Online optimization of the time scale in adaptive Laguerre-based filters

Noël Tanguy, Riwal Morvan, Pierre Vilbé, Léon-Claude Calvez

Correspondence

Online Optimization of the Time Scale in Adaptive Laguerre-Based Filters

N. Tanguy, R. Morvan, P. Vilbé, and L. C. Calvez

***Abstract*—A new online method to optimize the free parameter in adaptive Laguerre-based filters is presented. It is based on the minimization of a criterion that is equivalent to an upper bound for the quadratic approximation error. The proposed technique presents a fast convergence and a good robustness.**

***Index Terms*—Adaptive filtering, Laguerre filters, pole optimization.**

I. Introduction

In recent years, there has been large interest in the use of discrete Laguerre functions in approximation, modeling, filtering, identification, and control. These functions, which involve a free parameter closely related to the time scale, form a complete orthogonal set in $\ell^2[0, \infty)$. In adaptive filtering or in modeling, when a N-terms Laguerre model is used, a design problem is to optimally select the free parameter to minimize the modeling error. In [1]–[4], offline methods of parameter optimization based on the minimization of the error energy were proposed. These methods, which supply the optimal value of Laguerre pole position, require a relatively high computational cost. In another way, Fu and Dumont [5] and Tanguy [6] have proposed a suboptimal offline method for choosing the free parameter. The method is based on the minimization of an upper bound for the modeling error. It has received considerable interest owing to its simplicity, its low computational cost, and its relatively good efficiency. From this method, we derive, in this correspondence, a technique for suboptimal online optimization of the Laguerre filters parameter. We show that the proposed algorithm is efficient and presents good stability and convergence and numerical robustness.

This correspondence is organized as follows. In Section II, we recall the principles of the suboptimal method for adjusting the pole location of the Laguerre model. In Section III, we extend the method to model reduction in case where the full model is given by its Laguerre spectrum. In Section IV, we present a new method for online optimization of the Laguerre free parameter, and we give some illustrative examples in Section V before concluding in Section VI.

II. Background

We implicitly assume that all functions and sequences vanish for negative time or spectral index ($k < 0$ or $n < 0$). Suppose that $f(k)$ is a well-behaved real-valued discrete time signal that can be represented by the infinite expansion

$$f(k) = \sum_{n=0}^{\infty} c_n(a)\varphi_n(k, a) \tag{1}$$

where $\varphi_n(k, a)$ are the discrete Laguerre functions defined by the z transform

$$\sum_{k=0}^{\infty} \varphi_n(k, a)z^{-k} = \sqrt{1-a^2}\frac{z}{z-a}\left(\frac{1-az}{z-a}\right)^n \tag{2}$$

and depending on the free parameter $a(|a| < 1)$. As the discrete Laguerre functions are orthonormal $\sum_{k=0}^{\infty} \varphi_n(k, a)\varphi_m(k, a) = \delta_{n,m}$, the coefficients $c_n(a)$ in (1) are given by

$$c_n(a) = \sum_{k=0}^{\infty} f(k)\varphi_n(k, a). \tag{3}$$

Now, consider the truncated series $\sum_{n=0}^{N-1} c_n(a)\varphi_n(k, a)$, which for a given a is the best N-terms approximation of $f(k)$ in the sense of minimizing the normalized quadratic error defined by

$$q_N(a) \triangleq \frac{\|f-\tilde{f}\|^2}{\|f\|^2} = \frac{\sum_{n=N}^{\infty} c_n^2(a)}{\|f\|^2} \tag{4}$$

where $\|f\|^2 \triangleq \sum_{k=0}^{\infty} f^2(k)$. A key point in [5] and [6] is to use the inequality $\sum_{n=0}^{\infty} nc_n^2(a) \geq N\sum_{n=N}^{\infty} c_n^2(a)$ to derive an upper bound for $q_N(a)$ as

$$q_N(a) \leq J_N(a) \triangleq \frac{\sum_{n=0}^{\infty} nc_n^2(a)}{N\|f\|^2}. \tag{5}$$

