+t} = \hat{x}_{n+1|n} + \sum_{i=1}^{t} K_{n+i}(i)\left(y_{n+i} - \hat{y}_{n+i|n+i-1}\right). \tag{18}$$

where $1 \leq t \leq L-1$. In real time applications, $\{z_n\}$, the reconstruction of the process $\{x_n\}$ subject to missing data, is defined as:

$$z_{n+1} = \hat{x}_{n+1|n+1} = \begin{cases} x_{n+1} & \text{if } x_{n+1} \text{ is available,} \\ \hat{x}_{n+1|n} & \text{otherwise,} \end{cases} \tag{19}$$

IV. Combined RLSL algorithm and Kalman filter

In this section, a new algorithm for online identification and reconstruction of non stationary processes subject to missing observations using the lattice structure of the filter is described. It is an extension of the RLSL algorithm given by equations (7) to the identification of signals with missing samples, combined with the Kalman filter mentioned in section III-B for the prediction of missing samples. The model thus identified guarantees at each time the stability of the corresponding filter.

Referring to equations (8a), (8b) and (8c) and to equations (1), the update of the reflection coefficients and the computation of the prediction errors uses the backward prediction errors calculated at the previous time. We deduce recursively that in order to calculate the reflection coefficients at time n, all the lattice filter variables must have been calculated at all previous times. However, the recursive equations on order (1) used for the computation of the prediction errors, are initialized at each time using the value of the signal (equation (1a)). When a sample is missing, an optimal estimation in the least mean square sense of the prediction errors is then required, since these values are used at the next time.

Thus the cost function minimized by the proposed algorithm is an estimation of the one in equation (5). It is given at time n by:

$$\hat{J}_n^{(l)} = \sum_{i=1}^{n} w_{n-i}\left(\hat{f}_n^{(l)2} + \hat{b}_n^{(l)2}\right). \tag{20}$$

where $\hat{f}_n^{(l)}$ and $\hat{b}_n^{(l)}$ are the estimates of $f_n^{(l)}$ and $b_n^{(l)}$ respectively.

A. Optimal estimation of the prediction errors

The forward and backward prediction errors are defined as:

$$f_n^{(0)} = b_n^{(0)} = x_n, \tag{21a}$$

$$f_n^{(l)} = x_n - \sum_{i=1}^{l} a_{i,n}^{(l)} x_{n-i}, \tag{21b}$$

$$b_n^{(l)} = x_{n-l} - \sum_{i=1}^{l} a_{i,n}^{(l)} x_{n-l+i}, \tag{21c}$$

Using the missing observations theorem [17], the optimal estimation of the prediction errors in the least mean square sense are given by the following equations:

$$\hat{f}_n^{(0)} = \hat{b}_n^{(0)} = \hat{x}_{n|n}, \tag{22a}$$

$$\hat{f}_n^{(l)} = \hat{x}_{n|n} - \sum_{i=1}^{l} a_{i,n}^{(l)} \hat{x}_{n-i|n}, \tag{22b}$$

$$\hat{b}_n^{(l)} = \hat{x}_{n-l|n} - \sum_{i=1}^{l} a_{i,n}^{(l)} \hat{x}_{n-l+i|n}, \tag{22c}$$

where, $\hat{x}_{n-i|n}$ is the optimal prediction in the least mean square sense of x_{n-i} knowing only the available samples until time n. The equations giving the optimal estimation of the prediction errors are analog to the definition equations (21). The difference is that the prediction errors and the values of the signal are replaced by their estimations. Hence, in the case of missing observations, the recursive equations of the RLSL algorithm at a time n become:

- Initialize for $l = 0$

$$\hat{f}_n^{(0)} = \hat{b}_n^{(0)} = \hat{x}_{n|n}, \tag{23a}$$

$$\hat{k}_n^{(0)} = 1, \tag{23b}$$

- For $l = 1$ to $\min(L, n)$

$$C_n^{(l)} = \lambda C_{n-1}^{(l)} + 2\hat{f}_n^{(l-1)}\hat{b}_{n-1}^{(l-1)}, \tag{24a}$$
$$D_n^{(l)} = \lambda D_{n-1}^{(l)} + \hat{f}_n^{(l-1)2} + \hat{b}_{n-1}^{(l-1)2}, \tag{24b}$$
$$\hat{k}_n^{(l)} = -\frac{C_n^{(l)}}{D_n^{(l)}}, \tag{24c}$$
$$\hat{f}_n^{(l)} = \hat{f}_n^{(l-1)} - \hat{k}_n^{(l)}\hat{b}_{n-1}^{(l-1)}, \tag{24d}$$
$$\hat{b}_n^{(l)} = \hat{b}_{n-1}^{(l-1)} - \hat{k}_n^{(l)}\hat{f}_n^{(l-1)}, \tag{24e}$$

- end.

