AII 20

INTERNATIONAL JOURNAL OF CIRCUIT THEORY AND APPLICATIONS
Int. J. Circ. Theor. Appl. 2000; **28**:481–500

On orthogonal realizations for adaptive IIR filters

J. Cousseau[1,*,†], P. S. R. Diniz[2], G. Sentoni[1] and O. Agamennoni[1]

[1] *Department of Electrical Engineering, Universidad Nacional del Sur, Av. Alem 1253, Bahia Blanca 8000, Argentina*
[2] *Prog. de Engenharia Elétrica e Depto. de Eletrônica, COPPE/EE/Federal University of Rio de Janeiro, C.P. 68504, Rio de Janeiro 21945970, Brazil*

SUMMARY

Convergence speed is one of the main concerns in adaptive IIR filters. Fast convergence can be closely related to adaptive filter realization. However, with the exception of the lattice realization that is based on the nice properties of Szëgo orthonormal polynomials, no other adaptive IIR filter realization using orthonormal characteristics seems to be extensively studied in the literature. Furthermore, many orthogonal realizations for adaptive FIR filters, that are particularly suitable for rational modelling, have been proposed in the past years. Since rational orthogonal basis functions are a powerful tool for efficient system representation they seem attractive for adaptive IIR filters.

In this paper, we present some theoretical results related to the properties of a generalized orthonormal realization when used for mean-square output error minimization in a system identification application. One result is related to the low computational complexity of the updating gradient algorithm when some properties of the orthonormal realization are used. An additional result establishes conditions for the stationary points of the proposed updating algorithm.

In order to confirm the expected performance of the new realization, some simulations and comparisons with competing realizations in terms of computational complexity and convergence speed are presented. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: adaptive IIR filters; orthogonal realizations

1. INTRODUCTION

In order to achieve a suitable representation of physical systems, orthonormal representations are very powerful tools in system identification problems [1]. With this purpose a number of such representations were introduced in the past years. Many of the applications of these orthonormal representations have been oriented to improve the performance of FIR adaptive algorithms [2]. Of particular interest are those cases where the dynamics of the plant is partially known, i.e. the dominant poles of the plant can be measured or approximated. In those cases, in addition to the improvement in the conditioning of the covariance matrix related to the

* Correspondence to: J. Cousseau, Department of Electrical Engineering, Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca 8000, Argentina
† E-mail: iecousse@criba.edu.ar

Received 6 June 1999
Revised 25 April 2000

coefficient estimation, the proper choice of the orthonormal basis leads to a model with a minimum number of FIR (feedforward) coefficients [3].

Adaptive IIR filters are considered in this paper as an efficient alternative realization to the adaptive FIR filters. As a consequence, the orthonormal realizations discussed here are oriented to applications of adaptive IIR filters.

Among the many open problems, convergence speed and computational complexity are main concerns related to the wide use of adaptive IIR filters. Suitable updating algorithms for adaptive IIR filters were studied in the past decades and are today object of intensive research. Among a wide variety of algorithms it is worth mentioning: prediction error methods [4], equation error methods [5] and Steiglitz-McBride methods [6–8]. Also, interesting mixed methods have been proposed [9].

However, the main focus of this work is to address efficient adaptive IIR realizations, in the aspects related to convergence speed and computational complexity. As discussed in the paper, a realization that uses orthogonalization properties seems to be an interesting answer to achieve a good tradeoff between computational complexity and convergence speed.

In the recent past alternative efficient realizations have been proposed that either employ or approximate the structural orthogonalization idea. An example of this is the frequency-domain adaptive pole-zero filter [10]. In spite of being very promising from the point of view of obtained results for adaptive FIR filters, the preprocessing required by this realization leads to a high computational complexity. Meanwhile, other efficient non-orthogonal realizations have been studied and characterized, such as the parallel [11, 12] and cascade realizations [12].

The main drawback of the parallel realization is the introduction of equivalent local minima, placed at distinct subregions of the parameter space. These subregions are separated by boundaries representing reduced-order manifolds. A low complexity parallel realization with fast convergence speed, was proposed in Reference [13]. This realization utilizes appropriately configured second-order sections, in order to reduce the number of iterations required by the algorithm to move away from the reduced-order manifolds. The key idea behind the good performance of this approach is related to the different gradients of the various sections, even if the poles of the sections are at the same position. As analysed in Reference [11], the potential ill-conditioning of the parallel realization can be avoided with the introduction of a preprocessing or transform [10, 14]. In addition, this preprocessing provides some degree of orthogonalization among the sections of the parallel realization. This can be used in order to reduce the computational complexity of the Gauss–Newton-type algorithm without a noticeable loss in performance. However, no explicit orthogonalization is used in this realization.

Among the orthogonal realizations, the most widely known is the lattice realization [15], which has been extensively studied in both time-invariant [16] and time-variant cases [8]. Only recently it was found that the lattice realization can be implemented with $O(N)$ complexity [17]. Several variants of the lattice realization were proposed [8, 18]. These realizations use the Szëgo polynomials as orthogonal support. Structural stability and robustness properties of the lattice filter [8, 19] indicate that this realization is a suitable choice for many applications.

Inspired by the ideas of orthogonality of backward errors on stationary AR modelling, well explored in the lattice realization, and the preprocessing in the frequency-domain adaptive filter, the realization proposed is generated through a set of discrete orthogonal rational functions, in such a way that fast convergence speed can be expected. Rational orthonormal basis functions [3] are used in this work with two main purposes. From one point of view, orthonormal representation of rational systems allows the use of some powerful tools related to analytical spaces [8,20], mainly in $\mathscr{H}_2$ space. On the other hand, it will be shown that a general orthonormal representation of adaptive IIR filters [3] leads to an efficient realization that has computational complexity comparable with both the direct form and the lattice realizations. The family of proposed realizations are a generalized version of the one proposed in Reference [21]. Some variant of first-order sections or second-order sections are discussed. Indeed, stationary point study for mean-square output error minimization indicates that interesting gradient simplification can be introduced without modifying the stationary points.

The main contributions of this work are the following:

- Present a new adaptive IIR filter realization that uses a rational orthogonal basis support. Its relationship with Laguerre, Kautz and other realizations is discussed.
- Present results related to the stationary points that allow convenient simplifications of a MSOE updating algorithm.
- Provide a set of computer simulations to illustrate the expected performance of the proposed realization and to compare convergence speed with other realizations.

The paper is organized as follows: In Section 2 a general discussion on orthogonal rational basis functions is presented. Specific results that will be useful for an efficient realization are also presented. In Section 3 two adaptive IIR filter realizations are introduced. They are based on first-order and second-order rational orthogonal basis functions. In Section 4 an efficient realization of these adaptive filters is discussed. In Section 5, a discussion of the results and some computer simulations examples to confirm the expected performance are included. In particular, some specific comparisons with other algorithms in terms of convergence speed are presented. Finally, in Section 6 some conclusions can be found.

2. ON ORTHONORMAL BASIS FOR RATIONAL SYSTEMS

The construction of rational orthonormal basis functions for the representation of dynamic systems is a well-known issue in the area of system identification [22,23]. Assume that a proper stable linear time-invariant dynamic system is described by $y(n) = H(z)\,u(n)$, where $u(n)$ is the input signal, $y(n)$ is the output signal. The transfer function $H(z) = B(z)/A(z)$, with $B(z) = b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}$ and $A(z) = 1 + a_1 z^{-1} + \cdots + a_M z^{-M}$ is assumed to be asymptotically stable so that the system belongs to $\mathscr{H}_2$.

