論文 / 著書情報
Article / Book Information

Title	Multi-Band Decomposition of the Linear Prediction Error Applied to Adaptive AR Spectral Estimation
Authors	Fernando Gil. V RESENDE, Keiichi Tokuda, Mineo KANEKO, AKINORI NISHIHARA
Citation	IEICE Trans. Fundamentals, Vol. E80-A, No. 2, pp. 365-376
Pub. date	1997, 2
URL	http://search.ieice.org/
Copyright	(c) 1997 Institute of Electronics, Information and Communication Engineers

PAPER

Multi-Band Decomposition of the Linear Prediction Error Applied to Adaptive AR Spectral Estimation

Fernando Gil V. RESENDE Jr.†, *Nonmember*, **Keiichi TOKUDA**††, **Mineo KANEKO**†††, *and* **Akinori NISHIHARA**†, *Members*

SUMMARY A new structure for adaptive AR spectral estimation based on multi-band decomposition of the linear prediction error is introduced and the mathematical background for the solution of the related adaptive filtering problem is derived. The presented structure gives rise to AR spectral estimates that represent the true underlying spectrum with better fidelity than conventional LS methods by allowing an arbitrary trade-off between variance of spectral estimates and tracking ability of the estimator along the frequency spectrum. The linear prediction error is decomposed through a filter bank and components of each band are analyzed by different window lengths, allowing long windows to track slowly varying signals and short windows to observe fastly varying components. The correlation matrix of the input signal is shown to satisfy both time-update and order-update properties for rectangular windowing functions, and an RLS algorithm based on each property is presented. Adaptive forward and backward relations are used to derive a mathematical framework that serves as a basis for the design of fast RLS algorithms. Also, computer experiments comparing the performance of conventional and the proposed multi-band methods are depicted and discussed.

key words: *digital signal processing, AR spectral estimation, filter banks, adaptive filtering, recursive least-squares algorithms*

1. Introduction

Adaptive filter theory has been a continuously growing field with applications in several areas, ranging from system identification to interference cancelling through prediction and spectral estimation [1]. In parametric estimation, the autoregressive (AR) model is the most popular because of the variety of fast algorithms that have been designed to solve the adaptive filtering problem involved in the estimation process [2], [3]. The least-squares (LS) modeling criterion is widely used in connection with the AR model because it is particularly suitable for solving the related optimization problem.

Conventional LS methods typically use a single window in the analysis of the linear prediction error implying a fixed trade-off between variance of spectral estimates and tracking ability of the estimator, herein referred to as frequency resolution and time resolution, respectively. Since the statistical properties of the analyzed signal may vary along the frequency spectrum, the possibility of having different windowing functions for distinct bands of the frequency spectrum can improve the fidelity of spectral estimates.

Recently, wavelet decomposition [4] of the linear prediction error was used to improve the quality of spectral estimates by trading-off time and frequency resolutions in a wavelet-like form, that is, long windows for low-frequency bands and short windows for high-frequency ones [5]. Even though the wavelet-based LS method was shown to improve the quality of AR spectral estimates as compared to those obtained with conventional LS methods for input signals composed of long-duration low-frequency and short-duration high-frequency components, the trade-off between time and frequency resolutions is fixed by the wavelet transform structure itself limiting its applicability. Also, since the correlation matrix of the input signal related to the wavelet-based method does not satisfy the so-called time-update property [6], the design of fast wavelet-based LS algorithms, if feasible, is not similar to conventional fast LS methods. Furthermore, two-channel finite impulse response (FIR) perfect reconstruction quadrature mirror filter (QMF) banks, such as those used in the wavelet transform, cannot simultaneously have linear phase and paraunitary properties, unless the filters have fairly trivial forms [7], what leads to a non-constant group delay for different components of the filtered linear prediction error. In this paper, a new structure for adaptive AR multi-band LS spectral estimation which overcomes the above mentioned limitations is introduced and analyzed.

The multi-band structure presented in this work enables different windowing functions to be used in the analysis of distinct frequency bands of the linear prediction error through its decomposition with a power complementary filter bank. With this approach, fine details of the prediction error signal can be analyzed through short windows and slowly varying components can be observed by long windows independently of their frequencies. In other words, through the multi-band approach introduced in this paper, time and frequency resolutions can be arbitrarily traded-off, giving

Manuscript received May 23, 1996.

Manuscript revised July 28, 1996.

†The authors are with the Department of Physical Electronics, Tokyo Institute of Technology, Tokyo, 152 Japan.

††The author is with the Department of Computer Science, Nagoya Institute of Technology, Nagoya-shi, 466 Japan.

†††The author is with the School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa-ken, 923–12 Japan.

rise to spectral estimates that represent the true underlying spectrum with more fidelity than presently used LS methods.

The LS solution for the presented structure is obtained through a set of normal equations, where the input signal correlation matrix is shown to satisfy both time-update and order-update recursion formulae. Based on the time-update recursion formula of the correlation matrix an algorithm similar to the conventional RLS algorithm [1] is derived. Based on the order-update recursion formula of the correlation matrix an algorithm similar to the Levinson RLS algorithm [8] is presented. Also, a study on the relations between forward and backward linear prediction for the multi-band structure sets the mathematical background for the design of fast recursive LS (RLS) algorithms [9]. Computer experiments comparing the performance of conventional and multi-band LS methods are shown and discussed.

This paper is organized as follows. In the next section, the multi-band structure is introduced, its cost function is defined and compared to conventional and wavelet-based cost functions and the optimization problem is solved through a set of normal equations. In Sect. 3, a multi-band decomposition-based RLS algorithm based on the time-update property of the correlation matrix is described. In Sect. 4, the order-update property of the correlation matrix is used to design an RLS algorithm based on the Levinson RLS algorithm. Adaptive forward and backward linear prediction for the multi-band structure is studied in Sect. 5, and the mathematical background for the design of fast AR multi-band LS algorithms is exposed. In Sect. 6, computer experiments are shown and discussed. Section 7 presents our conclusions and on-going research. From Appendix A to Appendix F a thorough explanation of the mathematical results contained in Sects. 3, 4 and 5 is undertaken. The nomenclature used in this paper follows the standard contained in [1] with modifications that naturally arise from the multi-band characteristic of the presented structure. For convenience of notation, the presentation adopts the system identification scheme until Sect. 3, from where the terminology peculiar to spectral estimation theory is applied.

2. Multi-Window LS and Normal Equations

In Sect. 2.1, conventional and wavelet-based LS methods are briefly exposed and the multi-band LS method is introduced. In Sect. 2.2, the minimization condition for the multi-band LS cost function defined in the previous subsection is applied and a set of normal equations is obtained.

2.1 Conventional and Multi-Window LS

System identification, as well as adaptive AR spectral estimation, are based upon the update of model parameters subject to the minimization of a cost function [1], [2]. The LS criterion is widely used in adaptive filtering theory and its cost function is defined as

$$\varepsilon(n) = \sum_{i=-\infty}^{\infty} \upsilon(n-i)e^2(i), \tag{1}$$

for real values of the estimation error $e(i)$, where $\upsilon(k)$ is a windowing function, whose z-transform is represented by $\Upsilon(z)$. A block diagram scheme for (1) is depicted in Fig. 1 (a). From now on, estimation methods for which the cost function is defined as shown in (1) will be called conventional LS methods.

For an AR modeling of the input signal, in system identification problems, $e(i)$ can be expressed as

$$\begin{aligned} e(i) &= d(i) - y(i) \\ &= d(i) - \sum_{k=0}^{M-1} w_k(n)u(i-k), \quad i \leqq n, \end{aligned} \tag{2}$$

where M is the order of the model, $d(i)$ is the desired response, $y(i)$ is the estimated value obtained by filtering the input signal $u(i)$ with an FIR adaptive filter, whose coefficients at time-index n are represented by $w_k(n), k = 0, \ldots, M-1$. For rectangular sliding windows of length L, $\upsilon(k)$ may be written as

$$\upsilon(k) = \begin{cases} 1; & 0 \leqq k \leqq L-1, \\ 0; & \text{otherwise.} \end{cases} \tag{3}$$

A particular disadvantage of methods based on the cost function described in (1) is that the square of $e(i)$ is filtered by a single sliding window $\upsilon(k)$, which determines a unique value for the relation between the variance of the model parameters estimates and the tracking ability of the system for the whole frequency spectrum at every time-index n. The physical limitation of this relation is similar in concept to the trade-off between time resolution and frequency resolution that, in accordance with the uncertainty principle, are inversely proportional while having a lower bounded product [10].