We propose to use the Kalman filter presented in section III-B for an optimal prediction of the signal. The Kalman filter uses the estimated parameters of the filter direct realization. Since we identify the signal in the lattice filter form, therefore, at each time, a transformation of the reflection coefficients to the AR parameters through equations (2) is required.

Recall here that some terms are added to the prediction of a missing sample thanks to the Kalman filtering of the state when a sample is available during the $L-1$ following time instants. Hence an optimal prediction of a missing sample at time n, $\hat{x}_{n|n}$, is updated during the $L-1$ subsequent steps if some samples are available. In real time applications, the updated estimate of a missing sample, $\hat{x}_{n|n+t}$, is used only for the prediction of subsequent samples. The signal is reconstructed using its prediction provided by the Kalman filter at time n, $\hat{x}_{n|n}$. On the other hand, the forward and backward prediction errors estimated using the prediction of a missing sample, are used in the subsequent time step. Therefore, as for the prediction of the sample, their value must also be updated in the subsequent step. In order to avoid computing their values many times, we propose the following:

If at time n a sample is missing, it is predicted using the Kalman filter. However the prediction errors and the reflection coefficient are not computed until the prediction of the sample is updated, i.e. until a new sample is available. When a sample is available at time $n+t$, the state is updated thanks to the Kalman filter. It is now convenient to calculate all the lattice variables corresponding to the missing samples between the last available sample, at time $n-h$ and the current time $n+t$.

It must be noted here that if the time step between two available samples is greater than L ($t-h+1 > L$), only the predictions of the last $L-1$ missing samples (those contained within the state) are updated. The predictions of the preceding missing samples (from times $n-h+1$ till time $n+t-L$) are not updated. However, the lattice filter variables are not yet calculated at these times. But, in order to calculate the reflection coefficient at time $n+t$, all lattice filter variables must have been calculated for all previous times. Therefore, it is required to calculate the lattice filter variables from time $n-h+1$ till time $n+t-L$ even if there are no new information.

B. Proposed algorithm

The proposed algorithm can be summarised as follows. At time $n+1$, the first line of the matrix A is replaced by $\hat{\boldsymbol{a}}_n^{(L)\top}$, the vector of the parameters estimated at time n. The matrix is then named A_{n+1}.

$$A_{n+1} = \left[\begin{array}{ccc|c} \hat{a}_{1,n}^{(L)} & \dots & \dots & \hat{a}_{L,n}^{(L)} \\ \hline 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{array}\right], \tag{25}$$
$$P_{n+1|n} = A_{n+1}P_{n|n}A_{n+1}^\top + R_\epsilon,$$
$$\hat{\boldsymbol{x}}_{n+1|n} = A_{n+1}\hat{\boldsymbol{x}}_{n|n}$$
$$\hat{y}_{n+1|n} = c_{n+1}\hat{x}_{n+1|n}$$

a) If x_{n+1} is available, i.e. $c_{n+1} = 1$:

$$K_{n+1} = P_{n+1|n}\boldsymbol{c}_{n+1}(\boldsymbol{c}_{n+1}^\top P_{n+1|n}\boldsymbol{c}_{n+1})^{-1}, \tag{26a}$$
$$P_{n+1|n+1} = (I_d - K_{n+1}\boldsymbol{c}_{n+1}^\top)P_{n+1|n}, \tag{26b}$$
$$\hat{\boldsymbol{x}}_{n+1|n+1} = \hat{\boldsymbol{x}}_{n+1|n} + K_{n+1}(y_{n+1} - \hat{y}_{n+1|n}) \tag{26c}$$

The predictions of the previous missing data up to time $n-L+1$ are updated thanks to the filtering of the state in equation (26c). It is convenient now to calculate all the variables of the lattice filter since the last available observation at time $n-h$, where $h \geq 0$ depends on the observation pattern. At each time t, for $n-h+1 \leq t \leq n+1$, the recursive equations of the RLSL algorithm given by (28) are applied to estimate the different reflection coefficients $\hat{k}_t^{(l)}$ and prediction errors $\hat{f}_t^{(l)}, \hat{b}_t^{(l)}$ for $1 \leq l \leq L$. The values of the forward and backward prediction errors are initialized using the updated estimates of the missing samples (those contained within the filtered state $\hat{\boldsymbol{x}}_{n+1|n+1}$), i.e. $\hat{f}_t^{(0)} = \hat{b}_t^{(0)} = \hat{x}_{t|n+1}$.

