XXXVII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBrT2019, 29/09/2019–02/10/2019, PETRÓPOLIS, RJ

Adaptive linear predictors in cascade for blind deconvolution of non-stationary and non-minimum-phase channels

Renan D. B. Brotto, Kenji Nose-Filho, Romis Attux, João M. T. Romano

***Abstract*— Linear prediction plays a fundamental role in digital signal processing due to its interesting theoretical and practical aspects. An important application is the problem of predictive blind deconvolution. However, it is well known that the classical predictive technique, which assumes the use of the mean squared error (MSE) criterion together with a linear FIR (finite impulse response) fails when the distortion system is non-minimum-phase. In previous works, we have investigated alternative criteria for blind predictive deconvolution, replacing the MSE, which is related to the ℓ_2 norm, by the generalized ℓ_p norm, with $p \neq 2$. The results were effective for some non-minimum-phase systems, but not all of them, which clearly indicated a limitation of the linear FIR structure. In the present paper, we propose to employ a cascade of forward and backward linear predictors. The approach is applied to the blind equalization of communication channels. Due to the characteristics of the transmitted signals, the ℓ_p criterion must be considered with $p \to \infty$. We opt to use $p = 4$, which corresponds to the MFE (Mean Fourth Error) criterion, as a smooth approximation of the ℓ_∞ norm. Also, it allows applying the LMF (Least Mean Fourth) adaptive algorithm, in order to track non-stationary behaviors. Simulation results show that the proposed solution is able to deal effectively with the blind equalization of non-stationary and non-minimum-phase channels.**

***Keywords*— adaptive linear prediction, cascade structure, least mean fourth (LMF) algorithm, non-stationary and non-minimum-phase channels, predictive blind deconvolution.**

I. Introduction

Linear prediction (LP) naturally evokes the works of Kolmogorov [1] and Wiener [2], which established the principles of linear optimal filtering. The classical paper by Makhoul [3], as well as the nice books of M. Bellanger [4] and P. P. Vaidyanathan [5], provides an excellent survey about the theoretical relevance and several applications of LP. One of the first uses of this method was in seismic signal processing, in the predictive deconvolution problem [6]. The predictive approach has also been used for the blind channel equalization problem in communication systems [7]. However, it is well known that the main limitation of predictive deconvolution concerns the use of the linear structure associated with the mean-squared error criterion (MSE), since this setup provides magnitude equalization, but the phase response remains distorted.

Renan D. B. Brotto (e-mail), Romis Attux and João, M. T. Romano are with the School of Electrical and Computer Engineering, University of Campinas (UNICAMP). Kenji Nose-Filho is with the Center of Engineering, Modeling and Applied Social Sciences, Federal University of ABC (UFABC).
The authors would like to thanks CAPES (finance process 001), CNPq (305621/2015-7), FAPESP (2017/13025-2) and FAEPEX for the financial support.

An alternative to perform phase equalization is to adopt a criterion different from the classical MSE, which is related to the ℓ_2 norm. In [8], [9] it has been shown that the ℓ_p norms, with $p \neq 2$, are able to perform deconvolution in some non-minimum-phase channels, but not in all of them.

With this limitation in mind, this work proposes the use of MFE (Mean Fourth Error) criterion, which is related to the ℓ_4 norm, as a smooth approximation of the ℓ_∞ norm, alongside with a cascade structure. This also allows the use of the LMF (Least Mean Fourth) adaptive algorithm, in order to deal with non-stationary channels.

Our work is organized as follows: first, we recall the predictive deconvolution problem, and the main limitations of the linear MSE approach. Then, we present the ℓ_p norms as alternative criteria for blind deconvolution, together with a procedure to choose the most suitable value for p. Having chosen the criterion, we present the cascade structure along with the rule to adjust its parameters. With our setup defined, we present our simulation results for the equalization of telecommunication signals. Finally, we present our concluding remarks.