Using the nice relationship for the discrete Laguerre functions (see, for example, [7] and [8])

$$\begin{aligned} n\varphi_n(k, a) = &-\frac{(k+1)a}{1-a^2}\varphi_n(k+1, a) \\ &+\frac{k\left(1+a^2\right)+a^2}{1-a^2}\varphi_n(k, a) \\ &-\frac{ka}{1-a^2}\varphi_n(k-1, a) \end{aligned} \tag{6}$$

it is a simple matter to show that

$$\sum_{n=0}^{\infty} nc_n^2(a) = \frac{(1+m_1)a^2 - 2m_2a + m_1}{1-a^2}\|f\|^2 \tag{7}$$

holds, where the "moments" m_1 and m_2 are defined as

$$\begin{aligned} m_1 &\triangleq \sum_{k=1}^{\infty} kf^2(k) \Big/ \|f\|^2 \\ m_2 &\triangleq \sum_{k=1}^{\infty} kf(k)f(k-1) \Big/ \|f\|^2. \end{aligned} \tag{8}$$

Therefore, the upper bound for $q_N(a)$ can be written

$$J_N(a) = \frac{(1+m_1)a^2 - 2m_2a + m_1}{(1-a^2)N}.$$

Let

$$\rho_0 \triangleq \frac{2m_1+1}{2m_2} \tag{9}$$

Manuscript received August 5, 1998; revised September 24, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Mahmood R. Azimi-Sadjadi.

The authors are with Laboratoire d'Electronique et Systèmes de Télécommunications, LEST UMR CNRS 6616, Faculté des Sciences et Techniques, Université de Bretagne Occidentale (UBO), Brest, France (e-mail: Noel.Tanguy@univ-brest.fr).

Publisher Item Identifier S 1053-587X(00)02369-2.

for which it has been shown [6], [8] that $|\rho_0| \geq 1$ is always true. The suboptimal value of the a parameter minimizing $J_N(a)$ is a root of equation $a^2 - 2\rho_0 a + 1 = 0$ and is given by

$$\begin{cases} a_0 = \rho_0 - \sqrt{\rho_0^2 - 1}, & \text{if } \rho_0 \geq 1 \\ a_0 = \rho_0 + \sqrt{\rho_0^2 - 1}, & \text{if } \rho_0 \leq -1. \end{cases} \tag{10}$$

Remarks: ρ_0 is a characteristic of the function $f(k)$ and is thus independent of the a parameter. The optimal value a_0 in (10) is independent of N and is guaranteed to lie in $[-1, 1]$. Furthermore, the bound $J_N(a_0)$ is the best that can be achieved when the specification of $f(k)$ is limited to the knowledge of m_1 and m_2 [6].

III. Pole Optimization Based on the Laguerre Spectrum

Before dealing with the problem of the online optimization of the a parameter in adaptive Laguerre-based filtering, we are interested in the expression of the moments m_1, m_2, and ρ_0. Indeed, a key point of the proposed technique is to write the characteristic constant ρ_0 in (9) as a function of the c_n coefficients of the Laguerre spectrum of $f(k)$, e.g., the impulse response of a filter. In this way, we will determine the optimal value of the parameter directly from the Laguerre spectrum $c_n(a)$ of $f(k)$. This important problem happens in most practical cases: when the spectral model is obtained via adaptive filtering or from its Laplace or z transform. Thus, an a-parameter optimization for model order reduction can be done directly using the coefficients of the Laguerre spectrum without calculating the samples of the impulse response $f(k)$. Owing to the properties of the Laguerre transform, it will be shown that the moments m_1, m_2, and ρ_0 can be written in terms of the coefficients of the Laguerre-spectrum assumed to be known for a given value of the a- parameter

$$\rho_0 = \frac{S_2(a) + \rho S_1(a)}{\rho S_2(a) + S_1(a)} \tag{11}$$

with $\rho \hat{=} (1 + a^2)/2a$ and

$$S_1(a) \hat{=} \sum_{n=0}^{\infty} (2n+1) c_n^2(a)$$
$$S_2(a) \hat{=} 2 \sum_{n=1}^{\infty} n c_n(a) c_{n-1}(a). \tag{12}$$