Hence,

- For $t = n-h+1$ to $n+1$
  - Initialize for $l = 0$

$$\hat{f}_t^{(0)} = \hat{b}_t^{(0)} = \hat{x}_{t|n+1}, \tag{27a}$$
$$\hat{k}_t^{(0)} = 1, \tag{27b}$$

  - For $l = 1$ to $\min(L, n)$

$$C_t^{(l)} = \lambda C_{t-1}^{(l)} + 2\hat{f}^{(l-1)_t}\hat{b}_{t-1}^{(l-1)}, \tag{28a}$$
$$D_t^{(l)} = \lambda D_{t-1}^{(l)} + \hat{f}_t^{(l-1)2} + \hat{b}_{t-1}^{(l-1)2}, \tag{28b}$$
$$\hat{k}_t^{(l)} = -\frac{C_t^{(l)}}{D_t^{(l)}}, \tag{28c}$$
$$\hat{f}_t^{(l)} = \hat{f}_t^{(l-1)} - \hat{k}_t^{(l)}\hat{b}_{t-1}^{(l-1)}, \tag{28d}$$
$$\hat{b}_t^{(l)} = \hat{b}_{t-1}^{(l-1)} - \hat{k}_t^{(l)}\hat{f}_t^{(l-1)}, \tag{28e}$$

  - end
- end.

The AR parameters at time $n+1$, $(\hat{a}_{i,n+1}^{(L)})_{1\leq i\leq L}$, are deduced from the reflection coefficients $(\hat{k}_{n+1}^{(l)})_{1\leq i\leq L}$ using the Durbin Levinson recursions (equations (2)).

b) However if x_{n+1} is absent, $c_{n+1} = 0$: the predicted state, $\hat{\boldsymbol{x}}_{n+1|n}$, is not filtered by the Kalman filter, and the parameters are not updated since the reflection coefficients $(\hat{k}_{n+1}^{(l)})_{1\leq l\leq L}$ are not yet calculated,

$$K_{n+1} = 0, \tag{29a}$$
$$P_{n+1|n+1} = P_{n+1|n}, \tag{29b}$$
$$\hat{\boldsymbol{x}}_{n+1|n+1} = \hat{\boldsymbol{x}}_{n+1|n}, \tag{29c}$$
$$\hat{\boldsymbol{a}}_{n+1}^{(L)} = \hat{\boldsymbol{a}}_{n}^{(L)}. \tag{29d}$$

A flowchart describing the algorithm is presented in appendices A and B.

The cost function minimized by the Kalman pseudo-linear RLS algorithm [19] is the weighted mean of available samples prediction errors. Thus, the parameters are not updated by the pseudo-linear RLS algorithm if a data is missing. Indeed, the update of the parameters using RLS algorithms is proportional to the prediction error that cannot be calculated when a sample is missing. However, the cost function minimized by the proposed algorithm is the weighted mean of all quadratic prediction errors. When a sample is missing, the prediction error can not be calculated, it is replaced by its estimation. Indeed, recall that in order to update the reflection coefficients at a time n, the lattice filter variables must have been calculated at all previous times. Therefore, using the proposed algorithm, the lattice filter variables are estimated at all times even when a sample is missing. This suggests that its speed of convergence does not depend on the probability of losing samples. Indeed simulations show that it presents an excellent convergence behavior and have fast parameter tracking capability even for a large probability of missing a sample.

The computational complexity of both algorithms has been calculated. Since they both use the same Kalman filter for the reconstruction, their computational complexity differ by the identification algorithm used. For both algorithms, the computational complexity is found to be $O((1-q)NL^2)$, where q is the bernoulli's probability of losing a sample, N is the size of the signal and L the order of the AR model. Simulation of both algorithms implemented on Matlab showed close computation times. For an AR model of small order, the proposed algorithm is faster than the Kalman pseudo-linear RLS algorithm. However, it is slightly slower for high order models such as for speech signals.

V. Simulations

In this section, the proposed algorithm and the Kalman pseudo-linear RLS algorithm [19] are compared through three experiments. The test signal used in the first experiment is a non stationary AR process of order 2, with a high probability of missing samples. In the second experiment, both algorithms are applied to the reconstruction of a speech signal for different probabilities of missing samples. The test signal used in the third experiment is a music played on piano. For both experiments, their performance are compared in terms of the SNR given by:

$$(SNR) = 10log_{10}\left(\frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i - z_i)^2}\right). \tag{30}$$

A. Experiment 1

The test signal used is a non stationary AR(2) process generated over 15.10^3 samples. The parameters of the signal are $[1.5, -1]$ for the first 5.10^3 samples, $[-0.5, -1]$ for the next 5.10^3 samples and $[0.5, -1]$ for the last samples. The Bernoulli's probability of sample loss is $q = 0.7$. The forgetting factor used is $\lambda = 0.99$.