Now suppose that $\{\mathscr{F}_k(z)\}_{k=0}^{\infty}$ is an orthonormal family of functions in $\mathscr{H}_2$ and that $H(z) = \sum_{k=0}^{\infty} v_k \mathscr{F}_k(z)$. These functions can be constructed if we consider the state-space representation of $G(z)$ given by

$$\begin{bmatrix} \mathbf{x}(n+1) \\ y(n) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c} & d \end{bmatrix} \begin{bmatrix} \mathbf{x}(n) \\ u(n) \end{bmatrix} \tag{1}$$

or, alternatively

$$\begin{bmatrix} \mathbf{x}(n+1) \\ w(n) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{g} & v_0 \end{bmatrix} \begin{bmatrix} \mathbf{x}(n) \\ u(n) \end{bmatrix} \tag{2}$$

where $w(n)$ is an auxiliar variable. It is straightforward to verify that a filter H with state-space realization $(\mathbf{A}, \mathbf{b}, \mathbf{c}, d)$ is orthogonal and all-pass if and only if the matrix related to representation (1) or Equation (2) is an orthonormal matrix [19]. If the matrix of Equation (1) is in upper Hessemberg form [8], then it might describe a lattice filter. The orthonormal condition is satisfied if, for example, the state-space realization is input balanced [19] or if $\mathbf{APA}^{\mathrm{T}} + \mathbf{bb}^{\mathrm{T}} = \mathbf{P}$ has $\mathbf{P} = \mathbf{I}$ as solution.

Under these assumptions the basis functions can be chosen as [8, 24]

$$\mathscr{F}_k(z) = \begin{cases} \mathbf{e}_k^{\mathrm{T}}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, & k = 0, \ldots, M-1 \\ z^{M-k}V(z), & k \geqslant M \end{cases}$$

where $\mathbf{e}_k = [0, \ldots, 1, \ldots 0]^{\mathrm{T}}$ is a unit vector with a one at kth position and

$$V(z) = \left(\frac{z^{-M}A(z^{-1})}{A(z)}\right) = v_0 + \sum_{k=1}^{\infty} \mathbf{g}\mathbf{A}^{k-1}\mathbf{b}z^{-k}$$

is an all-pass filter.

Very important to our purposes is the fact that orthogonality and all-pass characteristics are preserved if several different connections of such filters are performed in order to build an orthonormal family of functions. Of particular interest is the preserving orthogonality of the serial or cascade connections [19], i.e. if $H_1(z)$ and $H_2(z)$ are two orthogonal all-pass filters with state-space realizations $(\mathbf{A}_1, \mathbf{b}_1, \mathbf{c}_1, d_1)$ and $(\mathbf{A}_2, \mathbf{b}_2, \mathbf{c}_2, d_2)$, respectively, then the serial connection $H_1(z)H_2(z)$ of the two filters is also orthogonal and all-pass. If $\mathbf{x}_k(n)$ are the states of the filter $H_k(z)$ in the cascade connection, with $\mathbf{x}_k(n) = [x_{k,1}(n), \ldots, x_{k,n_k}(n)]^{\mathrm{T}}$, where n_k is the McMillan degree of the all-pass filter $H_k(z)$, then the transfer functions $\mathscr{F}_{k,l}(z)$, defined by $x_{k,l}(n) = \mathscr{F}_{k,l}(q^{-1})u(n)$, will form an orthonormal family so that $\langle \mathscr{F}_{i,j}, \mathscr{F}_{k,l} \rangle = \delta_{i,k} \times \delta_{j,l}$.

A general description of orthogonal basis is based on the inner–outer factorization of $\mathscr{H}_2$ [20]. It contemplates the use of an *all-pass generator function* [24] or *all-pass completion function* [8], $V(z)$.

We consider here a basis function formed as follows:

$$\mathscr{F}_k(z) = \frac{N_k(z)}{D_k(z)} \prod_{i=1}^{k-1} \frac{\bar{D}_i(z)}{D_i(z)} \tag{3}$$

where $D_k(z)$ are P-order monic polynomials, $\bar{D}_k(z) = z^P D_k(z^{-1})$ and $N_k(z)$, depends on the coefficients of $D_k(z)$ and are chosen to maintain the orthogonality of the basis functions $F_k(z)$, i.e.

$$\|\mathscr{F}_k(z)\| = 1, \qquad \langle \mathscr{F}_k(z), \mathscr{F}_l(z) \rangle = \delta_{k,l}$$

In this particular case, $V(z) = \prod_{i=1}^{K} \bar{D}_i(z)/D_i(z)$, where $K = M/P - 1$, and M is the adaptive-filter order. In the following section, two particular sections are analyzed in this work, namely first-order polynomials, $D_k(z) = 1 + \beta_k z^{-1}$, and second-order polynomials, $D_k(z) = 1 + \beta_{1_k} z^{-1} + \beta_{2_k} z^{-2}$.

3. A GENERAL ORTHONORMAL REALIZATION

Some variants of the orthonormal realizations using identical all-pass filters were considered in References [22, 24]. Some other realizations with different all-pass filters were studied in Reference [3]. In spite of this basic difference, all of them use the ideas discussed in the previous section to build a family of orthogonal functions that approximates a rational function. If the cascade structure is built of identical first-order sections, the Laguerre basis is obtained [1] and if identical second-order sections are used, the two-parameter Kautz functions are obtained [23].

The proposed IIR orthonormal realization is based on the basis functions introduced in the previous section. In particular, the output of the IIR filter $y(n)$ can be written as

$$y(n)=\left[\sum_{k=1}^{K} v_k \mathscr{F}_k(z)\right] u(n)+v_0(n)u(n)$$
$$=\sum_{k=1}^{K} y_k(n)+v_0(n)u(n) \tag{4}$$

where both $v_k(n)$ and $\mathscr{F}_k(z)$ allow a complete characterization of the filter realization.

In order to use the proposed IIR filter realization in the minimization of the mean-squared output error (MSOE) $E[e^2(n)]$, we consider the prediction error $e(n)$ that can be expressed as follows:

$$e(n)=d(n)-y(n)=d(n)-\hat{H}(z)u(n)$$
$$=d(n)-\left(\sum_{k=1}^{K} y_k(n)+v_0(n)u(n)\right)$$

where $d(n)$ is the reference signal. For this minimization we can use the updating equation of the stochastic gradient algorithm that is given by

$$\theta(n+1)=\theta(n)-\mu\nabla(n) \tag{5}$$

where $\theta(n)$ is the coefficient vector, μ is the convergence factor and $\nabla(n)$ is a suitable estimate of the gradient vector.

In the following subsections, we analyse two particular cases of the proposed general realization for an adaptive IIR filter, employing first- and second-order sections. These realizations will be used for MSOE minimization, then as additional result gradient computation is analysed. Although similar from the point of view of the basis functions some particular differences related to orthogonalization and gradient computation lead us to present a separate study.