Proof: Using the well-known property for discrete Laguerre functions (see [8], for example)

$$\frac{\partial}{\partial a}\varphi_n(a) = \frac{1}{1-a^2}\left[(n+1)\varphi_{n+1}(a) - n\varphi_{n-1}(a)\right] \tag{13}$$

and (3), we deduce

$$\frac{\partial}{\partial a} S_1(a) = \frac{-2}{1-a^2} S_2(a)$$
$$\frac{\partial}{\partial a} S_2(a) = \frac{-2}{1-a^2} S_1(a). \tag{14}$$

Now, using (7), Parseval theorem, and (14), $S_1(a)$ and $S_2(a)$ in (12) are expressed as

$$S_1(a) = \frac{\|f\|^2}{1-a^2}\left[2\left(1+a^2\right)m_1 - 4a m_2 + 1 + a^2\right]$$
$$S_2(a) = 2\frac{\|f\|^2}{1-a^2}\left[-2a m_1 + \left(1+a^2\right)m_2 - a\right]. \tag{15}$$

Therefore, solving (15) for m_1 and m_2, we obtain

$$m_1 = \frac{\left(1+a^2\right)S_1 + 2aS_2 - (1-a^2)\|f\|^2}{2(1-a^2)\|f\|^2}$$
$$m_2 = \frac{2aS_1 + \left(1+a^2\right)S_2}{2(1-a^2)\|f\|^2}$$

and then, in view of definition (9), (11) follows. □

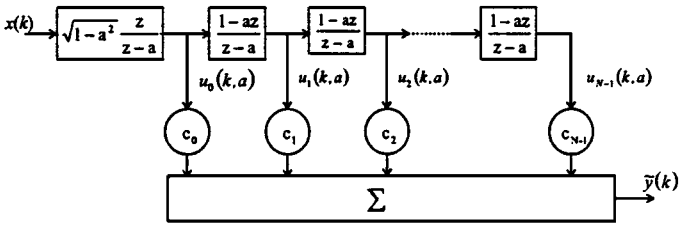

Fig. 1. Laguerre filter of order N.

Remark: Using (12) and (14), the derivative of $J_N(a)$ (5) with respect to a is written as

$$\frac{\partial}{\partial a} J_N(a) = \frac{-S_2(a)}{(1-a^2)N\|f\|^2}.$$

This implies that $S_2(a)$ is zero when the upper bound for the quadratic error is minimal. Because the minimum of $J_N(a)$ is unique for $|a| < 1$, the reciprocal is true; when $S_2(a)$ or $S_2(a)/S_1(a)$ are nearly zero, this means that a is near its optimal value. When the optimal value a_0 is obtained, the new coefficients can then be determined by the exchange relation [8]

$$c_n(a_0) = \sum_{j=0}^{\infty} c_j(a)\varphi_n\left(j, \frac{a_0 - a}{1 - a a_0}\right) \tag{16}$$

which holds for all a and a_0 in $]-1, 1[$. If the series is, in practice, truncated, the approximation would be relatively good if the number N of terms is great enough.