Discrete-time wavelet transform coefficients of the linear prediction error were used in [6] to define a cost function as follows

$$\varepsilon(n) = \sum_{j=1}^{B} \varepsilon^{(j)}(n) \tag{4}$$

$$\varepsilon^{(j)}(n) = \sum_{i=-\infty}^{n} \lambda^{n-i} \left(E_n^{(j)}(i)\right)^2 \tag{5}$$

$$E_n^{(j)}(i) = \begin{cases} e^{(j)}(n - 2^j(n-i)), & 1 \leqq j \leqq B-1 \\ e^{(j)}(n - 2^{j-1}(n-i)), & j = B \end{cases} \tag{6}$$

$$e^{(j)}(n) = h^{(j)}(n) * e(n) \tag{7}$$

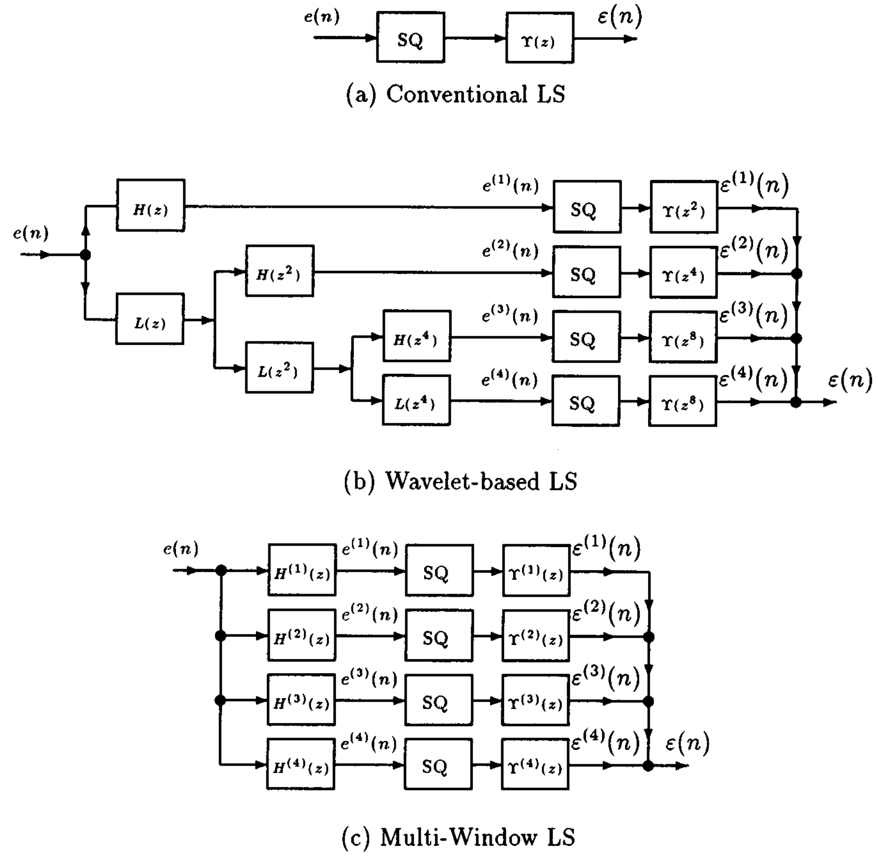


Fig. 1 Block-diagram representation of the cost function definition for (a) conventional LS, (b) wavelet-based LS with $B = 4$, and (c) the proposed multi-window LS method with $B = 4$. The "SQ" symbol represents the square operation and $\Upsilon(z)$ is the z-transform of the windowing function $v(k)$.

$$H^{(j)}(z) = \begin{cases} H\left(z^{2^{j-1}}\right)\displaystyle\prod_{l=0}^{j-2} L\left(z^{2^l}\right), & 1 \leqq j \leqq B-1, \\ \displaystyle\prod_{l=0}^{j-2} L\left(z^{2^l}\right), & j = B, \end{cases} \tag{8}$$

where the "$*$" symbol denotes the convolution operation and the estimation error $e(i)$, which is assumed to be real valued, is defined by (2) and decomposed into B bands by the analysis filters $h^{(j)}(n)$, $1 \leqq j \leqq B$, with z-transforms $H^{(j)}(z)$ as shown in Fig. 1 (b) for the 4-band case. The basic idea of the wavelet-based method is to improve the precision of estimates by using long windows to analyze low-frequency components of the estimation error and short windows for high-frequency ones.

The so-defined wavelet-based LS scheme achieves better estimates than conventional LS methods when the analyzed signal is mainly composed of long-duration low-frequency and/or short-duration high-frequency components. Whenever a different time-frequency resolution trade-off is needed, for example when the analyzed signal is composed of long-duration high-frequency and short-duration low-frequency components, the wavelet-based method is not suitable. Moreover, since the correlation matrix of the input signal for the wavelet-based structure does not satisfy the time-update property, design of fast algorithms cannot be undertaken by algebraic manipulations similar to those for conventional methods. Also, different bands of the filtered linear prediction error do not have a constant group delay since non-trivial paraunitary linear-phase filter banks cannot be obtained with the two-channel wavelet decomposition structure [7].

The following cost function, which is introduced and analyzed in this paper, overcomes the above mentioned limitations:

$$\varepsilon(n) = \sum_{j=1}^{B} \varepsilon^{(j)}(n) \tag{9}$$

$$\varepsilon^{(j)}(n) = \sum_{i=-\infty}^{\infty} v^{(j)}(n-i)\left(e^{(j)}(i)\right)^2 \tag{10}$$

$$e^{(j)}(i) = h^{(j)}(i) * e(i), \tag{11}$$

where the analysis bank $h^{(j)}(i)$, $j = 1, \ldots, B$, is a power complementary filter bank (see Fig. 1 (c)). By using a power complementary filter bank and proper weighting factors for each band we can guarantee a uniform response of the system in the sense that $\mathcal{E}[\varepsilon(n)] = c\mathcal{E}[e^2(n)]$, where $\mathcal{E}[.]$ denotes the mean value operator (see Appendix A for a proof of this statement). Phase-linearity of the analysis filter bank implies a constant group delay for all components of the filtered estimation error. Differently from the wavelet-based method, the windowing functions $v^{(j)}(k), j = 1, \ldots, B$ for this structure can be independently chosen, enabling the use of the most suitable time-frequency resolution grid according to the analyzed signal statistical properties. Furthermore, if rectangular windows are used, time-update and order-update relations for the input signal correlation matrix can be written, opening the way for the design of fast RLS algorithms.

In this paper, rectangular sliding-windows of distinct lengths are applied to different bands of the estimation error, that is,

$$v^{(j)}(k) = \begin{cases} \frac{1}{L^{(j)}}; & 0 \leqq k \leqq L^{(j)} - 1, \\ 0; & \text{otherwise,} \end{cases} \tag{12}$$

for $1 \leqq j \leqq B$, where the factor $\frac{1}{L^{(j)}}$ together with the use of a power complementary filter bank are necessary conditions to ensure a uniform response of the system for all frequencies in the sense mentioned in the above paragraph.

Once the structure of the new method is mathematically described in Eqs. (9)–(12), the optimization problem for the new cost function can be subsequently tackled.