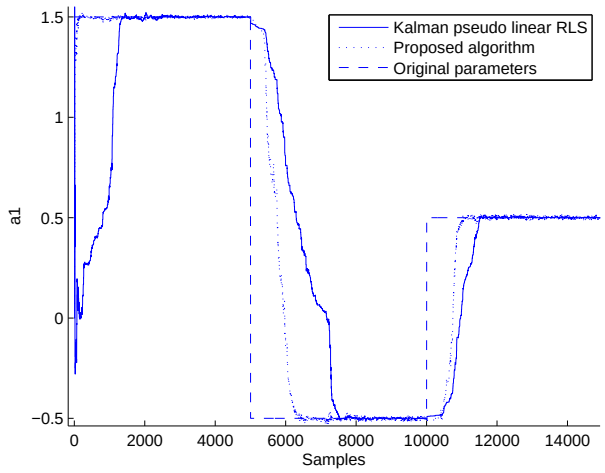


Fig. 2. Estimation of a_1 for the experiment 1.

Figure 2 shows the estimation of parameter a_1 with both algorithms. It shows that the proposed algorithm converges faster than the other one even for a large probability of missing a sample. Smaller values of the forgetting factor offer a faster convergence at the expense of a higher variance in the parameters estimation. For a high probability of losing samples, the variance can become arbitrarily large and may lead to a divergence using the Kalman pseudo-linear RLS algorithm. Simulations show that the speed of convergence of the Kalman pseudo-linear RLS algorithm depend highly on the observation pattern and on the probability of missing samples. Indeed this experiment was repeated for 100 regenerations of the observation pattern. The SNR obtained using the proposed algorithm is only one time less than 10 dB, however it is 36 times using the Kalman pseudo-linear RLS algorithm. This is due to a slow convergence or even to a divergence. Indeed, the Kalman pseudo-linear RLS failed to converge three times.

Moreover, the figure 2 seems to show that the proposed algorithm converges toward unbiased estimation of the parameters. However, due to the minimization of the prediction errors when samples are missing, the identification using this algorithm is biased. The bias decreases to zero as the parameter a_2 comes close to -1. Such signals correspond to resonant filters.

These characteristics suggest that the proposed algorithm might have good performance for speech signals reconstruction.

B. Experiment 2

The test signal used is the English sentence 'Mary had a little lamb, its fleece was white as snow' sampled at $F_s = 8$

kHz. This signal is reconstructed using both algorithms for different probabilities of missing a sample. For each probability, the observation pattern is simulated 1000 times. For each algorithm, the mean value and the variance of the SNR are calculated over the different simulations of the observation pattern for each probability of missing a sample. For both algorithms, the forgetting factor used is $\lambda = 0.99$. This value is empirical. Simulations show that both algorithms offer good performance in terms of reconstruction for this value of λ. The speech signal is modeled by an AR process of order 12.

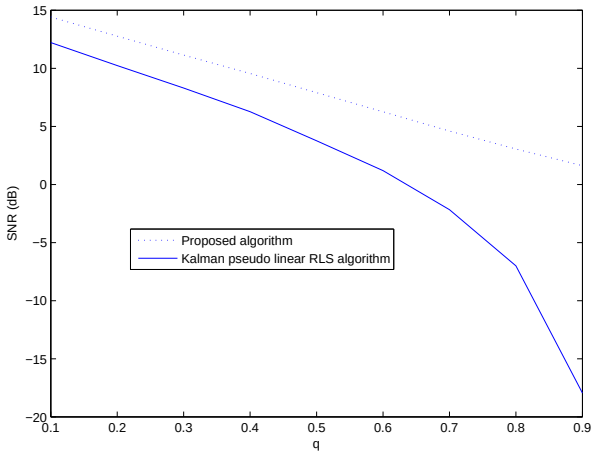


Fig. 3. Average SNR in terms of the probability of missing a sample for experiment 2

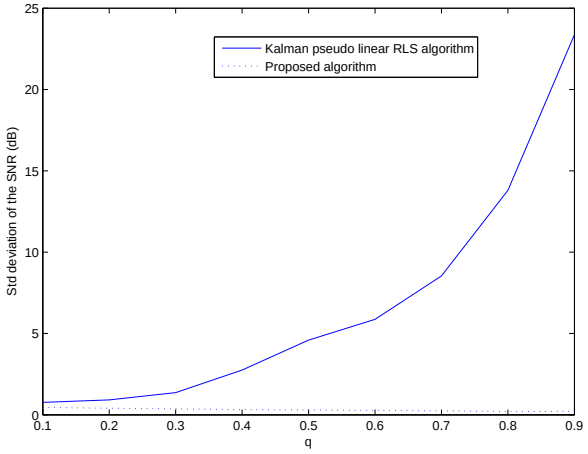


Fig. 4. Standard deviation of the SNR in terms of the probability of missing a sample for experiment 2

Figures 3 and 4 show respectively the mean value and the standard deviation of the SNR for both algorithms in terms of the Bernoulli's probability of missing a sample.