IV. Online Optimization of the a Parameter

We consider the filter of an adaptive Laguerre-based scheme (see Fig. 1). At the output, we have

$$\tilde{y}(k) = \sum_{n=0}^{N-1} c_n u_n(k, a). \tag{17}$$

The instantaneous performance of the system is measured by the error E, which is defined as

$$E \hat{=} \tfrac{1}{2} e(k)^2 = \tfrac{1}{2}\left(y(k) - \tilde{y}(k)\right)^2 \tag{18}$$

where $y(k)$ is the output of the unknown system. In the examples of Section V, the normalized least mean squares (NLMS) algorithm will be chosen. The update rule for the c_n coefficients is then given by $c_n(k+1) = c_n(k) - \mu e(k) u_n(k)$. The convergence is ensured if [9], [10]

$$0 < \mu(k) \leq \frac{1}{\sum_{n=0}^{N-1} u_n^2(k).}$$

A. Gradient Method to Optimize the a-Parameter

It is possible to determine the a-parameter variation gradient in a way that is similar to the gradient method that is used for the optimization of the c_n coefficients. In [11] a similar gradient descent algorithm to optimize the free parameter of gamma filters had been proposed. The least mean square (LMS) algorithm corrects the a parameter proportionally to the negative of the local gradient, i.e., the a-parameter update equation is in the direction of the negative gradient $\triangle a = -\eta(\partial E/\partial a)$, where η is a positive step size parameter. From (13), (17), and (18), it follows that

$$\triangle a = \frac{\eta e(k)}{1-a^2}\sum_{n=0}^{N-1} c_n(k)\left[(n+1)u_{n+1}(k, a) - n u_{n-1}(k, a)\right]. \tag{19}$$

This method has the following drawbacks.

- The convergence is very slow and is only ensured for values of η that are very small.
- The a-parameter convergence to its optimal value is not guaranteed, especially when η is small. The convergence can lead to a local minimum.
- The necessity of one supplementary filter because the computation of $(\partial/\partial a)u_{N-1}(k,\ a)$ requires the knowledge of $u_N(k,\ a)$ in (19).

As observed in numerous simulations, this method is very sensitive to the choice of η. Although good results and especially optimal value can be obtained, it remains, nevertheless, very difficult to use.

B. Suboptimal Method

The original idea for the optimization of the a parameter is to subject it to a small variation in the direction of its optimal value computed from the c_n-coefficient spectrum. In fact, to reduce somewhat the calculus and as $\rho = (1 + a^2)/2a$ is a monotonic function of the a parameter, we would rather work with ρ, whose optimal value is given by (11). The proposed algorithm initialized with $a_1 = 0.5$, for example, and $\rho_1 = (1 + a_1^2)/2a_1$ is as follows.

1) Compute the estimated "moments" $\tilde{S}_1(a_k)$ and $\tilde{S}_2(a_k)$ using

$$\tilde{S}_1(a_k) \hat{=} \sum_{n=0}^{N-1} (2n+1)c_n^2(k)$$

$$\tilde{S}_2(a_k) \hat{=} 2\sum_{n=1}^{N-1} nc_n(k)c_{n-1}(k).$$

2) Evaluate

$$\rho_{0,\,k} = \frac{\tilde{S}_2(a_k) + \rho_k \tilde{S}_1(a_k)}{\rho_k \tilde{S}_2(a_k) + \tilde{S}_1(a_k)}.$$

3) Calculate the new value $\rho_{k+1} = \rho_k + \eta(\rho_{0,\,k} - \rho_k)$.
4) Determine the corresponding value of the a parameter

$$a_{k+1} = \rho_{k+1} \pm \sqrt{\rho_{k+1}^2 - 1}.$$

The adaptation algorithm first adjusts the weights in order to match the desired output data of the unknown system in a least square sense and then corrects the a parameter in order to minimize the upper bound $J_N(a)$. The error surface relative to the weights adjust is quadratic for any given a-parameter in $]-1,\ 1[$. On the other hand, the upper bound $J_N(a)$ is an unimodal function of the a parameter in $]-1,\ 1[$. To ensure the convergence, η must be chosen relatively small. It must be all the more smaller than N is great and than the signal-to noise-ratio is high (a good value experimentally deduced is η ranging from 0.001–0.05). The computation of the new value of a requires $2N + 3$ additions, $4N + 3$ multiplications or divisions, and one square root (for the gradient method, $2N - 2$ additions and $3N + 4$ multiplications are required).