2.2 Normal Equations

It can be easily verified through Eqs. (2) and (9)–(11) that the cost function is a quadratic function of the model filter parameters, and minimization of $\varepsilon(n)$ is achieved when

$$\frac{\delta\varepsilon(n)}{\delta w_l(n)} = 0, \quad 0 \leqq l \leqq M-1. \tag{13}$$

Applying (2) and (9)–(12) in (13), we have

$$\sum_{k=0}^{M-1} w_k(n) \sum_{j=1}^{B} \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} u^{(j)}(n-i-k)u^{(j)}(n-i-l) = \sum_{j=1}^{B} \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} d^{(j)}(n-i)u^{(j)}(n-i-l) \tag{14}$$

for $0 \leqq l \leqq M-1$, where

$$u^{(j)}(i) = h^{(j)}(i) * u(i) \tag{15}$$

$$d^{(j)}(i) = h^{(j)}(i) * d(i). \tag{16}$$

This set of M equations, so-called normal equations, may be rewritten in matrix form as

$$\boldsymbol{\Phi}_M(n)\boldsymbol{w}_M(n) = \boldsymbol{\Theta}_M(n), \tag{17}$$

where

$$\boldsymbol{\Phi}_M(n) = \sum_{j=1}^{B} \boldsymbol{\Phi}_M^{(j)}(n) \tag{18}$$

$$\boldsymbol{\Phi}_M^{(j)}(n) = \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} \boldsymbol{u}_M^{(j)}(n-i) \left(\boldsymbol{u}_M^{(j)}(n-i)\right)^T \tag{19}$$

$$\boldsymbol{u}_M^{(j)}(n-i) = \begin{bmatrix} u^{(j)}(n-i) \\ u^{(j)}(n-i-1) \\ \vdots \\ u^{(j)}(n-i-M+1) \end{bmatrix} \tag{20}$$

$$\boldsymbol{w}_M(n) = [w_0(n), \ldots, w_{M-1}(n)]^T \tag{21}$$

$$\boldsymbol{\Theta}_M(n) = \sum_{j=1}^{B} \boldsymbol{\Theta}_M^{(j)}(n) \tag{22}$$

$$\boldsymbol{\Theta}_M^{(j)}(n) = \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} d^{(j)}(n-i)\boldsymbol{u}_M^{(j)}(n-i). \tag{23}$$

The solution of (17) for $\boldsymbol{w}_M(n)$ by inverting $\boldsymbol{\Phi}_M(n)$ for every time-index n is not computationally efficient. In the following, special properties of the correlation matrix $\boldsymbol{\Phi}_M(n)$ are studied and applied to the design of RLS algorithms.

3. An RLS Algorithm Based on Time-Update

The purpose of this section is to describe an RLS algorithm based on the time-update property of the correlation matrix related to the multi-band decomposition-based structure introduced in Sect. 2. To improve readability, a detailed derivation of time recursive relations for the correlation matrix and the cross-correlation vector, and the application of these equations in connection with the *matrix inversion lemma* (MIL) [1] to obtain the RLS algorithm described in the following is undertaken in Appendix B.

The RLS algorithm for the system identification multi-band LS method based on the time-update property of the correlation matrix can be summarized as follows

$$\boldsymbol{K}_M(n) = \boldsymbol{P}_M(n-1)\boldsymbol{U}_M(n,n-L) \left(\boldsymbol{S}^{-1} + \boldsymbol{U}_M^T(n,n-L)\boldsymbol{P}_M(n-1)\boldsymbol{U}_M(n,n-L)\right)^{-1} \tag{24}$$

$$\boldsymbol{\alpha}(n) = \boldsymbol{d}(n,n-L) - \boldsymbol{U}_M^T(n,n-L)\boldsymbol{w}_M(n-1) \tag{25}$$

$$\boldsymbol{w}_M(n) = \boldsymbol{w}_M(n-1) + \boldsymbol{K}_M(n)\boldsymbol{\alpha}(n) \tag{26}$$

$$\boldsymbol{P}_M(n) = \boldsymbol{P}_M(n-1) - \boldsymbol{K}_M(n)\boldsymbol{U}_M^T(n,n-L)\boldsymbol{P}_M(n-1) \tag{27}$$

It should be noted that in this form, the algorithm has computation complexity on the order of (M^2+B^3), but if each band is separately time-updated in two steps, i.e. if the MIL is applied $2B$ times, it is not necessary to compute any matrix inversion and the computational complexity of the resulting algorithm is on the order of BM^2.

A similar RLS algorithm for AR spectral estimation based on forward or backward linear prediction parameters can be obtained through straightforward substitution by the corresponding quantities. For example, in the forward prediction case the desired response vector is given in (A· 48) while the input signal is one time-step delayed with respect to the input signal for system identification structures. Time-update recursions for the coefficient vector and the *a priori* forward prediction error vector are shown in (28) and (29), respectively. Forward and backward prediction analysis is studied in Sect. 5.

4. An RLS Algorithm Based on Order-Update

In this section an RLS algorithm with complexity on the order of $(BM + M^2)$, based on the Levinson RLS algorithm [8], for AR spectral estimation is exposed. We assume the use of rectangular sliding windowing functions for every band, but other windowing functions, such as exponential windows, can also be used with a similar algorithm [11]. Detailed derivations related to the order-update property of the correlation matrix and the time-update recursion formula for the autocorrelation vector are shown in Appendix C.

Table 1 An RLS algorithm for AR multi-band LS spectral analysis based on the order-update property of the input signal correlation matrix.

For each time-index $n = 1, 2, 3, \dots$ compute:

For $j = 1, 2, \dots, B$

Calculate $u^{(j)}(n)$ using (15)

$$\mathbf{c}^{(j)}_{M+1}(n) = \mathbf{c}^{(j)}_{M+1}(n-1) + \frac{1}{L^{(j)}}\left(u^{(j)}(n)\mathbf{u}^{(j)}_{M+1}(n) - u^{(j)}(n-L^{(j)})\mathbf{u}^{(j)}_{M+1}(n-L^{(j)})\right),$$

End

$$\mathbf{c}_{M+1}(n) = \mathbf{c}^{(1)}_{M+1}(n) + \mathbf{c}^{(2)}_{M+1}(n) + \cdots + \mathbf{c}^{(B)}_{M+1}(n)$$

$$\mathcal{F}_0(n) = \mathcal{B}_0(n) = c_0(n)$$

$$\mathbf{a}_0(n) = \mathbf{g}_0(n) = [1]$$

For $m = 1, 2, \cdots, M+1$ compute :

$$C_{m-1}(n) = [c_1(n), \dots, c_m(n)]\mathbf{g}_{m-1}(n-1)$$

$$K^f_m(n) = -C_{m-1}(n)/\mathcal{B}_{m-1}(n-1)$$

$$K^b_m(n) = -C_{m-1}(n)/\mathcal{F}_{m-1}(n)$$

$$\mathcal{F}_m(n) = \mathcal{F}_{m-1}(n) + K^f_m(n)C_{m-1}(n)$$

$$\mathcal{B}_m(n) = \mathcal{B}_{m-1}(n-1) + K^b_m(n)C_{m-1}(n)$$

$$\mathbf{a}_m(n) = \begin{bmatrix} \mathbf{a}_{m-1}(n) \\ 0 \end{bmatrix} + K^f_m(n)\begin{bmatrix} 0 \\ \mathbf{g}_{m-1}(n-1) \end{bmatrix}$$

$$\mathbf{g}_m(n) = \begin{bmatrix} 0 \\ \mathbf{g}_{m-1}(n-1) \end{bmatrix} + K^b_m(n)\begin{bmatrix} \mathbf{a}_{m-1}(n) \\ 0 \end{bmatrix}$$

End

The entire algorithm is summarized in Table 1. The forward prediction vector of order m at time-index n is represented by $\boldsymbol{a}_m(n)$ and defined in (A· 39). The corresponding backward prediction quantity is represented by $\boldsymbol{g}_m(n)$. The forward prediction cost function $\mathcal{F}_m(n)$ is defined in (A· 32), and $\mathcal{B}_m(n)$ is the corresponding quantity for the backward prediction case. It is interesting to note that since this algorithm is initialized from the time recursively updated autocorrelation coefficients rather than from the data sequence itself, the application of arbitrary recursive windows is allowed, what may improve its steady-state and tracking behavior [8], [11]. In Sect. 6 computer experiments using this algorithm are used to compare conventional and multi-band methods.

In the next section the mathematical basis for the design of fast multi-band RLS algorithms is presented.

5. Adaptive Forward and Backward Linear Prediction for Multi-Band LS

Relations between forward and backward linear prediction are commonly used to derive fast RLS algorithms to solve adaptive filtering problems. In the following, these relations are applied to the multi-band method to build a mathematical framework for the design of fast multi-band RLS algorithms. A detailed explanation of the following equations may be seen in Appendices D through F.