3.1. First-order sections

In this case $K=M$, and we assume for simplicity that the model has real poles (the complex pole case can be solved using the second-order sections). Then

$$B_k(z)=\frac{N_k(z)}{D_k(z)}\prod_{i=1}^{k-1}\frac{\bar{D}_i(z)}{D_i(z)}$$

$$D_k(z) = 1 + \beta_k z^{-1}$$

$$N_k(z) = \alpha_k = \sqrt{1 - |\beta_k|^2}$$

the coefficient vector is

$$\theta(n) = [v_0(n) v_1(n) \cdots v_M(n)\ \beta_1(n) \cdots \beta_M(n)]^{\mathrm{T}}$$

and $\nabla(n)$ is

$$\nabla(n) = -2e(n) \left[\frac{\partial y(n)}{\partial v_0} \cdots \frac{\partial y(n)}{\partial v_M} \frac{\partial y(n)}{\partial \beta_1} \cdots \frac{\partial y(n)}{\partial \beta_M} \right]^{\mathrm{T}} \tag{6}$$

Using the instantaneous squared value of the output error as an estimate of $E[e^2(n)]$, the components of the gradient vector are given by

$$\frac{\partial y(n)}{\partial v_k} = B_k(n) u(n) \tag{7}$$

$$\frac{\partial y(n)}{\partial \beta_j} = \sum_{k=1}^{M} v_k(n) \frac{\partial y_k(n)}{\partial \beta_j} \tag{8}$$

then after some reordering

$$\frac{\partial y(n)}{\partial \beta_j} = v_j C_j(z) u(n) - \frac{z^{-1}}{D_j(z)} \sum_{i=j}^{M} v_i B_i(z) u(n)$$

$$+ \frac{z^{-1}}{D_j(z)} \sum_{i=j}^{M-j} v_i \frac{N_{i+1}(z)}{D_{i+1}(z)} \prod_{k=1, k \neq j}^{i} \frac{\bar{D}_k(z)}{D_k(z)} u(n) \tag{9}$$

where $C_j(z) = (N_j'(z)/D_j(z)) \prod_{i=1}^{j-1} \bar{D}_i(z)/D_i(z)$ and $N_j' = \partial N_j(z)/\partial \beta_j$ (note that $N_j(z)$ depends on β_j only).

Remark.

1. If we use the definition of the basis function $B_k(z)$, for $k = 1, \ldots, M$, then the last term of the equation above, defined as $B_k^{M-j}(z)$ is a basis function of an $(M-j)$-dimensional basis space, i.e. with $k = 1, \ldots, M-j$. Thus, Equation (9) can be written as

$$\frac{\partial y(n)}{\partial \beta_j} = v_j C_j(z) u(n) - \frac{z^{-1}}{D_j(z)} \sum_{i=j}^{M} v_i B_i(z) u(n) + \frac{z^{-1}}{D_j(z)} \sum_{i=j}^{M-j} v_i B_i^{M-j}(z) u(n) \tag{10}$$

2. Formally, by assuming the usual small step convergence approximation, Equations (7), (10) and (5) describe the updating the first-order sections orthogonal adaptive filtering algorithm.

3. This constrained-real poles adaptive IIR filter is mainly proposed to show the particular properties of this kind of orthogonal realizations.
4. Equation (10) is not an efficient form to obtain the gradient $\partial y(n)/\partial\beta_j$ but, as will be shown in the next section, is useful to introduce simplified alternative computations.
5. Assuming $u(n)$ white noise the stationary points of the proposed algorithm are given by

$$\langle B_k(n),\ (H(z)-\hat{H}(z))\rangle = 0 \tag{11}$$

$$\left\langle v_j C_j(z) - \frac{z^{-1}}{D_j(z)}\sum_{i=j}^{M} v_i B_i(z), (H(z)-\hat{H}(z))\right\rangle + \left\langle \frac{z^{-1}}{D_j(z)}\sum_{i=j}^{M-j} v_i B_i^{M-j}(z), (H(z)-\hat{H}(z))\right\rangle = 0 \tag{12}$$

3.2. *Second-order sections*

In this case $K = M/2$, and for second-order sections, the basis functions can be defined as follows:

$$F_k(z) = \begin{cases} \dfrac{N_k(z)}{D_k(z)}\displaystyle\prod_{i=1}^{k-1}\frac{\bar{D}_i(z)}{D_i(z)} & \text{for } k \text{ odd} \\ \dfrac{N_k'(z)}{D_k(z)}\displaystyle\prod_{i=1}^{k-1}\frac{\bar{D}_i(z)}{D_i(z)} & \text{for } k \text{ even} \end{cases} \tag{13}$$

$$D_k(z) = 1 + \beta_{1_k} z^{-1} + \beta_{2_k} z^{-2} \tag{14}$$

$$N_k(z) = \alpha_{1_k} + \alpha_{2_k} z^{-1} \tag{15}$$

$$N_k'(z) = \alpha'_{1_k} + \alpha'_{2_k} z^{-1} \tag{16}$$

where

$$\begin{aligned} \alpha_{1_k} &= -\tfrac{1}{2}\sqrt{c}(\sqrt{a}+\sqrt{b}), \quad \alpha'_{1_k} = \alpha_{2_k} \\ \alpha_{2_k} &= \tfrac{1}{2}\sqrt{c}(\sqrt{a}-\sqrt{b}), \quad \alpha'_{2_k} = \alpha_{1_k} \end{aligned} \tag{17}$$

with $a = 1-\beta_{1_k}+\beta_{2_k}$, $b = 1+\beta_{1_k}+\beta_{2_k}$ and $c = 1-\beta_{2_k}$. Note that the definitions of α_{1_k}, α_{2_k}, α'_{1_k} and α'_{2_k} are required in order to maintain the basis functions $F_k(z)$ orthonormal, i.e. $\|F_k(z)\| = 1$, $\langle F_k(z), F_l(z)\rangle = \delta_{k,l}$. As stated in Reference [3], the procedure to obtain the coefficients of $N_k(z)$ is well defined. In particular, the procedure is an extension of the first-order basis functions discussed in the previous subsection when complex poles are considered. This can be seen as a combination of two successive first-order basis functions $B_k(z)$ and $B_{k+1}(z)$, $k = 1,\ldots,M-1$, to obtain the associated second-order basis functions, $F_{2k-1}(z)$ and $F_{2k}(z)$, $k = 1, \ldots, M/2 - 1$. This can be written as

$$\begin{bmatrix} F_{2k-1} \\ F_{2k} \end{bmatrix} = \begin{bmatrix} d_0 & d_1 \\ d'_0 & d'_1 \end{bmatrix} \begin{bmatrix} B_k \\ B_{k+1} \end{bmatrix}$$

where the constants d_0, d'_0, d_1 and d'_1 are complex numbers. Then the unit norm constraint on $F_{2k-1}(z)$ and $F_{2k}(z)$ leads to $|d_0|^2 + |d_1|^2 = 1$ and the orthogonality leads to $d_1 d_1'^* + d_0 d_0'^* = 0$. Note that

$$d_0 = \frac{\alpha_{1_k} + \beta_k^* \alpha_{2_k}}{1 - \beta_k^{2*}}, \qquad d_1 = \frac{\beta_k^* \alpha_{1_k} + \alpha_{2_k}}{1 - \beta_k^{2*}}$$

$$d'_0 = \frac{\alpha'_{1_k} + \beta_k \alpha'_{2_k}}{1 - \beta_k^2}, \qquad d'_1 = \frac{\beta_k \alpha'_{1_k} + \alpha'_{2_k}}{1 - \beta_k^2}$$

where β_k is the complex pole considered. In terms of the unit norm constraint, assuming α_{1_k}, α_{2_k} are real, we obtain

$$(1 + |\beta_k|^2)(\alpha_{1_k}^2 + \alpha_{2_k}^2) + 2(\beta_k^* + \beta_k)\alpha_{1_k}\alpha_{2_k} = 1 + |\beta_k|^4 - (\beta_k^2 + \beta_k^{2*}) \tag{18}$$

Considering the specific solution of Reference [3] for the other orthogonality constraint we can relate α_{1_k} and α_{2_k} with α'_{1_k} and α'_{2_k}, as

$$\begin{bmatrix} \alpha'_{1_k} \\ \alpha'_{2_k} \end{bmatrix} = \frac{1}{\sqrt{1-\gamma}} \begin{bmatrix} \gamma & 1 \\ -1 & -\gamma \end{bmatrix} \begin{bmatrix} \alpha_{1_k} \\ \alpha_{2_k} \end{bmatrix}$$

where $\gamma = (\beta_k + \beta_k^*)/(1 + |\beta_k|^2)$. Finally a possible solution for Equation (18) is

$$\alpha_{1_k} = -\sqrt{1 - \gamma + |\beta_k|^2} - \sqrt{1 + \gamma + |\beta_k|^2}$$

$$\alpha_{2_k} = \sqrt{1 - \gamma + |\beta_k|^2} - \sqrt{1 + \gamma + |\beta_k|^2}$$

that in terms of the second-order denominator coefficients gives the result of Equation (17).