Figure 3 shows that on average the proposed algorithm offers better performance than the other one in terms of reconstruction error and particularly for high probabilities of missing a sample. The average SNR obtained using the proposed algorithm decreases linearly with the probability of sample loss. However, it decreases exponentially using the Kalman pseudo linear RLS algorithm. Indeed, when 60% of samples are missing, the SNR obtained using the proposed algorithm is of 3 dB larger than the one obtained using the other algorithm. When 85% of the samples are missing, this difference is of 9 dB. It must be noted here that the reconstructed signal obtained using the proposed algorithm is still understandable even when 85% of the samples are missing.

Moreover, referring to figure 4, the standard deviation of the SNR obtained with the proposed algorithm is small comparing to the one obtained using the other algorithm. In addition, it is almost constant for any probability of missing a sample. This shows that the proposed algorithm is robust to a large probability of missing a sample and to the observation pattern. However, the standard deviation of the Kalman pseudo-linear RLS algorithm increases exponentially with the probability of missing samples. The Kalman pseudo-linear RLS algorithm is then highly dependent on the observation pattern for high probabilities of missing a sample. Indeed, for a speech signal, the AR parameters are stationary for a small period of time, therefore when a large number of samples are missing, the Kalman pseudo linear RLS algorithm may fail to track the variations of the parameters.

C. Experiment 3

In this experiment, the test signal used is a musical signal. It is the toccata in C min of J. Bach played on piano by Glen Gould. In musical signals, the sounds are always voiced, corresponding to resonant filters. This suggests that it must have also a good performance for musical signals reconstruction. The same experiment as in experiment 2 is repeated for the musical signal.

Table 1 illustrates the mean value and the standard deviation of the SNR for a samples loss of 60%, 70%, 80% and 90%. M_1 and M_2 are the mean values of the SNR obtained using the proposed algorithm and the other one respectively. S_1 and S_2 are the corresponding standard deviations.

TABLE I
COMPARISON OF BOTH ALGORITHMS FOR EXPERIMENT 3

M_1 and S_1 are the mean value and the standard deviation of the SNR obtained using the proposed algorithm, M_2 and S_2 are obtained using the other one.

q	0.6	0.7	0.8	0.9
$M_1 \pm S_1$ (dB)	20 $\pm$ 0.16	15 $\pm$ 0.14	10.27 $\pm$ 0.11	5.4 $\pm$ 0.35
$M_2 \pm S_2$ (dB)	16 $\pm$ 0.41	11 $\pm$ 2.46	-0.18 $\pm$ 12.36	-38 $\pm$ 31.8

Referring to table I, the proposed algorithm shows, as expected, good performance in reconstruction of musical signals. As for the previous experiment, the average SNR obtained using the proposed algorithm decreases linearly with the probability of sample loss. However, it decreases exponentially using the Kalman pseudo linear RLS algorithm.

Listening tests have revealed that the prediction errors, using the proposed algorithm, are practically inaudible for this test and for many other signals. However, for high probabilities of missing samples, the errors of reconstruction using the Kalman pseudo linear RLS algorithm are pulse shaped and the voice is

sometimes deformed. This is usually the case when the model parameters are changed.

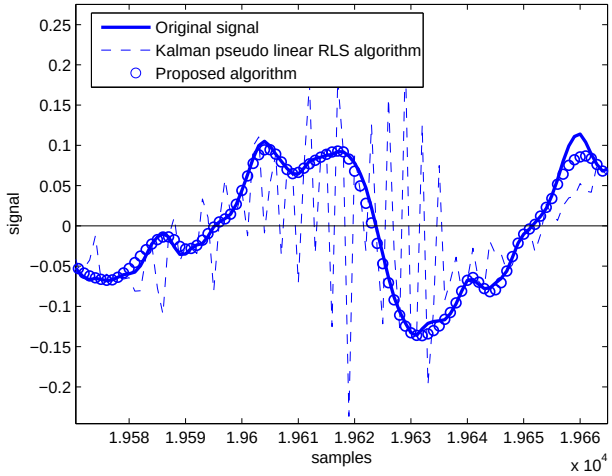


Fig. 5. Reconstruction of the signal using both algorithms for experiment 3 at the beginning of a new note.