Adaptive Forward Linear Prediction:

$$\boldsymbol{a}_M(n) = \boldsymbol{a}_M(n-1) - \begin{bmatrix} \mathbf{0}_{1\times 2B} \\ \boldsymbol{K}_M(n-1) \end{bmatrix} \boldsymbol{\eta}_M(n) \quad (28)$$

$$\boldsymbol{\eta}_M(n) = \boldsymbol{U}^T_{M+1}(n, n-L)\boldsymbol{a}_M(n-1) \quad (29)$$

$$\mathcal{F}_M(n) = \mathcal{F}_M(n-1) + \boldsymbol{f}^T_M(n)\boldsymbol{S}\boldsymbol{\eta}_M(n) \quad (30)$$

Adaptive Backward Linear Prediction:

$$\boldsymbol{g}_M(n) = \boldsymbol{g}_M(n-1) - \begin{bmatrix} \boldsymbol{K}_M(n) \\ \mathbf{0}_{1\times 2B} \end{bmatrix} \boldsymbol{\psi}_M(n) \quad (31)$$

$$\boldsymbol{\psi}_M(n) = \boldsymbol{U}^T_{M+1}(n, n-L)\boldsymbol{g}_M(n-1) \quad (32)$$

$$\mathcal{B}_M(n) = \mathcal{B}_M(n-1) + \boldsymbol{b}^T_M(n)\boldsymbol{S}\boldsymbol{\psi}_M(n) \quad (33)$$

Gain Matrix:

$$\boldsymbol{K}_{M+1}(n) = \begin{bmatrix} 0 \\ \boldsymbol{K}_M(n-1) \end{bmatrix} + \frac{1}{\mathcal{F}_M(n)}\boldsymbol{a}_M(n)\boldsymbol{f}^T_M(n)\boldsymbol{S} \quad (34)$$

$$\boldsymbol{K}_{M+1}(n) = \begin{bmatrix} \boldsymbol{K}_M(n) \\ 0 \end{bmatrix} + \frac{1}{\mathcal{B}_M(n)}\boldsymbol{g}_M(n)\boldsymbol{b}^T_M(n)\boldsymbol{S} \quad (35)$$

Conversion Matrix:

Definition:

$$\boldsymbol{f}_M(n) = \boldsymbol{\Gamma}_M(n-1)\boldsymbol{\eta}_M(n) \quad (36)$$

$$\boldsymbol{b}_M(n) = \boldsymbol{\Gamma}_M(n)\boldsymbol{\psi}_M(n) \quad (37)$$

Update:

$$\boldsymbol{\Gamma}_{M+1}(n) = \boldsymbol{\Gamma}_M(n-1) - \frac{1}{\mathcal{F}_M(n)} \boldsymbol{f}_M(n) \boldsymbol{f}_M^T(n) \boldsymbol{S} \tag{38}$$

$$\boldsymbol{\Gamma}_{M+1}(n) = \boldsymbol{\Gamma}_M(n) - \frac{1}{\mathcal{B}_M(n)} \boldsymbol{b}_M(n) \boldsymbol{b}_M^T(n) \boldsymbol{S} \tag{39}$$

The relations described above show the basic recursive formulae for the forward and backward linear prediction problems for the presented multi-band decomposition-based structure. In particular, these equations can be seen as a generalization of conventional "single-window" LS methods for the case when the linear prediction error is analyzed through several windows.

A fast transversal filters (FTF) algorithm can be directly derived from the equations above, and this was undertaken in [9]. However, even though the algorithms presented in Sects. 3, 4, and 5 are exact solutions of the same RLS problem and equivalent results are obtained with infinite precision arithmetic, because of the well known numerical instability of the FTF algorithm under finite precision arithmetic [12]–[17], different results can be expected. Moreover, the error feedback mechanism [13]–[17] is only able to stabilize the FTF algorithm under a very restricted range for the forgetting factor values [1], hindering its application in several situations. For the exponential weighting factor λ, this range is defined by:

$$1 - \frac{1}{2M} < \lambda < 1, \tag{40}$$

where M is the order of the AR model [1]. In the particular case of the proposed method, which has its advantage over conventional approaches exactly when distinct values of the forgetting factor are used in different bands, the restriction described in (40) makes FTF implementations of multi-band decomposition-based LS methods unpractical. The forward and backward recursive relations presented in this section, however, are not only of theoretical interest but also may serve as a mathematical framework for the design of fast RLS algorithms with better numerical properties, such as the lattice LS algorithm [1].

6. Computer Experiments and Discussion

Computer experiments are shown in Fig. 2 and Fig. 3. In order to be able to evaluate the convergence behavior of the presented methods while avoiding extremely large figures, even though estimation was undertaken at every sample of the analyzed signal, plots are shown for every 10 samples in Fig. 2 and Fig. 3. The algorithm described in Table 1 was used to compute estimates for the multi-band LS method in these figures. Through computer experiments with an FTF implementation of

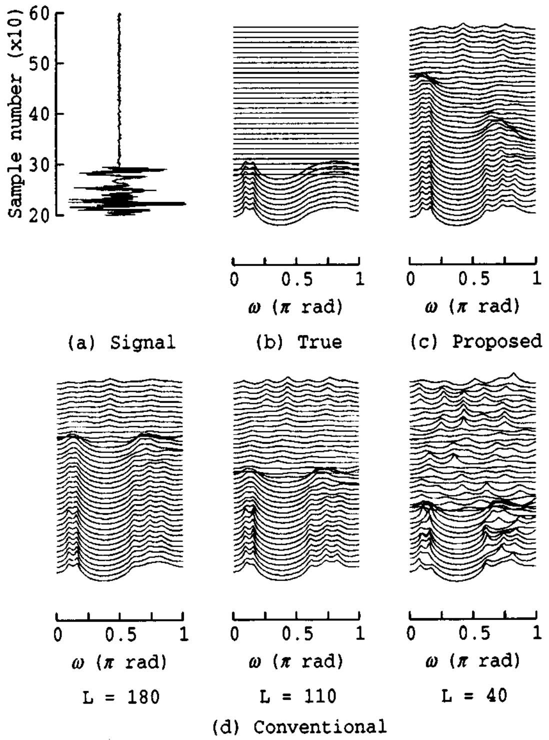


Fig. 2 Time-varying AR spectral estimates for the test signal displayed in (a) (time domain) and (b) (frequency domain), obtained through (c) multi-window LS with L = [180, 160, 140, 120, 100, 80, 60, 40] (from lowest to highest frequency band) using an 8-channel linear-phase paraunitary filter bank designed in [18], and (d) conventional LS with L = 180, 110, and 40.

multi-band LS methods directly obtained from the relations derived in Sect. 5, we have verified that, even though correct tracking is obtained during some iterations, its tendency to diverge makes it unsuitable for nonstationary signal analysis. In Fig. 2, an 8-channel linear-phase paraunitary analysis filter bank designed in [18] is used to decompose the linear prediction error. The input signal is shown in Fig. 2 (a) (time domain) and Fig. 2 (b) (frequency domain). AR spectral estimates for multi-band and conventional LS (with 3 different window lengths) are shown in Fig. 2 (c) and Fig. 2 (d), respectively. The AR model for both methods has 10 coefficients. This simulation high-lights the combining effect that multi-band LS has, making its behavior vary according to the windowing function applied to each band, in this case, resembling long-windowed conventional LS for low-frequencies and short-windowed conventional LS for high-frequencies.