As can be noted, with this particular choice the normalization is not computationally expensive (except for a square-root table look up, it is directly related to the stability check of the second-order section).

In order to use the proposed IIR filter realization in the minimization of the mean-squared output error (MSOE) $E[e^2(n)]$, we consider the prediction error $e(n)$ that can be expressed as follows:

$$e(n) = d(n) - \left[\sum_{k=1}^{M/2} (v_{2k-1} F_{2k-1}(z) u(n) + v_{2k} F_{2k}(z) u(n)) + v_0(n) u(n) \right]$$

Using the updating Equation (5), we define

$$\theta(n)=[v_0(n)v_1(n)\cdots v_M(n)\ \beta_{1_1}(n)\beta_{2_1}(n)\cdots\beta_{1_{M/2}}(n)\beta_{2_{M/2}}(n)]^{\mathrm{T}}$$

as the coefficient vector and $\nabla(n)$ is a suitable estimate of the gradient vector given by

$$\nabla(n)=-2e(n)\left[\frac{\partial y(n)}{\partial v_0}\cdots\frac{\partial y(n)}{\partial v_M}\frac{\partial y(n)}{\partial\beta_{1_1}}\frac{\partial y(n)}{\partial\beta_{2_1}}\cdots\frac{\partial y(n)}{\partial\beta_{1_{M/2}}}\frac{\partial y(n)}{\partial\beta_{2_{M/2}}}\right]^{\mathrm{T}} \tag{19}$$

Using the instantaneous squared value of the output error as an estimate of $E[e^2(n)]$, the components of the gradient vector are given by

$$\frac{\partial y(n)}{\partial v_k}=\frac{\partial y_k(n)}{\partial v_k}=F_k(n)u(n) \tag{20}$$

$$\frac{\partial y(n)}{\partial\beta_{1_j}}=\sum_{k=1}^{M/2}\left(v_{2k-1}(n)\frac{\partial F_{2k-1}(n)}{\partial\beta_{1_j}}+v_{2k}(n)\frac{\partial F_{2k}(n)}{\partial\beta_{1_j}}\right)u(n) \tag{21}$$

$$\frac{\partial y(n)}{\partial\beta_{2_j}}=\sum_{k=1}^{M/2}\left(v_{2k-1}(n)\frac{\partial F_{2k-1}(n)}{\partial\beta_{2_j}}+v_{2k}(n)\frac{\partial F_{2k}(n)}{\partial\beta_{2_j}}\right)u(n) \tag{22}$$

$k=0,\ldots,M$ $(F_0(z)=1)$, $j=1,\ldots,M/2$. After some straightforward calculations and ordering, Equations (21) and (22) can be written in a similar form that for the first-order sections as

$$\begin{aligned}\frac{\partial y(n)}{\partial\beta_{1_j}}&=(v_{2j-1}G^1_{2j-1}(z)u(n)+v_{2j}G^1_{2j}(z)u(n))\\&-\frac{z^{-1}}{D_j(z)}\sum_{k=j}^{M/2}(v_{2k-1}F_{2k-1}(z)u(n)+v_{2k}F_{2k}(z)u(n))\\&+\frac{z^{-1}}{D_j(z)}\sum_{k=j}^{M/2-j}(v_{2k-1}F^{M/2-j}_{2k-1}(z)u(n)+v_{2k}F^{M/2-j}_{2k}(z)u(n))\end{aligned} \tag{23}$$

$$\begin{aligned}\frac{\partial y(n)}{\partial\beta_{2_j}}&=(v_{2j-1}G^2_{2j-1}(z)u(n)+v_{2j}G^2_{2j}(z)u(n))\\&-\frac{z^{-2}}{D_j(z)}\sum_{k=j}^{M/2}(v_{2k-1}F_{2k-1}(z)u(n)+v_{2k}F_{2k}(z)u(n))\\&+\frac{1}{D_j(z)}\sum_{k=j}^{M/2-j}(v_{2k-1}F^{M/2-j}_{2k-1}(z)u(n)+v_{2k}F^{M/2-j}_{2k}(z)u(n))\end{aligned} \tag{24}$$

where $j = 1, \ldots, M/2-1$. Also, the notation of orthogonal basis functions of $M/2-j$-dimensional basis space, i.e. $F_k^{M/2-j}$, $k = 1, \ldots, M/2-j$ was introduced. In addition

$$G_j^1(z) = \begin{cases} \dfrac{\partial N_j(z)}{\partial \beta_{1_j}} \left(\prod_{i=1}^{j} \dfrac{\bar{D}_{i-1}(z)}{D_i(z)} \right) & \text{for } j \text{ odd} \\ \dfrac{\partial N_j'(z)}{\partial \beta_{1_j}} \left(\prod_{i=1}^{j} \dfrac{\bar{D}_{i-1}(z)}{D_i(z)} \right) & \text{for } j \text{ even} \end{cases}$$

$$G_j^2(z) = \begin{cases} \dfrac{\partial N_j(z)}{\partial \beta_{2_j}} \left(\prod_{i=1}^{j} \dfrac{\bar{D}_{i-1}(z)}{D_i(z)} \right) & \text{for } j \text{ odd} \\ \dfrac{\partial N_j'(z)}{\partial \beta_{2_j}} \left(\prod_{i=1}^{j} \dfrac{\bar{D}_{i-1}(z)}{D_i(z)} \right) & \text{for } j \text{ even} \end{cases}$$

with $\partial N_k(z)/\partial \beta_{1_j} = \partial \alpha_{1_k}/\partial \beta_{1_j} + (\partial \alpha_{2_k}/\partial \beta_{1_j}) z^{-1}$ and $(\partial N_k(z)/\partial \beta_{2_j}) \partial \alpha_{1_k}/\partial \beta_{2_j} + (\partial \alpha_{2_k}/\partial \beta_{2_j}) z^{-1}$ for k even, $\partial N_k(z)/\partial \beta_{1_j} = \partial \alpha'_{1_k}/\partial \beta_{1_j} + (\partial \alpha'_{2_k}/\partial \beta_{1_j}) z^{-1}$ and $\partial N_k(z)/\partial \beta_{2_j} = \partial \alpha'_{1_k}/\partial \beta_{2_j} + (\partial \alpha'_{2_k}/\partial \beta_{2_j}) z^{-1}$ for k odd. Also,

$$\frac{\partial \alpha_{1_k}}{\partial \beta_{1_j}} = \frac{1}{2} \frac{\alpha_{2_k}}{\sqrt{ab}}, \qquad \frac{\partial \alpha_{2_k}}{\partial \beta_{1_j}} = \frac{1}{2} \left(\frac{\alpha_{1_k}}{c} - \frac{\alpha_{1_k}}{\sqrt{ab}} \right)$$

$$\frac{\partial \alpha_{1_k}}{\partial \beta_{2_j}} = \frac{1}{2} \frac{\alpha_{1_k}}{\sqrt{ab}}, \qquad \frac{\partial \alpha_{2_k}}{\partial \beta_{1_j}} = -\frac{1}{2} \left(\frac{\alpha_{2_k}}{c} - \frac{\alpha_{2_k}}{\sqrt{ab}} \right)$$

and the computations of $\partial \alpha'_{1_k}/\partial \beta_{1_j}$, $\partial \alpha'_{1_k}/\partial \beta_{2_j}$, $\partial \alpha'_{2_k}/\partial \beta_{1_j}$ and $\partial \alpha'_{2_k}/\partial \beta_{2_j}$ come from the equations above.

Remark.