V. Simulation Results

Example 1: The first system we consider has the transfer function

$$F(z) = \frac{z^2 - 1.1z - 0.2}{(z - 0.9)^2}.$$

The input signal $x(k)$ is obtained by passing a Gaussian white noise of unit variance through a first-order lowpass filter $N(z) = (z - 0.6)/(z - 0.7)$. A fourth-order Laguerre model filter is used with an initial a parameter equal to 0.5. The system is simulated using 1200 samples. The algorithm is started after the second sample. The variation of the a parameter during the simulation is presented in Fig. 2. With $\eta = 1/50$, we observe a rapid convergence of the proposed algorithm to a relatively good value of the a parameter. The gradient algorithm presents a troubled variation of the a parameter that has a pernicious influence on the convergence of the a parameter as well as of the c_n coefficients. Note that our method supplies a relatively stable value of the a parameter after only 500 iterations. After 1200 iterations, our method gives $a = 0.9220$ with a normalized quadratic error $q_N = 0.0018$ with $q_N \hat{=} \|f - \tilde{f}\|^2/\|f\|^2$ and where f and $\tilde{f}$ are, respectively, the impulse responses of the unknown system and its Laguerre model. The gradient method with the well-chosen value $\eta = 1.0 * 10^{-5}$ yields $a = 0.8412$ with $q_N = 0.0120$. In Fig. 3, we present the variation of the normalized quadratic error q_N as a function of the a parameter. It shows the quality improvement obtainable by a good choice of the a parameter.

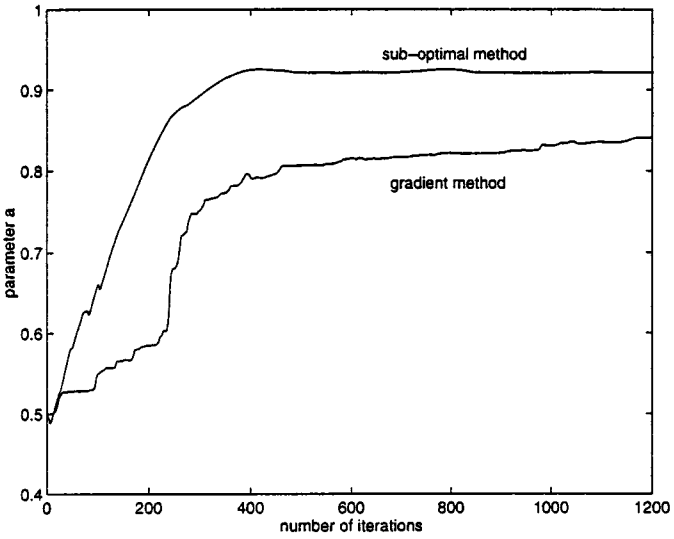

Fig. 2. Comparison of a-parameter adaptive curve.

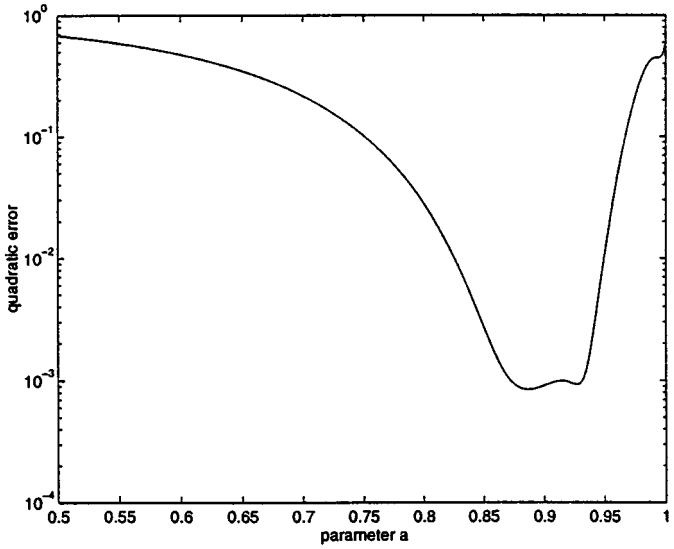

Fig. 3. Variation of the normalized quadratic error as a function of the a-parameter.