Computer experiments displayed in Fig. 3 compare the performance of conventional LS (Fig. 3 (d)) and multi-band LS for two 4-band linear-phase analysis banks, both designed in [18], one of which is a paraunitary filter bank (Fig. 3 (b)), and the other has a near-perfect-reconstruction characteristic (Fig. 3 (c)). The reason we have used these two filter banks in our analysis is that while power complementariness im-

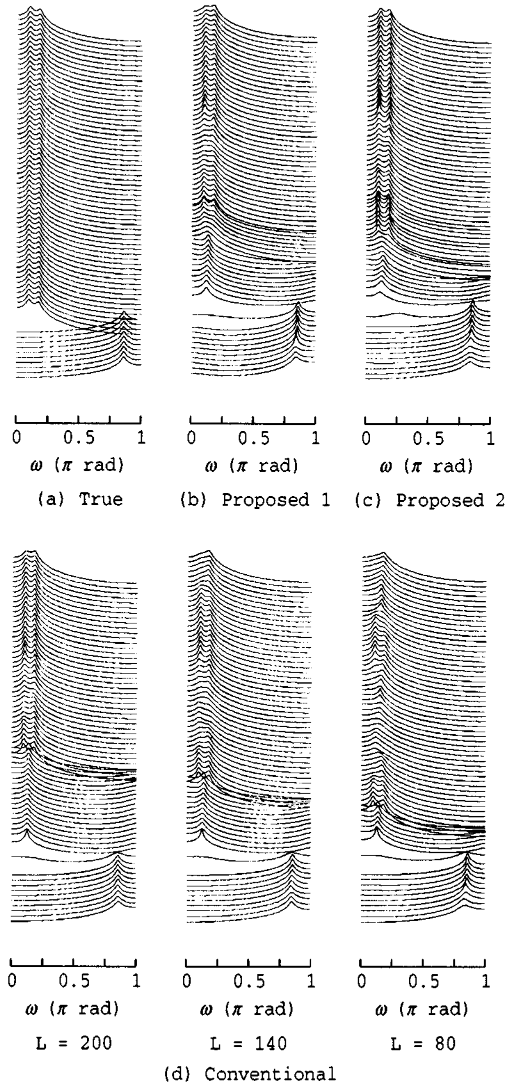


Fig. 3 Comparison between (a) the true running spectrum, and AR spectral estimates based on (b) multi-window LS with $L =$ [200, 160, 120, 80] (from lowest to highest frequency band), using a 4-channel linear-phase paraunitary filter bank (filter length $= 8$)[18], (c) multi-window LS with $L =$ [200, 160, 120, 80] (from lowest to highest frequency band), using a 4-channel linear-phase near-perfect-reconstruction filter bank (filter length $=$ 32) [18], and (d) conventional LS with three different window lengths: 200, 140 and 80.

plies uniformity in the spectral analysis in the sense that $\mathcal{E}[\varepsilon(n)] = c\mathcal{E}[e^2(n)]$, near-perfect-reconstruction designs offer a much better stopband response emphasizing the difference in the analysis for each band obtained through different window lengths. The test signal used begins with a high-frequency component which stops when two spatially-close low-frequency components start. Fig. 3 (a) depicts the true spectrum. In Fig. 3 (b) and in Fig. 3 (c), a paraunitary and a near-perfect-reconstruction filter have been respectively used in connection with AR multi-band LS. Fig. 3 (d) depicts the results obtained with the AR conventional LS method for different window lengths. An AR model of order 4 was used for all methods in Fig. 3.

Through a first analysis of Fig. 3, it is clearly observed that the proposed method with both filter banks yields faster convergence than conventional methods in the estimation of the two low-frequency components. Careful comparison between the results of Fig. 3 (c) (proposed 2) with the others highlights the fact that it takes one more spectral line for the system of Fig. 3 (c) as compared to the other systems to show that the high-frequency component stopped being part of the analyzed signal. On the other hand, the same system (proposed 2) is the one which gives the fastest convergence in the estimation of the two spatially-close low-frequency components. The explanation for these two results follows.

The relative one spectral line delay mentioned in the above paragraph is due to the group delay effect caused by the convolution of the linear prediction error with a filter bank in the proposed method. This effect is not observed in Fig. 3 (b) because while the filters used in Fig. 3 (c) have length 32, those for Fig. 3 (b) have length 8 giving rise to a relatively short group delay.

The faster convergence observed in Fig. 3 (c) in the estimation of the two low-frequency components can be explained by the following facts:

- the influence of any component in the computation of estimates, after it stops being part of the analyzed signal when conventional methods are used with sliding windows, only vanishes completely after a number of iterations equal to the window length used in the analysis plus the AR model order. As a consequence, from the point it actually finishes, the high-frequency component, when conventional methods are used in Fig. 3 (d) (AR model order $=$ 4), will take 204, 144 and 84 iterations to be totally forgotten in the evaluation of spectral estimates of systems with $L =$ 200, 140 and 80, respectively.
- the window length for the proposed method varies with the frequency localization of the analyzed component. If an ideal frequency-selective 4-channel filter bank, with filter length equal to F and AR model order equal to M were used and, from low to high frequencies, we had $L =$ [200, 160, 120, 80], it would take (80 $+M$ $+$ F) iterations for the effect of the high-frequency component of Fig. 3 to completely vanish when using the proposed method. With practical filter banks, as can be seen in Fig. 3 (b) and Fig. 3 (c), the actual window length value used in the spectral analysis suffers the influence of window lengths of neighboring bands. This influence, however, can be reduced by using filters with good stopband attenuation at the expense of increasing group delay, and this trade-off is observed through a comparison of Fig. 3 (b)

with Fig. 3 (c).

Computer experiments evidence the fact that while with conventional LS methods one has a fixed trade-off relation between frequency and time resolutions, with the proposed multi-band LS one has the flexibility to trade-off time and frequency resolutions along the frequency spectrum giving rise to spectral estimates closer to the true underlying running spectrum whenever the characteristics of the statistics of the analyzed signal varies along the frequency spectrum.

7. Conclusion

A new method for adaptive AR spectral analysis that increases the fidelity of estimates through the trade-off of variance and tracking ability of spectral estimates along the frequency spectrum was introduced. The basic idea is to decompose the linear prediction error into several frequency bands and analyze the components of each band through a different window.

Time-update and order-update properties of the input signal correlation matrix were used to derive RLS algorithms for rectangular sliding windowing functions. Forward and backward linear prediction relations were derived to build a mathematical framework for the design of fast recursive algorithms for AR multi-band LS. The flexibility obtained with multi-band LS was exemplified with computer experiments and compared to conventional LS methods.

Designing least-mean-square (LMS) [19], least-squares-lattice (LSL), and QR-decomposition-based LSL (QRD-LSL) algorithms for the multi-band method is subject of ongoing research. The construction of algorithms that manage to independently self-adjust window lengths along the time according to variations on the statistical properties of each band of the input signal is also subject of present research [11].

Acknowledgment

The authors would like to express their appreciation to Profs. N. Fujii and S. Takagi of the Tokyo Institute of Technology and Prof. P.S.R. Diniz of COPPE/UFRJ for their input and encouragement. The authors are also grateful to the anonymous reviewers for their stimulating critiques. This work was partially supported by Monbusho, Ministry of Education of Japan.

References

[1] S. Haykin, "Adaptive Filter Theory," Prentice-Hall, Englewood Cliffs, N.J., 1991.
[2] S.L. Marple, "Digital Spectral Analysis with Applications," Prentice-Hall, Englewood Cliffs, N.J., 1987.
[3] P. Strobach, "New forms of Levinson and Shur algorithms," IEEE ASSP Magazine, vol.8, no.1, Jan. 1991.
[4] I. Daubechies, "Ten Lectures on Wavelets," SIAM, Philadelphia, Pennsylvania, 1992.
[5] F.G. Resende, K. Tokuda, and M. Kaneko, "AR spectrum estimation based on wavelet representation," Proc. IEEE ISCAS, London, England, 1994.
[6] F.G. Resende, K. Tokuda, and M. Kaneko, "Adaptive AR spectral estimation based on wavelet decomposition of the linear prediction error," IEICE Trans. Fundamentals, vol.E79-A, no.5, pp.665–673, May 1996.
[7] P.P. Vaidyanathan, "Multirate Systems and Filter Banks," Prentice-Hall, Englewood Cliffs, N.J., 1993.
[8] P. Strobach, "Recursive triangular array ladder algorithms," IEEE Trans. Signal Processing, vol.39, no.1, Jan. 1991.
[9] F.G. Resende, K. Tokuda, and M. Kaneko, "AR spectral estimation based on multi-window analysis of the linear prediction error," Proc. 38th Midwest Symposium on Circuits and Systems, Rio de Janeiro, Brazil, 1995.
[10] R. Bracewell, "The Fourier Transform and its Applications," McGraw-Hill, New York, N.Y., 1986.
[11] F.G. Resende, P.S.R. Diniz, M. Kaneko, and A. Nishihara, "Adaptive AR spectral estimation based on multi-band decomposition of the linear prediction error with variable forgetting factors," to appear in Proc. of IEEE ICASSP, Munich, Germany, 1997.
[12] S. Ljung and L. Ljung, "Error propagation properties of recursive least-squares adaptation algorithms," Automatica, vol.21, pp.157–167, 1985.
[13] J.L. Botto and G.V. Moustakides, "Stabilization of fast recursive least-squares transversal filters for adaptive filtering," Proc. of ICASSP, pp.403–407, April 1987.
[14] J.L. Botto and G.V. Moustakides. "Stabilization of fast Kalman algorithms," IEEE Trans. ASSP, vol.ASSP-37, pp.1342–1348, Sept. 1989.
[15] D.T.M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," IEEE Trans. Signal Processing, vol.39, no.1, pp.92–114, Jan. 1991.
[16] D.T.M. Slock, "The backward consistency concept and roundoff error propagation dynamics in RLS algorithms," Optical Engineering, vol.31, pp.1153–1169, 1992.
[17] D.T.M. Slock and T. Kailath, "Fast Transversal RLS Algorithms," in Adaptive System Identification and Signal Processing Algorithms, eds. N. Kalouptsidis and S. Theodoridis, pp.123–190, Prentice-Hall, Englewood Cliffs, N.J., 1993.
[18] A.K. Soman, P.P. Vaidyanathan, and T.Q. Nguyen, "Linear phase paraunitary filter banks: Theory, factorizations and designs," IEEE Trans. Signal Processing, vol.41, no.12, Dec. 1993.
[19] F.G. Resende, P.S.R. Diniz, K. Tokuda, M. Kaneko, and A. Nishihara, "LMS-based algorithms with multi-band decomposition of the linear prediction error," Proc. 11th DSP Symposium Tottori, Japan, 1996.