- Although Equations (23) and (24) are not suitable for direct implementation, they are useful for the analysis of possible simplifications in the updating algorithm.
- The second term of the right-hand side of both Equations (23) and (24) represents an important factor of increase in computational complexity, if the exact gradient so obtained is implemented. However, preliminary results obtained by heuristic simplifications of the above formulae have shown that considerable simplifications can be introduced in the previous algorithm without impairing the minimization of the MSOE [25]. A formal proof of this result is presented in the next section.
- As previously noted, the stability check for this structure is very simple. However, as noted in Reference [25], the mapping between the coefficients of the direct-form realization and the coefficients of the orthogonal realization is not unique, then the problem of manifolds needs to be considered, as for example with a suitable initialization of the updating algorithm.

4. AN EFFICIENT PARTIAL GRADIENT ALGORITHM

In this section is presented a simplified *partial gradient* algorithm that makes explicit use of the properties related to the IIR filter structures studied in the previous section.

The first result is presented in the following theorem that is applicable to the second-order sections orthogonal realization. An extension of this result for realizations with first-order sections is straightforward.

Theorem. The third term of the following inner product is equal to zero, i.e.

$$\left\langle \frac{z^{-1}}{D_j(z)} \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ (H(z)-\hat{H}(z)) \right\rangle = 0 \tag{25}$$

for $j = 1, \ldots, M/2$.

The proof is simple if we use the fact that a rational transfer function in $\mathscr{H}_2$ [20] can be decomposed as $V(z)g(z)$, where $g(z) \in \mathscr{H}_2$ does not have any common zeros with $V(z)$. The properties of this decomposition are called *Beurling–Lax* Theorem [8]. Since the error $H(z) - \hat{H}(z)$ is also a rational function, we can write

$$\left\langle \frac{1}{D_j(z)} \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ (H(z)-\hat{H}(z)) \right\rangle$$
$$= \left\langle \frac{1}{D_j(z)} \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ V(z)g(z) \right\rangle$$

where $V(z) = V_{M/2}(z) = \prod_{k=1}^{M/2} \bar{D}_k(z)/D_k(z)$. Then

$$\left\langle \frac{1}{D_j(z)} \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ V(z)g(z) \right\rangle$$
$$= \left\langle \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ \frac{1}{D_j(z^{-1})} V_{M/2}(z)g(z) \right\rangle$$
$$= \left\langle \sum_{k=j}^{M/2-j} (v_{2k-1}F_{2k-1}^{M/2-j}(z) + v_{2k}F_{2k}^{M/2-j}(z)),\ V_{M/2-1}(z)g'(z) \right\rangle = 0$$

where $g'(z) = (1/D_j(z))g(z) \in \mathscr{H}_2$. The last equation vanishes because the second equality is a linear combination of the basis functions $F_k^{M/2-j}$, with $j > 1$ that is orthogonal to $V_{M/2-1}(z)$. □

The theorem can be applied to eliminate the last term of Equations (23) and (24) related to the stationary point of the updating algorithm with respect to the coefficients β_{1_k} and β_{2_k}.

This result motivates us to present a new algorithm, related to the orthonormal realization of Equation (4). The *simplified partial gradient orthogonal algorithm* can be written as

$$\theta(n+1) = \theta(n) - \mu\nabla'(n) \tag{26}$$

where

$$\nabla'(n) = -2e(n)\left[\frac{\partial y(n)}{\partial v_0}\cdots\frac{\partial y(n)}{\partial v_M}\,\frac{\partial y(n)}{\partial \beta_{1_1}}\,\frac{\partial y(n)}{\partial \beta_{2_1}}\cdots\frac{\partial y(n)}{\partial \beta_{1_{M/2}}}\,\frac{\partial y(n)}{\partial \beta_{2_{M/2}}}\right]^{\mathrm{T}} \tag{27}$$

and

$$\frac{\partial y(n)}{\partial v_k} = \frac{\partial y_k(n)}{\partial v_k} = F_k(n)u(n) \tag{28}$$

$$\frac{\partial y(n)}{\partial \beta_{1_j}} = (v_{2j-1}G^1_{2j-1}(z)u(n) + v_{2j}G^1_{2j}(z)u(n)) - \frac{z^{-1}}{D_j(z)}\sum_{k=j}^{M/2}(v_{2k-1}F_{2k-1}(z)u(n) + v_{2k}F_{2k}(z)u(n)) \tag{29}$$

$$\frac{\partial y(n)}{\partial \beta_{2_j}} = (v_{2j-1}G^2_{2j-1}(z)u(n) + v_{2j}G^2_{2j}(z)u(n)) - \frac{z^{-2}}{D_j(z)}\sum_{k=j}^{M/2}(v_{2k-1}F_{2k-1}(z)u(n) + v_{2k}F_{2k}(z)u(n)) \tag{30}$$

The result is an efficient orthonormal algorithm with computational complexity similar to the direct-form (or lattice) realization.

Note that, the performance of the proposed algorithm is expected to be close to the full algorithm of Equation (5) except for the iterations before the stationary behaviour is reached. In these iterations the results obtained through the theorem cannot be applied. However, the following result is available:

Lemma. The stationary points of the partial gradient orthogonal algorithm coincide with the stationary points related to the minimization of $E[e^2(n)]$.

The proof is straightforward because the new algorithm was designed to maintain the stationary points of the full algorithm, that determines the minimization of $E[e^2(n)]$. □

Although the full orthogonal algorithm can be shown to converge globally to a local minima of the MSOE, only local convergence can be demonstrated for this algorithm. Similar results to that obtained in Reference [8] can be applied to analyse the convergence of the proposed algorithm. Specially useful are those related to approximation (interpolation) [20] of the specified impulse response even for reduced-order system identification.

5. DISCUSSION, SIMULATION AND COMPARISONS

From the previous sections some advantages and drawbacks of the new realizations are evident, some of them related to the associated updating algorithm and others related to the basis functions that define the realization.

Perhaps a possible limitation of the proposed orthogonal realization as defined for the first- and second-order sections is the fact that multiple poles cannot be modelled. There are many ways in which this problem could be circumvented. A straightforward one is to contemplate *a priori* for possible multiple pole sections (with the associated increase in complexity and loss of modularity). Another form that shows interesting results is to use the concept related to the *polyphase* realization. In this vein, additional results related to the orthogonal realization, in this case using the Steiglitz–McBride updating algorithm, were presented in References [26,27].

Recently, other approaches that use all-pass filters as the basic module of the adaptive IIR realization were proposed [28]. In spite of the interesting characteristics of this realization, the orthogonal properties of the realization are mainly oriented to reduce the transversal adaptive filter order part of the realization, and not to minimize (locally) the mean-squared output error as is the case in our proposal. How the estimate obtained by the realization of Reference [28] relate to MSOE remains an open issue.

As previously noted, the proposed realization is an extension of many results based on Laguerre and Kautz functions or Generalized basis functions [3,22,24]. Some of these realizations could also be used with fixed poles at arbitrary locations known *a priori* to shorten the order of the remaining adaptive FIR (feedforward) coefficients.