Example 2: In the second example, the unknown system is given by

$$F(z) = \frac{z^3 - 0.8z^2 - 1.2z + 0.9}{(z - 0.7)(z - 0.8)(z^2 - 1.8z + 0.9)}.$$

The input signal is a Gaussian white noise of unit variance. A 15th-order Laguerre model filter is used with an initial a parameter equal to 0.5. The system is simulated using 4000 samples. Fig. 4 presents the variation of the a parameter during the simulation. In this example, the gradient algorithm presents a very slow convergence with $\eta = 1.0 * 10^{-6}$, which was well chosen after different tests. After 4000 iterations,

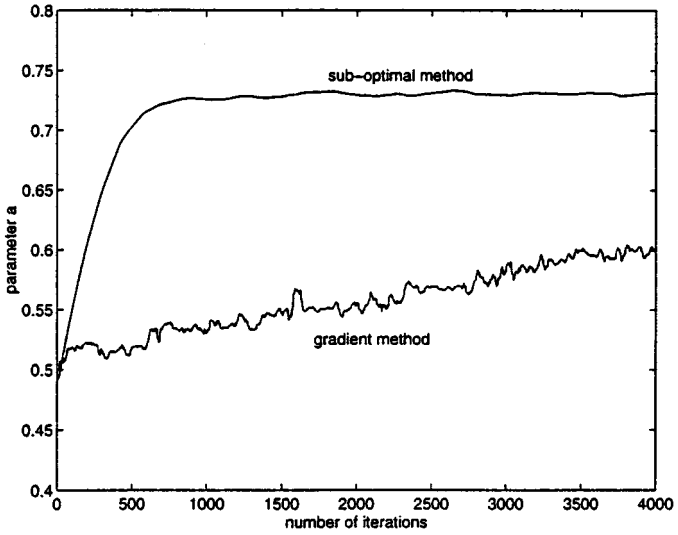

Fig. 4. Comparison of a-parameter adaptive curve.

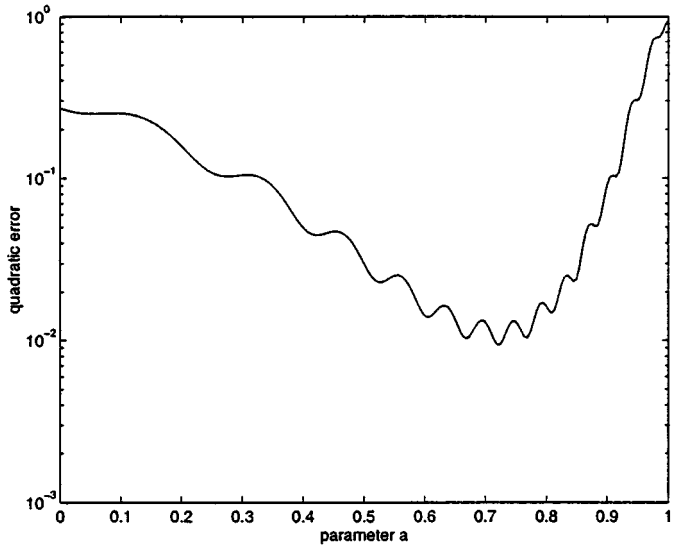

Fig. 5. Variation of the normalized quadratic error as a function of the a-parameter.

our suboptimal algorithm with $\eta = 1/200$ gives $a = 0.7311$ with $q_N = 0.0283$. Note that our technique converges rapidly, and after only 800 samples, a good value for the a parameter is obtained. In Fig. 5, we present the variation of the normalized quadratic error q_N versus the a parameter.