Figure 5 shows the reconstruction of the signal using both algorithms at the beginning of a musical note. The beginning of a new musical note corresponds to a jump in the AR parameters. For a high probability of missing samples, the Kalman pseudo linear RLS algorithm fails to adapt quickly to the jump in the estimation of the parameters. This may result in a transient instability leading to transient strong oscillations in the reconstructed signal as shown in figure 5. However, the proposed algorithm presents a fast and stable adaptation to the change in the parameters. The signal reconstructed using the proposed algorithm is close to the original one.

In the case of noisy data, an observation noise must be considered in the state space representation of the process. Generally, the observation noise is considered additive and independent of the process. The Kalman filter equations will then be modified taking into consideration the observation noise. This requires some information about the noise, for example its variance. Since the model thus obtained is not an AR process, a bias on the estimated parameters may be introduced. Thus, the identification algorithm must also be modified to deal with this case.

VI. Conclusion

A new algorithm for adaptive stable identification and robust reconstruction of an AR process subject to missing observations is proposed. At our knowledge, it is the first algorithm that permits on line processing of signals subject to missing observations using a lattice structure of the filter. It is an extension of the RLSL algorithm to the case of missing observations combined with a Kalman filter for the prediction. This algorithm guarantees the stability of the model identified. In addition, it is robust to large number of missing observations, and simulation shows that it offers a fast convergence and parameter tracking for any Bernoulli's probability of missing samples. However, this is at the expense of a bias in the identified model due to the cost function minimized by the algorithm in the case of missing observations. This algorithm has been compared to the Kalman pseudo linear RLS algorithm [19]. The latter uses a direct realization of the filter, hence the stability of the filter identified is not guaranteed. In addition, simulation shows that its performance depends highly on the observation pattern particularly for high probability of missing samples. Simulation on speech and musical signals shows the advantage of the proposed algorithm in terms of reconstruction error and particularly for a high probability of missing samples. Listening tests approves the results obtained. Indeed, when 85% of the samples are missing, a speech signal reconstructed using the proposed algorithm is still intelligible. The proposed algorithm may be modified to deal with noisy data, this may be the subject of a further publication.

References

[1] P. Albertos and R. Sanchis and A. Sala, "Output prediction under scarce data operation: control application", *Automatica, 35*, pp. 1671–1681, 1999.

[2] P. Bondon, "Prediction with incomplete past of a stationnary process," *Elsevier, Stochastic processes and their applications*, pp. 67–76, 2002.

[3] P. Brockwell and R. Davis, "Time Series: Theory and methods," *Springer-Verlag*, 1991.

[4] J.P. Burg, "Maximum entropy spectral analysis," *Ph.D. dissertation, Geophysics Dept., Stanford Univ., Stanford, CA*, May 1975.

[5] J. Durbin, "The fitting of time series models," *Rev. Inst. Int. Statist.*, vol. 28, pp. 233–243, 1960.

[6] B. Friedlander, "Lattice filters for adaptive processing," in *Proceedings of the IEEE*, vol.70, No. 8, pp. 829–867, 1982.

[7] A. Isaksson, "Identification of ARX Models Subject to Missing Data," *IEEE Transactions on Automatic Control*, vol. 38, pp. 813–819, 1993.

[8] A. Isaksson, "A Recursive EM Algorithm for identification subject to missing data," in *SYSID*, Copenhagen, Denmark, 1994, vol. 2, pp. 679–684.

[9] F. Itakura and S. Saito, "A statistical method for estimation of speech spectral density and formant frequencies," in *Electron. Commun.*, vol. 53-A, pp. 36-43, 1970.

[10] D. Lee and M. Morf and B. Friedlander, "Recursive least squares ladder estimation algorithms," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 29, pp. 627–641, 1981.

[11] N. Levinson, "The Wiener (Root Mean Square) Error Criterion in Filter Design and Prediction," *J. Math. Phys.*, vol. 25, pp. 261–278, 1947.

[12] S. Mirsaidi and G. Fleury and J. Oksman, "LMS Like AR modeling in the case of missing observations," *IEEE Transactions on Signal Processing*, vol. 45, No. 6, pp. 1574–1583, 1997.

[13] J. Makhoul, "Stable and efficient lattice methods for linear prediction," *IEEE Trans. on Acoustics, Speech and Signal Processing*, pp. 423–428, 1977.

[14] J. Makhoul and R. Viswanathan, "Adaptive lattice methods for linear prediction," *in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Tulsa, OK*, pp. 83–86, April 1978.