5.1. A low complexity gradient realization

In order to show the potential of a low complexity gradient computation for the general orthogonal realization proposed, we consider the usual *slow convergence* approximation in Equations (4) to study some *heuristic simplifications*. We consider also in this set of comparisons a Gauss–Newton-type algorithm, with the usual additional (matrix-inversion lemma based) update equation related to the estimative of the inverse Hessian matrix [4]. This discussion will be useful to illustrate the expected performance of the theoretical results of the last section and also to compare the performance of different approximations using the orthogonal realization:

1. *The complete gradient realization*: The second-order-based orthogonal adaptive IIR filter realization is illustrated in Figure 1. In this discussion we simplify the numerator coefficients for each section to $N_k(z)=(1+z^{-1})$ and $N_k'(z)=(1-z^{-1})$ (see Equations (15) and (16)). In this form orthogonality is maintained with the advantage of reduced complexity. We call this implementation *complete gradient realization*. Using the instantaneous squared value of the output error as an estimate of $E[e^2(n)]$, is illustrative to consider the computation of

$$\nabla_{\beta_{1i}}(n)=\frac{\partial e^2(n)}{\partial \beta_{1i}}=-2e(n)\sum_{k=1}^{M/2} v_k(n)\frac{\partial y_k(n)}{\partial \beta_{1i}}, \quad i=1,2,\ldots,2\lfloor M/2\rfloor$$

$$\nabla_{\beta_{2i}}(n)=\frac{\partial e^2(n)}{\partial \beta_{2i}}=-2e(n)\sum_{k=1}^{M/2} v_k(n)\frac{\partial y_k(n)}{\partial \beta_{2i}}, \quad i=1,2,\ldots,2\lfloor M/2\rfloor$$

It is useful to introduce the signals $g_k(n)$ (see Figure 1) that are related by $g_k(n)=F_k(z)u(n)$, where for example

$$\frac{\partial y_k(n)}{\partial \beta_{1i}}=g_k^{\beta_{1i}}(n)+g_k^{\beta_{1i}}(n-1) \quad \text{for } k \text{ odd}$$

$$\frac{\partial y_k(n)}{\partial \beta_{1i}}=g_k^{\beta_{1i}}(n)-g_k^{\beta_{1i}}(n-1) \quad \text{for } k \text{ even}$$

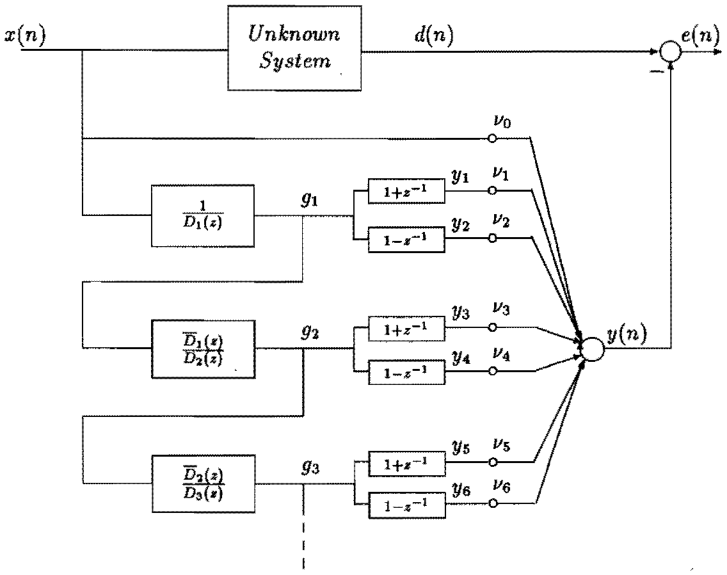

Figure 1. Second-order-based orthogonal adaptive IIR filter realization (simplified numerators).

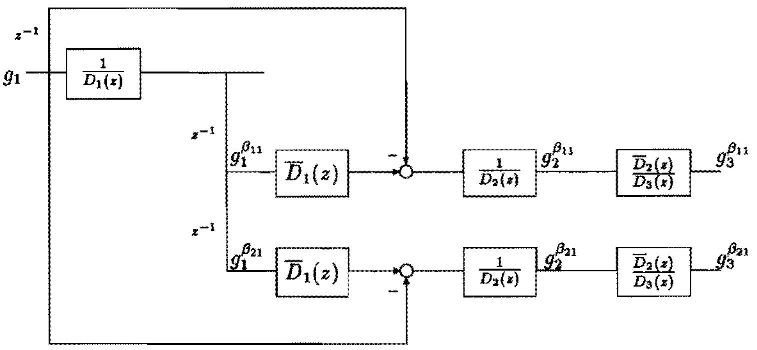

Figure 2. Complete gradient computation: $g_i^{\beta_{11}}$ and $g_i^{\beta_{21}}$ for a sixth-order example.

where $g_k^{\beta_{1i}}(n) = \partial g_k(n)/\partial \beta_{1i}(n)$. Differentiation of the signal $g_1(n)$ (the most complex conputationally) w.r.t. the coefficients β_{11} and β_{21} can be obtained in a sixth-order filter as illustrated in Figure 2.

A first simplification to the gradient computation can be achieved by using the slow convergence approximation, where the coefficients of the adaptive filter are considered approximately the same in neighboring iterations. Using the same sixth-order filter example, the simplified gradient computation w.r.t. the coefficients β_{11} and β_{21} is illustrated in Figure 3.

2. *The partial gradient orthogonal realization*: More interesting, following the results of Section 4 the lower branch in Figure 3 can be disregarded (for this and for each $g_i^{\beta_{1j}}$ and $g_i^{\beta_{2j}}$) to obtain the *partial gradient orthogonal realization* of Equations (28)–(30).

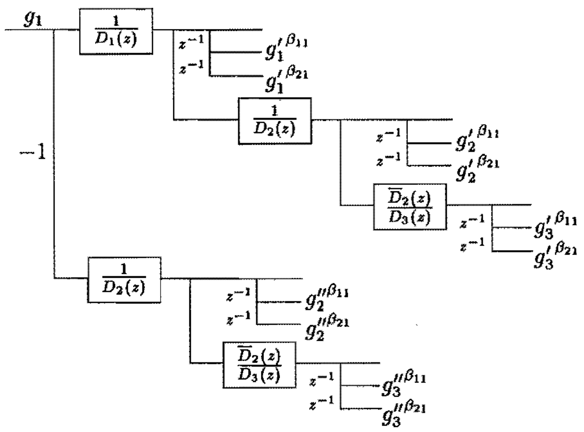

Figure 3. Gradient computation: $g_i^{\beta_{11}} = g_i'^{\beta_{11}} + g_i''^{\beta_{11}}$ and $g_i^{\beta_{21}} = g_i'^{\beta_{11}} + g_i''^{\beta_{11}}$ for a sixth-order example.

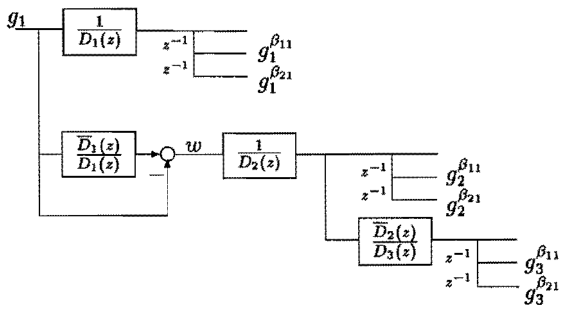

Figure 4. Redraw of Figure 2 using slow convergence approximation: $g_i^{\beta_{11}}$ and $g_i^{\beta_{21}}$ for a sixth-order example.

3. *The reduced complexity orthogonal realization*: Another heuristic simplification is possible for the cases where the poles to be identified are, as usually, close to the unit circle. Consider the gradient computation of the coefficients β_{11} and β_{21} using the complete orthogonal realization of Figure 2 with the slow convergence approximation, as depicted in Figure 4. The signal w in Figure 4, is defined by $w = (\bar{D}_1(z)/D_1(z))g_1 - g_1 = ((\beta_{21} - 1)(z^2 - 1)/D_1(z))g_1$. Then, if the value of the coefficient β_{21} is close to one, as is the case in some usual IIR filtering applications, we can consider the cancellation of this branch. In this case, we need only one all-pole block per section in order to compute the gradient. Then previous equations are reduced to the following:

$$g_k^{\beta_{1i}}(n) = \left\{ \frac{z^{-1}}{D_k(z)} \right\} g_k(n)$$

$$g_k^{\beta_{2i}}(n) = \begin{cases} \left\{ \dfrac{z^{-2}}{D_k(z)} \right\} g_k(n) \\ 0 \quad \text{otherwise} \end{cases}$$

The remaining coefficient derivatives are special cases of the equations above. We call the resulting realization *reduced complexity orthogonal realization*. Obviously the expected performance in terms of convergence speed is worst than in the previous cases. On the other hand, the computational complexity in this case is lower than for the complete and partial orthogonal realizations, since we need only two filters per section in order to obtain the respective gradient components.