Appendix A: Proof that $\mathcal{E}[\varepsilon(n)] = c\mathcal{E}[e^2(n)]$

A proof that power complementariness of the filter bank used in the proposed method is a sufficient condition for the validity of the relation $\mathcal{E}[\varepsilon(n)] = c\mathcal{E}[e^2(n)]$ is in the following.

Proof: From the definition of the cost function, given in Eqs. (9)–(12), we have

$$\mathcal{E}[\varepsilon(n)] = \sum_{j=1}^{B} \mathcal{E}[\varepsilon^{(j)}(n)] \tag{A·1}$$

$$= \sum_{j=1}^{B} \mathcal{E}\left[\sum_{i=0}^{L^{(j)}-1} \frac{1}{L^{(j)}} \left(e^{(j)}(i)\right)^2\right] \tag{A·2}$$

$$= \sum_{j=1}^{B} \underbrace{\sum_{i=0}^{L^{(j)}-1} \frac{1}{L^{(j)}}}_{1} \frac{1}{2\pi} \int_{-\pi}^{\pi} P_e(\omega) \left|H^{(j)}(e^{j\omega})\right|^2 d\omega \tag{A·3}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} P_e(\omega) \sum_{j=1}^{B} \left|H^{(j)}(e^{j\omega})\right|^2 d\omega, \tag{A·4}$$

where $P_e(\omega)$ is the power spectrum of $e(n)$. If the analysis bank satisfies the power complementary property, that is,

$$\sum_{j=1}^{B} \left|H^{(j)}(e^{j\omega})\right|^2 = c, \tag{A·5}$$

where c is a positive constant, then it can be seen from (A·4) that

$$\mathcal{E}[\varepsilon(n)] = c\mathcal{E}[e^2(n)]. \tag{A·6}$$

Appendix B: Time-Update Relations

In this appendix, time recursive relations for the correlation matrix and the cross-correlation vector are derived and applied to the *matrix inversion lemma* (MIL) to obtain the basic relations underlying the RLS algorithm for multi-band LS methods shown in Sect. 3, which can be seen as a generalization of the "single-window" conventional RLS algorithm [1].

Simple manipulations of (18)–(20) gives

$$\begin{aligned}\boldsymbol{\Phi}_M(n) &= \boldsymbol{\Phi}_M(n-1) \\ &\quad + \sum_{j=1}^{B} \frac{1}{L^{(j)}} \left(\boldsymbol{u}_M^{(j)}(n) \left(\boldsymbol{u}_M^{(j)}(n)\right)^T \right. \\ &\quad \left. - \boldsymbol{u}_M^{(j)}(n-L^{(j)}) \left(\boldsymbol{u}_M^{(j)}(n-L^{(j)})\right)^T\right)\end{aligned} \tag{A·7}$$

which can be rewritten as

$$\begin{aligned}\boldsymbol{\Phi}_M(n) &= \boldsymbol{\Phi}_M(n-1) \\ &\quad + \boldsymbol{U}_M(n,n-L)\boldsymbol{S}\boldsymbol{U}_M^T(n,n-L)\end{aligned} \tag{A·8}$$

where $\boldsymbol{U}_M(n,n-L)$ is an $M\times 2B$ matrix and $\boldsymbol{S}$ is a $2B\times 2B$ matrix as follows

$$\boldsymbol{U}_M^T(n,n-L) = \begin{bmatrix} \left(\boldsymbol{u}_M^{(1)}(n)\right)^T \\ \left(\boldsymbol{u}_M^{(1)}(n-L^{(1)})\right)^T \\ \vdots \\ \left(\boldsymbol{u}_M^{(B)}(n)\right)^T \\ \left(\boldsymbol{u}_M^{(B)}(n-L^{(B)})\right)^T \end{bmatrix} \tag{A·9}$$

$$\boldsymbol{S} = \text{diag}\underbrace{\left[\frac{1}{L^{(1)}}, \frac{-1}{L^{(1)}}, \ldots, \frac{1}{L^{(B)}}, \frac{-1}{L^{(B)}}\right]}_{2B}, \tag{A·10}$$

where diag[$*$] is the representation for a diagonal matrix with main diagonal equal to $*$.

From (22) and (23) we can write the following recursive time-update relation for the cross-correlation vector

$$\begin{aligned}\boldsymbol{\Theta}_M(n) &= \boldsymbol{\Theta}_M(n-1) \\ &\quad + \sum_{j=1}^{B} \frac{1}{L^{(j)}} \left(d^{(j)}(n)\boldsymbol{u}_M^{(j)}(n)\right. \\ &\quad \left. - d^{(j)}(n-L^{(j)})\boldsymbol{u}_M^{(j)}(n-L^{(j)})\right)\end{aligned} \tag{A·11}$$

which may be rewritten in matrix form as

$$\begin{aligned}\boldsymbol{\Theta}_M(n) &= \boldsymbol{\Theta}_M(n-1) \\ &\quad + \boldsymbol{U}_M(n,n-L)\boldsymbol{S}\boldsymbol{d}(n,n-L)\end{aligned} \tag{A·12}$$

where the $2B\times 1$ vector $\boldsymbol{d}(n,n-L)$ is defined as

$$\boldsymbol{d}(n,n-L) = \begin{bmatrix} d^{(1)}(n) \\ d^{(1)}(n-L^{(1)}) \\ \vdots \\ d^{(B)}(n) \\ d^{(B)}(n-L^{(B)}) \end{bmatrix}. \tag{A·13}$$

For simplicity of notation, the subscript "$_M$" is omitted until the end of this appendix.

Application of the MIL [1] in (A·8) gives

$$\begin{aligned}\boldsymbol{P}(n) &= \boldsymbol{\Phi}^{-1}(n) = \boldsymbol{\Phi}^{-1}(n-1) \\ &\quad - \boldsymbol{\Phi}^{-1}(n-1)\boldsymbol{U}(n,n-L) \\ &\quad \left(\boldsymbol{S}^{-1} + \boldsymbol{U}^T(n,n-L)\boldsymbol{\Phi}^{-1}(n-1)\boldsymbol{U}(n,n-L)\right)^{-1} \\ &\quad \boldsymbol{U}^T(n,n-L)\boldsymbol{\Phi}^{-1}(n-1).\end{aligned} \tag{A·14}$$