[15] J. Makhoul and L. Cosell, "Adaptive lattice analysis of speech," *IEEE Trans. on circuits and systems*, vol. 28, pp. 494–499, 1981.

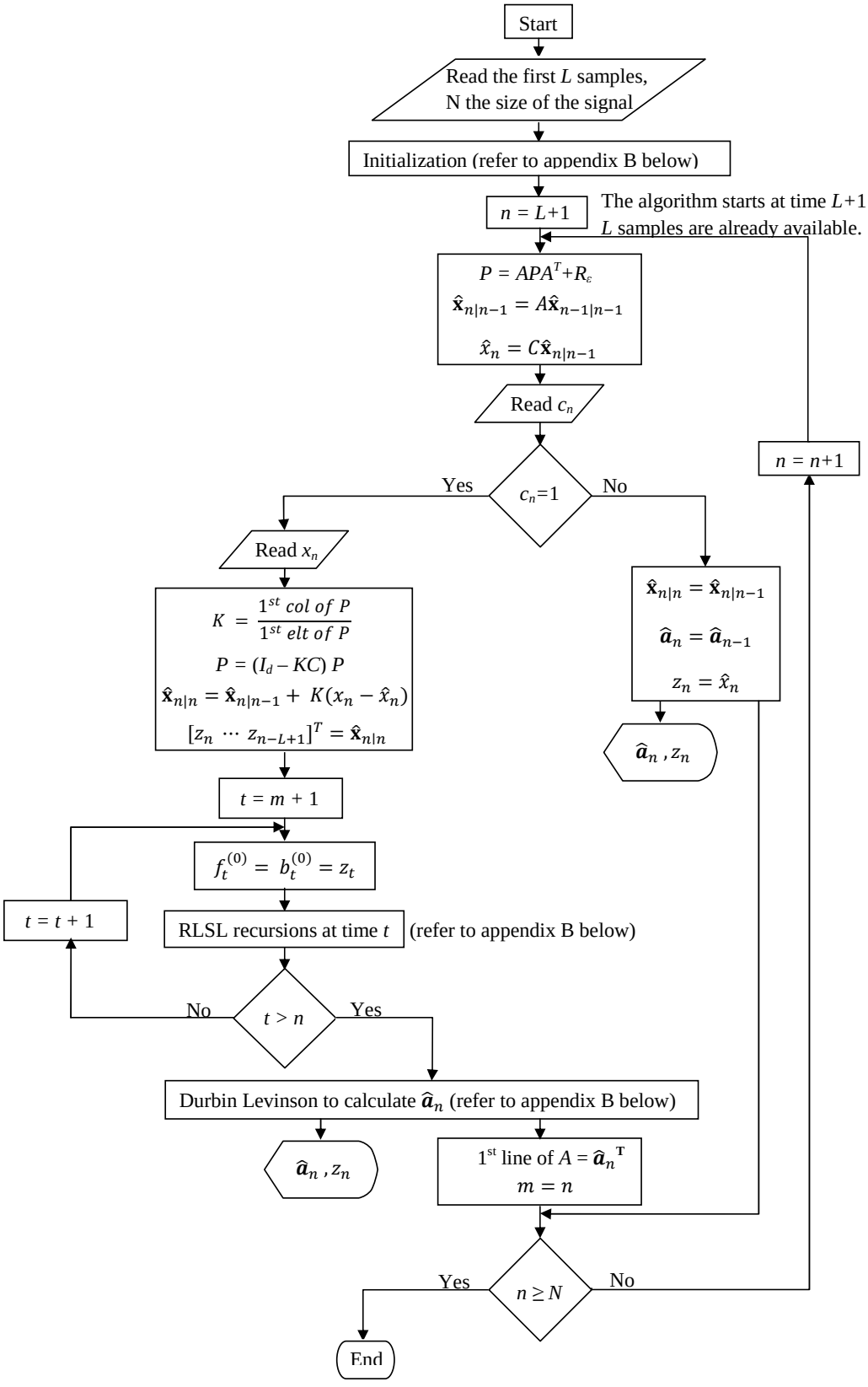
[16] R. Sanchis and P. Albertos, "Recursive identification under scarce measurements - convergence analysis," *Automatica, 38*, pp. 535–544, 2002.

[17] C. Therrien, "The missing observations theorem and a new proof of Levinson's recursion," *in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), Kobe, Japan*, pp. 2307–2308, 2005.

[18] R. Zgheib and G. Fleury and E. Lahalle, "New fast recursive algorithms for simultaneous reconstruction and identification of AR processes with missing observations," *Proc. Fourteenth European Signal Processing Conference, EUSIPCO, Florence, Italy*, 2006.