4. *The block diagonal simplification*: Owing to the approximate orthogonalization of the internal signals, for a Gauss–Newton-type algorithm, we can use with the reduced complexity orthogonal realization an additional approach that uses a *block diagonal simplification* to the estimate of the inverse of the Hessian matrix using the proposed gradient computation, without a noticeable loss in convergence speed.

To illustrate the behaviour of the complete orthogonal realization, the partial gradient, the reduced complexity orthogonal realizations and the use of a block diagonal simplification for the estimative of the inverse of the Hessian matrix, we present a set of computer simulations in system identification. For completeness, we also include in the comparisons the standard parallel realization. Different initial values for the coefficients of each section are used in order to reduce the manifold effects. We consider a zero mean, Gaussian measurement noise $v(n)$ of -40 dB, and a zero mean white noise input $x(n)$ with unit variance.

For the first example, the plant to be identified has the following transfer function:

$$H(z)=\frac{z(z-0.9)(z^2+0.81)}{(z^2-1.13z+0.64)(z^2+0.9z+0.81)}$$

with zeros at $z_0=0$, $z_1=0.9$, $z_{2,3}=\pm j0.9$, and poles at $p_{0,1}=0.8\angle\pm 45°$ and $p_{2,3}=0.9\angle\pm 120°$. The learning curves for the complete orthogonal realization, the partial gradient orthogonal realization, the reduced complexity realization, the block diagonal simplification and the standard parallel realization are depicted in Figure 5, where the ensemble average of 20 simulations are shown. As can be noted the convergence speed for all realizations are comparable, despite of the different computational complexity.

In a second example, we consider the identification of the following transfer function:

$$H(z)=\frac{z^2-1.75z+0.765}{z^6-1.31z^5+2.0879z^4-1.6169z^3+1.662z^2-0.784z+0.5740} \tag{31}$$

with zeros at $z_0=0.9$, $z_1=0.85$ and poles at $p_{0,1}=0.8718\angle\pm 109.78°$, $p_{2,3}=0.95\angle\pm 43.23°$ and $p_{4,5}=0.91\angle\pm 73.74°$. The learning curves for the realizations of the previous example are depicted in Figure 6, where also the averages of 20 simulations are shown. Even for this higher-order case, the convergence speed of the proposed partial gradient orthogonal realization are comparable to the complete realization, and faster than the standard parallel realization.

5.2. The partial gradient orthogonal realization

In this section we compare the proposed partial gradient orthogonal realization using second-order sections, in its simplified version of Section 4 with the *simplified gradient lattice* (SPGL) realization proposed and extensively studied in Reference [8]. As a result of the similar approach followed for the proposed realization and the SPGL realization, the performance in terms of convergence speed is similar in both cases.

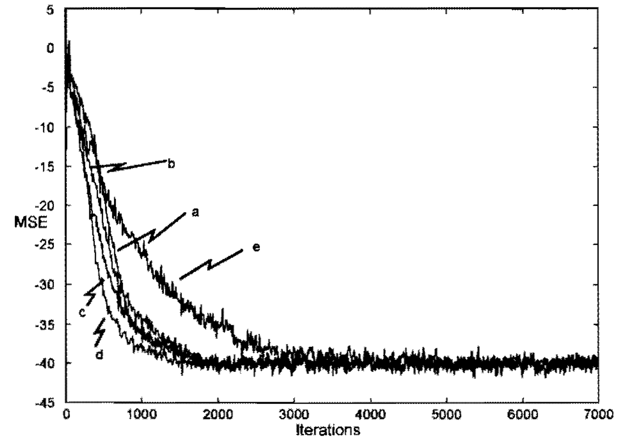

Figure 5. Gauss–Newton-type algoritm comparison (fourth-order example): (a) parallel realization; (b) complete orthogonal realization; (c) partial gradient realization; (d) reduced complexity (e) block diagonal inverse Hessian estimative matrix.

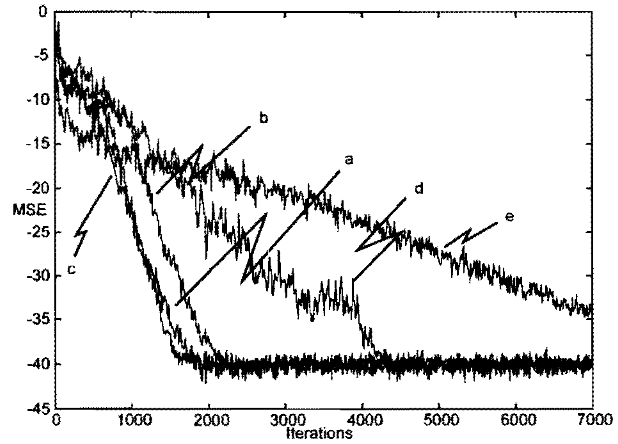

Figure 6. Gauss–Newton-type algoritm comparison (sixth-order example): (a) partial gradient orthogonal realization, (b) reduced complexity, (c) complete orthogonal realization, (d) parallel realization (e) block diagonal inverse Hessian estimative matrix.

Unlike from the previous section, we compare both realizations with an stochastic gradient-based update equation as in Equation (26). In this case in order to illustrate the learning behaviour 20 independent computer run were averaged and measurement noise was not included. In a fourth-order system identification example, the following transfer function was used:

$$H(z)=\frac{z^4-1.39z^3+1.92z^2-1.71z+0.98}{z^4-0.22z^3+0.77z^2-0.21z+0.68}$$

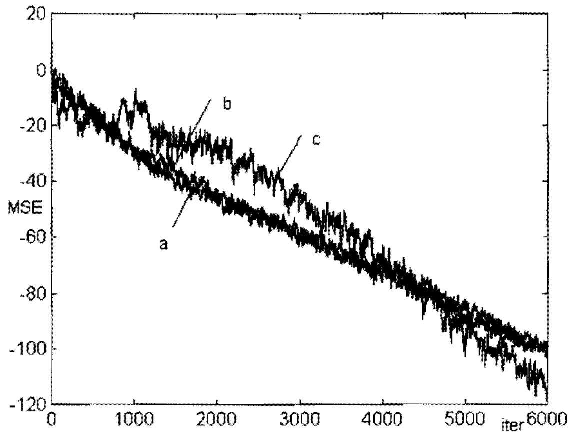

Figure 7. Stochastic gradient-type algorithm comparison (fourth-order example): (a) complete orthogonal realization; (b) partial gradient orthogonal realization; (c) simplified partial gradient lattice realization.

with zeros at $z_{0,1} = 1.1\angle \pm 91.8°$, $z_{2,3} = 0.9\angle \pm 36°$, and poles at $p_{0,1} = 0.92\angle \pm 117°$ and $p_{2,3} = 0.9\angle \pm 54°$. The convergence factor μ was optimized to obtain the fastest convergence speed for each algorithm. In Figure 7 are illustrated the learning curves for the complete orthogonal realization, the partial gradient orthogonal realization and the SPGL realization. As can be noted, the performance in terms of convergence speed is very similar for both the complete orthogonal algorithm and the simplified version of Section 4, as expected. Indeed, the performance of the proposed simplified realization is also similar to the SPGL realization.