VI. Conclusion

In this correspondence, we have developed a new online parameter optimization algorithm in adaptive Laguerre-based models. The proposed method highly exhibits better performance than that achieved by a gradient algorithm and requires only some supplementary calculus. The optimal value of the parameter is obtained directly from the Laguerre spectrum of the function/system without calculating the samples of the impulse response. The results of our simulation indicate that the method supplies a rapid convergence to a good value of the a parameter. Different simulation conditions have been tested that show the robustness and the numerical stability of the algorithm. Moreover, the proposed algorithm has been efficiently implemented on a digital signal processor and can be efficiently used by itself or as a preliminary approach for an optimal method.

References

[1] M. A. Masnadi-Shirazi and N. Ahmed, "Laguerre approximation of non recursive discrete-time systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, Apr. 3–6, 1990, pp. 1309–1312.

[2] ——, "Optimum Laguerre networks for a class of discrete-time systems," *IEEE Trans. Signal Processing*, vol. 39, pp. 2104–2108, Aug. 1991.

[3] M. A. Masnadi-Shirazi, "Optimum synthesis of linear discrete-time systems using orthogonal Laguerre sequences," Ph.D. dissertation, Univ. New Mexico, Albuquerque, May 1990.

[4] T. Oliveira e Silva, "On the determination of the optimal pole position of Laguerre filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 2079–2087, Sept. 1995.

[5] Y. Fu and G. A. Dumont, "An optimum time scale for discrete Laguerre network," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 934–938, June 1993.

[6] N. Tanguy, P. Vilbé, and L. C. Calvez, "Optimum choice of free parameter in orthonormal approximations," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1811–1813, 1995.

[7] J. Ragot, M. Roesch, and C. Humbert, "Algorithm for identifying the dynamic characteristics of objects by means of orthogonal functions," *J. A*, vol. 18, no. 3, pp. 156–168, 1977.

[8] N. Tanguy, "La transformation de Laguerre discrète," Thèse de doctorat, Univ. Brest, Brest, France, Dec. 1994.

[9] H. Perez and S. Tsujii, "IIR adaptative filtering via discrete Legendre functions," *Electron. Lett.*, vol. 24, no. 8, pp. 450–451, 1988.

[10] ——, "A system identification algorithm using orthogonal functions," *IEEE Trans. Signal Processing*, vol. 39, pp. 752–756, Apr. 1991.

[11] J. C. Principe, B. De Vries, and P.G. de Oliveira, "The gamma filter—A new class of adaptive IIR filters with restricted feedback," *IEEE Trans. Signal Processing*, vol. 41, pp. 649–656, Feb. 1993.

A Two-Stage Algorithm for MIMO Blind Deconvolution of Nonstationary Colored Signals

Chor Tin Ma, Zhi Ding, and Sze Fong Yau

***Abstract*—A new two-stage algorithm is proposed for the deconvolution of multi-input multi-output (MIMO) systems with colored input signals. While many blind deconvolution algorithms in the literature utilize high order statistics of the output signal for white input signals, the additional information contained in colored input signals allows the design of second-order statistical algorithms. In fact, practical signal sources such as speech signals do have distinct, nonstationary, colored power spectral densities. We present a two-stage signal separation approach in which the first step utilizes a matrix pencil between output auto-correlation matrices at different delays, whereas the second stage adopts a subspace method to identify and deconvolve MIMO systems.**

***Index Terms*—Blind deconvolution, multiuser systems, signal separation.**

Manuscript received August 5, 1998; revised October 18, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ali H. Sayed.

C. T. Ma and S. F. Yau are with the Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: ctma@ee.ust.hk; eeyau@ux-mail.ust.hk).

Z. Ding is with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242 USA (e-mail: zding@eng.uiowa.edu).

Publisher Item Identifier S 1053-587X(00)02354-0.