[19] R. Zgheib and G. Fleury and E. Lahalle, "New fast algorithm for simultaneous identification and optimal reconstruction of non stationary AR processes with missing observations," *IEEE Proc. Twelveth Digital Signal Processing Workshop, Wyoming, USA*, pp. 371–376, 2006.

APPENDIX A


Start
Read the first L samples,
N the size of the signal
Initialization (refer to appendix B below)
n = L+1
The algorithm starts at time L+1 after L samples are already available.
P = APA^T+R_ε
x̂_{n|n−1} = Ax̂_{n−1|n−1}
x̂_n = Cx̂_{n|n−1}
Read c_n
c_n=1
Yes
No
Read x_n
K = 1st col of P / 1st elt of P
P = (I_d − KC) P
x̂_{n|n} = x̂_{n|n−1} + K(x_n − x̂_n)
[z_n ⋯ z_{n−L+1}]^T = x̂_{n|n}
x̂_{n|n} = x̂_{n|n−1}
â_n = â_{n−1}
z_n = x̂_n
â_n , z_n
n = n+1
t = m + 1
f_t^(0) = b_t^(0) = z_t
t = t + 1
RLSL recursions at time t (refer to appendix B below)
No
t > n
Yes
Durbin Levinson to calculate â_n (refer to appendix B below)
â_n , z_n
1st line of A = â_n^T
m = n
Yes
n ≥ N
No
End

B

Initialization: is at time L since the algorithm starts at time L+1.

$$\hat{\mathbf{x}}_{L|L} = [x_L \cdots x_1]^T,$$

$$[z_L \cdots z_1]^T = \hat{\mathbf{x}}_{L|L},$$

$$A = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix},$$

$$R_\varepsilon = [1\ 0 \ldots 0\ ]^T [1\ 0 \ldots 0],$$

$$P = 10^3 \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix},$$

$$[k^{(1)} \quad \cdots \quad k^{(L)}] = [0 \quad \cdots \quad 0],$$

$$\left[b_L^{(0)} \quad \cdots \quad b_L^{(L)}\right] = [0 \quad \cdots \quad 0],$$

$$\left[f_L^{(0)} \quad \cdots \quad f_L^{(L)}\right] = [0 \quad \cdots \quad 0],$$

$$m = L,\ \lambda = 0.99,$$

$$[d^{(1)} \quad \cdots \quad d^{(L)}] = [0 \quad \cdots \quad 0],$$

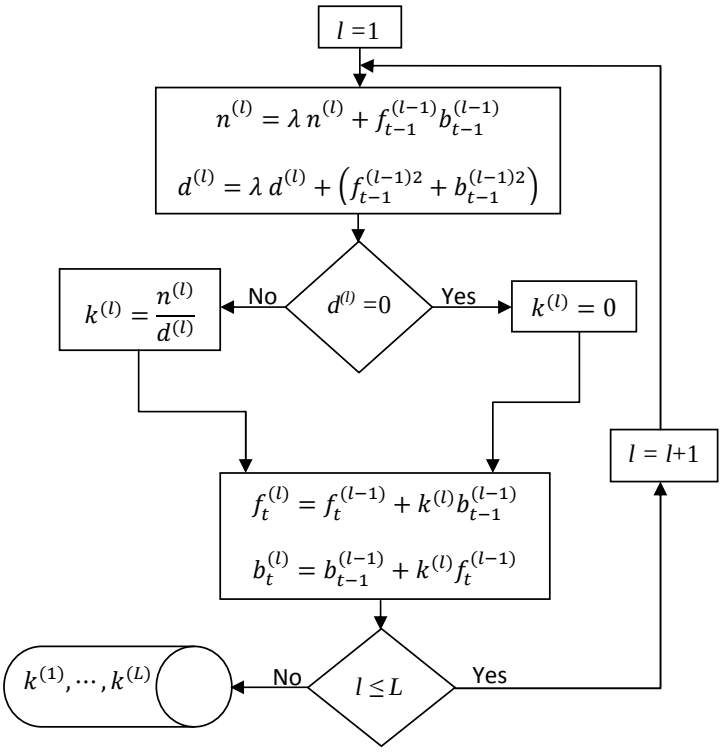
$$[n^{(1)} \quad \cdots \quad n^{(L)}] = [0 \quad \cdots \quad 0]$$

z_n is the signal reconstruction process

m is for each time, the instant of the last previous available sample

$n^{(l)}, d^{(l)}$ are the numerator and the denominator of the reflection coefficient $k^{(l)}$. They are updated at each time recursively.

RLSL recursions at time t :



Durbin Levinson at time n :