In a second example, the transfer function to be identified is the same as that for the previous sixth-order example of Equation (31). Figure 8 illustrates the performance of the simplified orthogonal realization and the SPGL realization for the sixth-order example.

The simulations include the complete orthogonal realization, the simplified orthogonal realization and the SPGL realization. As can be noted, the learning curve for the complete orthogonal realization indicates a noticeably slower convergence speed than for the other cases. More interesting, the simplified orthogonal realization shows the fastest initial convergence speed for this example, although the SPGL realization reaches the noise floor in first place.

6. CONCLUSIONS

A new efficient realization suitable for adaptive IIR filters is presented. The efficiency of the new realization is based on the orthonormal basis functions representation adopted, particularly real orthonormal first- or second-order sections. The proposed realization for the adaptive IIR filter uses a generalization of some orthogonal realizations based on Laguerre or Kautz basis functions. Two main results were presented, both related to stationary point characterization and expected convergence behaviour. As a result an efficient low complexity associated gradient computation is proposed, that we defined the *partial gradient orthogonal realization.*

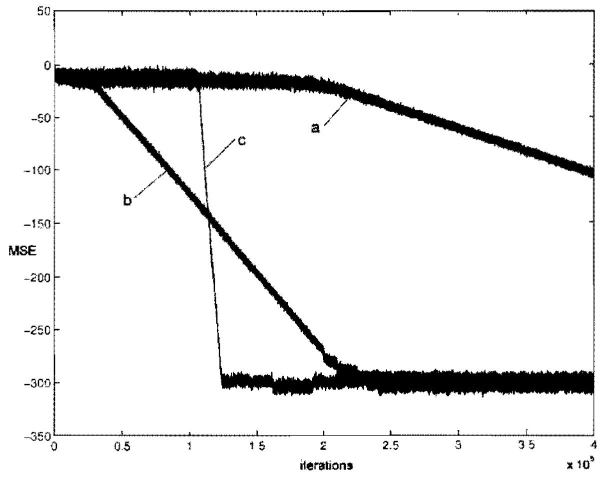

Figure 8. Stochastic gradient-type algorithm comparison (sixth-order example): (a) complete orthogonal realization; (b) partial gradient orthogonal realization; (c) simplified partial gradient lattice realization.

In order to confirm the expected performance of the proposed realization, in terms of convergence speed, a set of computer simulations are included. Indeed, some heuristic simplifications in order to make use of the orthogonal properties of the proposed realization were discussed. The comparison results of comparison with other practical adaptive IIR realizations demonstrate that the new structures are useful for practical applications.

REFERENCES

1. Wahlberg B. System identification using Kautz models. *IEEE Transactions on Automatic Control* 1994; **39**:1276–1282.
2. Davidson GW, Falconer DD. Reduced complexity echo cancellation using orthogonal functions. *IEEE Transactions on Circuits and Systems* 1991; **38**:20–28.
3. Ninness B, Gustafsson F. A unifying construction of orthonormal bases for system identification. *IEEE Transactions on Automatic Control* 1997; **42**:515–521.
4. Ljung L, Söderström T. *Theory and Practice of Recursive Identification*. The MIT Press: Cambridge, 1983.
5. Mendel JM. *Discrete Techniques of Parameter Estimation: The Equation Error Formulation*. Marcel Dekker: New York, 1973.
6. Steiglitz K, McBride LF. A technique for the identification of linear systems. *IEEE Transactions on Automatic Control* 1965; **AC-10**:461–464.
7. Fan H, Jenkins WK. A new adaptive IIR filter. *IEEE Transactions on Circuits and Systems* 1986; **CAS-33**:939–947.
8. Regalia PA. *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker: New York, 1995.
9. López-Valcarce R, Pérez-González F. New schemes and theoretical analysis of the master-slave family of recursive identification algorithms. *Signal Processing* 1997; **58**:79–94.
10. Shynk JJ, Gooch RP. Frequency-domain adaptive pole-zero filtering. *Proceedings of IEEE* 1985; **73**(10):1526–1528.
11. Shynk JJ. Adaptive IIR filtering using parallel form realizations. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 1989; **37**:519–533.
12. Nayeri M, Jenkins K. Alternate realizations to adaptive IIR filters and properties of their performance surfaces. *IEEE Transactions on Circuits and Systems* 1989; **36**:485–496.
13. Diniz PSR, Cousseau JE, Antoniou A. Parallel realization for fast adaptive IIR filters. *IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing* 1994; **41**:561–567.
14. Diniz PSR, Cousseau JE, Antoniou A. Improved parallel realization for adaptive IIR filters. *Proceedings of IEE Part G: Circuits, Devices and Systems* 1993; **140**:322–328.

15. Parikh D, Ahmed N, Stearns SD. An adaptive lattice algorithm for recursive filters. *IEEE Transactions on Acoustics, Speech and Signal Processing* 1980; **ASSP-28**:110–112.
16. Regalia P, Mitra SK, Vaidyanathan PP. The digital all-pass filter: a versatile signal processing building block. *Proceedings of IEEE* 1988; **76**.
17. Rodriguez-Fonollosa JA, Masgrau E. Simplified gradient calculation in adaptive IIR lattice filters. *IEEE Transactions on Signal Processing* 1991; **39**:1702–1705.
18. Miao KX, Fan H, Doroslovački M. Cascade lattice adaptive IIR filters. *IEEE Transactions on Signal Processing* 1994; **42**:721–742.
19. Roberts RA, Mullis CT. *Digital Signal Processing*. Addison-Wesley: Reading, MA, 1987.
20. Chui CK, Chen G. *Discrete* H^{∞} *Optimization*. Springer: Houston, 1997.
21. Fahmy MA *et al.* System identification using discrete orthogonal functions. *International Journal of Circuit Theory and Applications* 1993; **21**(4):317–329.
22. Bodin P, Oliveira TAM, Wahlberg B. On the construction of orthonormal basis functions for system identification. *Technical Report IR-S3-REG*-9605, S3-Automatic Control, The Royal Institute of Technology, 1994.
23. den Brinker AC, Benders FPA, Oliveira TAM. Optimality conditions for truncated Kautz series. *IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing* 1996; **43**:117–122.
24. Heuberger PSC, Van den Hof PMJ, Bosgra O. A generalized orthonormal basis for linear dynamical systems. *Proceedings of the 32nd Conference on Decision and Control*, San Antonio, Texas, 1993; 2850–2855.
25. Cousseau J, Sentoni G, Diniz PSR, Agamennoni O. On orthogonal parallel realization for adaptive IIR filters. *Proceedings of IEEE International Conference on Electronics, Circuits Systems*, Rodos, Greece, 1996; 856–859.
26. Salama L, Cousseau J. Efficient echo cancellation based on an orthogonal adaptive IIR realization. presented at *IEEE/SBT International Telecommunication Symposium*, Rio de Janeiro, August, 1998.
27. Salama L, Cousseau J. Comparison of orthonormal adaptive FIR and IIR filter realizations. presented at *IEEE Symposium on Advances in Digital Filtering and Signal Processing*, Victoria, B.C., Canada, June 1998; 77–80.
28. Kobayashi M, Takagi Y, Okello J, Itoh Y. An IIR adaptive filter based on estimation of all-pass system. *IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing* **45**:676–684.
29. Cousseau J, Diniz PSR, Sentoni G, Agamennoni O. An orthogonal adaptive IIR realization for efficient MSOE optimization. *Proceedings of IEEE International Conference on Electronics, Circuits Systems*, Lisbon, Portugal, 1998; 221–224.
30. Nayeri M. Improvements in adaptive IIR filtering: theory and application. *Ph.D. Thesis*, University of Illinois at Urbana, Champaign, 1988.
31. Partington JR. *An Introduction to Hankel Operators*. Cambridge University Press: London, 1988.