Defining the gain matrix $\boldsymbol{K}(n)$ as

$$\begin{aligned}\boldsymbol{K}(n) &= \boldsymbol{P}(n-1)\boldsymbol{U}(n,n-L) \\ &\quad \left(\boldsymbol{S}^{-1} + \boldsymbol{U}^T(n,n-L)\boldsymbol{P}(n-1)\boldsymbol{U}(n,n-L)\right)^{-1}\end{aligned} \tag{A·15}$$

we obtain

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) - \boldsymbol{K}(n)\boldsymbol{U}^T(n,n-L)\boldsymbol{P}(n-1) \tag{A·16}$$

which is a recursive update relation for the inverse of the correlation matrix. From (A·15), it can be seen that

$$\boldsymbol{K}(n)\left(\boldsymbol{S}^{-1}+\boldsymbol{U}^T(n,n-L)\boldsymbol{P}(n-1)\boldsymbol{U}(n,n-L)\right)$$
$$=\boldsymbol{P}(n-1)\boldsymbol{U}(n,n-L) \qquad (A\cdot 17)$$
$$\boldsymbol{K}(n)=\boldsymbol{P}(n)\boldsymbol{U}(n,n-L)\boldsymbol{S}. \qquad (A\cdot 18)$$

The recursive relation for the weight vector $\boldsymbol{w}(n)$ is obtained as follows

$$\begin{aligned}\boldsymbol{w}(n) &= \boldsymbol{P}(n)\boldsymbol{\Theta}(n) &(A\cdot 19)\\ &= \boldsymbol{P}(n)\boldsymbol{\Theta}(n-1)\\ &\quad +\boldsymbol{P}(n)\boldsymbol{U}(n,n-L)\boldsymbol{S}\boldsymbol{d}(n,n-L)\\ &= \boldsymbol{w}(n-1)+\boldsymbol{K}(n)\boldsymbol{\alpha}(n) &(A\cdot 20)\end{aligned}$$

where $\boldsymbol{\alpha}(n)$ is a $2B\times 1$ vector called the *a priori* error vector and is defined as

$$\boldsymbol{\alpha}(n)=\boldsymbol{d}(n,n-L)-\boldsymbol{U}^T(n,n-L)\boldsymbol{w}(n-1). \qquad (A\cdot 21)$$

Appendix C: Order-Update Relations

This appendix details the order-update relation satisfied by the correlation matrix and explains the time-update recursion formula for the autocorrelation vector, both used in the derivation of the algorithm described in Sect. 4.

The augmented correlation matrix, as defined in (A·37), satisfies the order-update property, that is,

$$\begin{aligned}\boldsymbol{\Phi}_{M+1}(n) &= \begin{bmatrix}\mathcal{U}_1(n) & \boldsymbol{\Theta}_1^T(n)\\ \boldsymbol{\Theta}_1(n) & \boldsymbol{\Phi}_M(n-1)\end{bmatrix}\\ &= \begin{bmatrix}\boldsymbol{\Phi}_M(n) & \boldsymbol{\Theta}_2(n)\\ \boldsymbol{\Theta}_2^T(n) & \mathcal{U}_2(n-M)\end{bmatrix} &(A\cdot 22)\end{aligned}$$

where

$$\mathcal{U}_1(n)=\sum_{j=1}^{B}\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}\left(u^{(j)}(n-i)\right)^2 \qquad (A\cdot 23)$$

$$\mathcal{U}_2(n-M)=\sum_{j=1}^{B}\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}\left(u^{(j)}(n-i-M)\right)^2 \qquad (A\cdot 24)$$

$$\boldsymbol{\Theta}_1(n)=\sum_{j=1}^{B}\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}u^{(j)}(n-i)\boldsymbol{u}_M^{(j)}(n-i-1) \qquad (A\cdot 25)$$

$$\boldsymbol{\Theta}_2(n)=\sum_{j=1}^{B}\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}u^{(j)}(n-i-M)\boldsymbol{u}_M^{(j)}(n-i). \qquad (A\cdot 26)$$

Since the correlation matrix satisfies the order-update property described in (A·22), the Levinson algorithm can be used to update the AR model filter coefficients from the current value of the autocorrelation vector $\boldsymbol{c}_{M+1}(n)$. In the following, a recursive algorithm to evaluate the autocorrelation vector, which is required as the input for the Levinson RLS algorithm, is obtained. The autocorrelation vector is given by

$$\begin{aligned}\boldsymbol{c}_{M+1}(n) &= [c_0(n),\cdots,c_M(n)]^T &(A\cdot 27)\\ &= [\mathcal{U}_1(n)\ \boldsymbol{\Theta}_1^T(n)]^T &(A\cdot 28)\\ &= \sum_{j=1}^{B}\boldsymbol{c}_{M+1}^{(j)}(n), &(A\cdot 29)\end{aligned}$$

where

$$\boldsymbol{c}_{M+1}^{(j)}(n)=\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}u^{(j)}(n-i)\boldsymbol{u}_{M+1}^{(j)}(n-i). \qquad (A\cdot 30)$$

From (A·30), a recursive formula for the evaluation of the autocorrelation vector may be written as follows

$$\begin{aligned}\boldsymbol{c}_{M+1}^{(j)}(n) &= \boldsymbol{c}_{M+1}^{(j)}(n-1)\\ &\quad+\frac{1}{L^{(j)}}\Big(u^{(j)}(n)\boldsymbol{u}_{M+1}^{(j)}(n)\\ &\quad-u^{(j)}(n-L^{(j)})\boldsymbol{u}_{M+1}^{(j)}(n-L^{(j)})\Big), &(A\cdot 31)\end{aligned}$$

for $j=1,\ldots,B$. Equations (A·29) and (A·31) are used in connection with the Levinson RLS algorithm [8] to derive an RLS algorithm on the order of $(BM+M^2)$ for adaptive AR multi-band LS spectral analysis as shown in Table 1.

Appendix D: Augmented Normal Equations

In the following, augmented normal equations for the forward linear prediction case are derived. From (9)–(12), we can define the cost function to be minimized for multi-band forward linear prediction as

$$\mathcal{F}_M(n)=\sum_{j=1}^{B}\frac{1}{L^{(j)}}\sum_{i=0}^{L^{(j)}-1}\left(f_M^{(j)}(n-i)\right)^2, \qquad (A\cdot 32)$$

where

$$f_M^{(j)}(i)=h^{(j)}(i)*f_M(i),\quad 1\leqq j\leqq B, \qquad (A\cdot 33)$$

$$f_M(i)=\sum_{k=0}^{M}a_k(n)u(i-k),\quad -\infty<i\leqq n, \qquad (A\cdot 34)$$

with $a_0(n)=1$. In the equations above, $f_M(i)$ is the *a posteriori* linear prediction error, $h^{(j)}(n)$, $1\leqq j\leqq B$, represents the analysis filter bank (see Fig. 1 (c)), $a_k(n)$, $k=1,\ldots,M$, are the AR model parameters at time-index n and $u(n)$ is the signal whose power spectrum estimates are been evaluated.

Similarly to conventional LS, the cost function defined in (A·32) is a quadratic function of the AR model parameters $a_k(n)$, and the minimization condition is given as

$$\frac{\delta \mathcal{F}_M(n)}{\delta a_l(n)} = 0, \quad 1 \leqq l \leqq M. \tag{A·35}$$

Substituting (A·32)–(A·34) into (A·35), the following set of so-called augmented normal equations is obtained:

$$\boldsymbol{\Phi}_{M+1}(n)\boldsymbol{a}_M(n) = [\mathcal{F}_M(n), 0, \ldots, 0]^T, \tag{A·36}$$

where the augmented correlation matrix is

$$\boldsymbol{\Phi}_{M+1}(n) = \sum_{j=1}^{B} \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} \boldsymbol{u}_{M+1}^{(j)}(n-i)\left(\boldsymbol{u}_{M+1}^{(j)}(n-i)\right)^T \tag{A·37}$$

with

$$\boldsymbol{u}_{M+1}^{(j)}(i) = \begin{bmatrix} \boldsymbol{u}_M^{(j)}(i) \\ u^{(j)}(i-M) \end{bmatrix} \tag{A·38}$$

$$\boldsymbol{a}_M(n) = [1, \ a_1(n), \cdots, a_M(n)]^T. \tag{A·39}$$

The AR spectral estimates are computed by substituting the solution of (A·36) for $\boldsymbol{a}_M(n)$ into

$$S_{AR,u}(w,n) = \frac{\sigma_v^2}{\left|1 + \sum_{k=1}^{M} a_k(n)e^{-jwk}\right|^2}, \tag{A·40}$$

where σ_v^2 represents the constant power spectrum of the white-noise process applied to the AR model.

Appendix E: Proof of (30) and (33)

In the following, a recursive formula for the sum of weighted error squares, which explains Eqs. (30) and (33) is presented. The subscript "$_M$" is omitted in this appendix for simplicity of notation.

Orthogonality between the estimation error and the estimate of the desired response allows us to write

$$\varepsilon_{min}(n) = \varepsilon_d(n) - \boldsymbol{\Theta}^T(n)\boldsymbol{w}(n) \tag{A·41}$$

where

$$\begin{aligned}\varepsilon_d(n) &= \sum_{j=1}^{B} \frac{1}{L^{(j)}} \sum_{i=0}^{L^{(j)}-1} \left(d^{(j)}(n-i)\right)^2 \\ &= \varepsilon_d(n-1) + \boldsymbol{d}^T(n,n-L)\boldsymbol{S}\boldsymbol{d}(n,n-L).\end{aligned} \tag{A·42}$$

From (A·12) and (26) we have

$$\begin{aligned}\boldsymbol{\Theta}^T(n)\boldsymbol{w}(n) &= (\boldsymbol{\Theta}(n-1) \\ &\quad + \boldsymbol{U}(n,n-L)\boldsymbol{S}\boldsymbol{d}(n,n-L))^T \boldsymbol{w}(n-1) \\ &\quad + \boldsymbol{\Theta}^T(n)\boldsymbol{K}(n)\boldsymbol{\alpha}(n).\end{aligned} \tag{A·43}$$

Substituting these results in (A·41), we get

$$\begin{aligned}\varepsilon_{min}(n) &= \varepsilon_{min}(n-1) \\ &\quad + \boldsymbol{d}^T(n,n-L)\boldsymbol{S}\boldsymbol{\alpha}(n) - \boldsymbol{\Theta}^T(n)\boldsymbol{K}(n)\boldsymbol{\alpha}(n).\end{aligned} \tag{A·44}$$

Observing that

$$\begin{aligned}\boldsymbol{\Theta}^T(n)\boldsymbol{K}(n) &= \boldsymbol{\Theta}^T(n)\boldsymbol{\Phi}^{-1}(n)\boldsymbol{U}(n,n-L)\boldsymbol{S} \\ &= \boldsymbol{w}^T(n)\boldsymbol{U}(n,n-L)\boldsymbol{S}\end{aligned} \tag{A·45}$$

we finally obtain

$$\varepsilon_{min}(n) = \varepsilon_{min}(n-1) + \boldsymbol{e}^T(n)\boldsymbol{S}\boldsymbol{\alpha}(n). \tag{A·46}$$

Appendix F: Complementary Derivations Related to (28)–(39)

This appendix complements the material presented in Sect. 5 and the related appendices with respect to the derivation of the recursive relations (28)–(39). Backward linear prediction related equations are not explicitly derived in the following since they may be obtained by trivial analogy with respective relations for forward linear prediction. The $2B \times 1$ forward prediction error vector is

$$\begin{aligned}\boldsymbol{f}_M(i) &= \begin{bmatrix} f^{(1)}(i) \\ f^{(1)}(i-L^{(1)}) \\ \vdots \\ f^{(B)}(i) \\ f^{(B)}(i-L^{(B)}) \end{bmatrix} \\ &= \boldsymbol{u}_M(i,i-L) - \boldsymbol{U}_M^T(i-1,i-L-1)\boldsymbol{w}_M(n)\end{aligned} \tag{A·47}$$

for $1 \leqq i \leqq n$, where

$$\boldsymbol{u}_M(i,i-L) = \begin{bmatrix} u^{(1)}(i) \\ u^{(1)}(i-L^{(1)}) \\ \vdots \\ u^{(B)}(i) \\ u^{(B)}(i-L^{(B)}) \end{bmatrix}. \tag{A·48}$$

Also, for

$$\boldsymbol{a}_M(n) = \begin{bmatrix} 1 \\ -\boldsymbol{w}_M(n) \end{bmatrix} \tag{A·49}$$

we may rewrite the *a posteriori* forward prediction error vector (A·47) as

$$\boldsymbol{f}_M(i) = \boldsymbol{U}_{M+1}^T(i,i-L)\boldsymbol{a}_M(n). \tag{A·50}$$

Eqs. (28) and (29) for the forward filter coefficients update and *a priori* forward prediction vector, respectively, may be easily obtained by application of the MIL as shown in Appendix B.

The order-update property of the correlation matrix given in (A·22) enables us to write[1]

$$\boldsymbol{\Phi}_{M+1}^{-1}(n) = \begin{bmatrix} 0 & \mathbf{0}_{1\times M} \\ \mathbf{0}_{M\times 1} & \boldsymbol{\Phi}_M^{-1}(n-1) \end{bmatrix} + \frac{1}{\mathcal{F}_M(n)}\boldsymbol{a}_M(n)\boldsymbol{a}_M^T(n) \quad \text{(A·51)}$$

which post-multiplied by $\boldsymbol{U}(n, n-L)\boldsymbol{S}$ gives the time-order update formula for the gain matrix shown in (34). For the conversion matrix formulae we first premultiply (26) by $\boldsymbol{U}^T(n, n-BL)$ and get

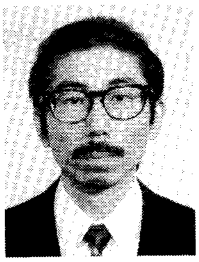
$$\boldsymbol{e}(n) = \boldsymbol{\Gamma}(n)\boldsymbol{\alpha}(n) \quad \text{(A·52)}$$

where we define the $2B\times 2B$ conversion matrix $\boldsymbol{\Gamma}(n)$ as

$$\boldsymbol{\Gamma}(n) = \boldsymbol{I} - \boldsymbol{U}^T(n, n-BL)\boldsymbol{K}(n). \quad \text{(A·53)}$$

For the forward prediction case, (36) can be obtained by a procedure similar to the one used to derive (A·52). Finally, pre-multiplying (34) by $\boldsymbol{U}^T(n, n-L)$, we obtain the recursive formula for the conversion matrix given in (38).

Fernando Gil V. Resende Jr. received the B.Sc. degree from Instituto Militar de Engenharia, Brazil, in 1990 and the M.Sc. degree from Tokyo Institute of Technology, Japan, in 1994, both in electrical and electronic engineering. Since 1994 he has been working towards his Ph.D. degree at Tokyo Institute of Technology. His research interests include adaptive filtering theory, spectral estimation and speech processing.



Keiichi Tokuda received the B.E. degree in electrical and electronic engineering from the Nagoya Institute of Technology, Nagoya, Japan, the M.E. and D.E. degrees in information processing from the Tokyo Institute of Technology, Tokyo, Japan, in 1984, 1986, and 1989, respectively. From 1989 to 1996 he was a Research Associate at the Department of Electrical and Electronic Engineering, Tokyo Institute of Technology. Since 1996 he has been with the Department of Computer Science, Nagoya Institute of Technology as Associate Professor. His research interests include speech coding, speech synthesis and recognition, and signal processing. He is a member of IEEE and ASJ.

Mineo Kaneko received the B.E., M.E. and Dr.E. degrees in electrical and electronic engineering from Tokyo Institute of Technology in 1981, 1983 and 1986, respectively. In 1986 he joined the Department of Electrical and Electronic Engineering, Tokyo Institute of Technology as a Research Associate. From 1992 to 1996, he was an Associate Professor in the same Department of Tokyo Institute of Technology. Since 1996 he has been an Associate Professor in the School of Information Science, Japan Advanced Institute of Science and Technology, Hokuriku. His research interests include high-speed and fault tolerant VLSI signal processing, computer aided design of integrated circuits and wafer scale integration. He received best paper awards from IEEE APCCAS'92 and APCCAS'94 in 1992 and 1994, respectively. He is a member of IEEE.

Akinori Nishihara was born in Fukuoka, Japan, on February 26, 1951. He received the B.E., M.E. and Dr.Eng. degrees in electronics from Tokyo institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is now Professor of the Center for Research and Development of Educational Technology. His main research interests are in filter design, 1D and multi-D signal processing, and modern applications of classical circuit theory. From 1990 to 1994 he served as an Associate Editor of the IEICE Trans. Fundamentals, and is now serving as an Associate Editor of the IEEE Trans. Circuits & Systems II. Dr. Nishihara is a member of IEEE, EURASIP, ECS and JET.