II. Recall on predictive deconvolution

In simple terms, in the discrete-time context, the forward (one-step) linear predictor can be implemented by the structure of an FIR filter, which corresponds to the following difference equation:

$$\hat{x}_f(n) = \sum_{k=1}^{K} a_k x(n-k), \tag{1}$$

where the vector $\mathbf{a}$ contains the parameters $a_1, \ldots, a_K$, referred to as the forward prediction coefficients. Therefore, we can define the forward prediction error as:

$$e_f(n) = x(n) - \sum_{k=1}^{K} a_k x(n-k). \tag{2}$$

In a similar way, we can define the backward predictor filter as being:

$$\hat{x}_b(n-K) = \sum_{k=1}^{K} b_k x(n-k+1), \tag{3}$$

where the vector $\mathbf{b}$ contains the parameters $b_1, \ldots, b_K$, referred as to the backward prediction coefficients, and the backward prediction error is given by:

$$e_b(n) = x(n-K) - \sum_{k=1}^{K} b_k x(n-k+1). \quad (4)$$

The classical procedure to calculate the optimal prediction coefficients $\mathbf{a}$ or $\mathbf{b}$ consists in minimizing the mean squared error (MSE), $E[e_f^2(n)]$ or $E[e_b^2(n)]$, which leads to the well-known Yule-Walker equations [4].

It is also known that the forward and backward prediction error filters (PEF), when associated with the linear structure and the MSE criterion, present two important properties [7]:

- The PEF works as a whitening filter, i.e., the error signal $e(n)$ tends to be uncorrelated as the number of coefficients K increases.
- The forward and the backward PEF are FIR minimum- and maximum-phase filters, respectively.

To see how these properties are related to the deconvolution task, let us first introduce this problem.

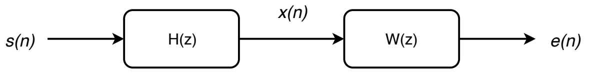


Fig. 1: Block diagram of blind deconvolution problem.

In Figure 1, we have an input signal $s(n)$ that is distorted by the channel $H(z)$, originating the signal $x(n)$. The objective is to project a deconvolution filter, $W(z)$, to compensate the channel, producing the signal $e(n)$, which is an estimate of $s(n)$. In the unsupervised version of the problem, the filter is adjusted without the complete knowledge of the input signal neither the channel.

Provided that the signal $s(n)$ has uncorrelated samples and the channel has minimum-phase (for the forward predictor) or maximum-phase (for the backward predictor) response, the PEF structure can be used as a deconvolution filter. As an example of predictive deconvolution, we have the pioneer work of Robinson [6], which uses a forward prediction error filter in geophysics signal processing.

However, decorrelation is sufficient for equalization only of minimum- or maximum-phase channels [7]. For more general scenarios, we must explore a stronger property: the independence between the samples of $s(n)$. In the following, we will present how the ℓ_p norms, with $p \neq 2$, explore the non-linear decorrelation, which goes further in the independence direction.

III. ALTERNATIVE CRITERIA FOR PREDICTIVE BLIND DECONVOLUTION

In [9], we investigated how the ℓ_p norms, with $p \neq 2$, can be used as alternative criteria for deconvolution. To do so, we have the following optimization problem:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_{n=0}^{T-1} |e(n)|^p = \arg\min_{\mathbf{w}} ||\mathbf{e}||_p^p. \quad (5)$$

Taking the gradient of (5) with respecto to $\mathbf{w}$, we have:

$$\nabla_{\mathbf{w}} J_{\ell_p} = \sum_{n=0}^{T-1} p|e(n)|^{p-1} sign\big(e(n)\big)\mathbf{x}(n-1). \quad (6)$$

Considering an ergodic process [10] and a large enough number of samples, the sum in (6) converges to an expectation operator [10], leading to:

$$\nabla_{\mathbf{w}} J_{\ell_p} = \mathbb{E}\Big[p|e(n)|^{p-1} sign\big(e(n)\big)\mathbf{x}(n-1)\Big]. \quad (7)$$

For the optimal predictor, we have:

$$\begin{gathered} \nabla_{\mathbf{w}} J_{\ell_p} = \mathbf{0} \\ \mathbb{E}\Big[p|e(n)|^{p-1} sign\big(e(n)\big)\mathbf{x}(n-1)\Big] = \mathbf{0}. \end{gathered} \quad (8)$$

Equation (8) shows that the minimization of (5) leads to a non-linear decorrelation between the prediction error and the input signal. As the past samples of the prediction error are linear combinations of $x(n-k)$, with $k \geq 1$, an ℓ_p PEF with an adequate order results in:

$$\mathbb{E}\Big[f\big(e(n)\big)e(n-k)\Big] = 0,\ k \geq 1, \quad (9)$$

with $f\big(e(n)\big) = p|e(n)|^{p-1} sign\big(e(n)\big)$.

According to equation (9), the ℓ_p prediction-error filters, with enough coefficients, perform a non-linear decorrelation between $e(n)$ and its past samples. Now, if we adopt

$$f(s) = \frac{d\big(log(p(s))\big)}{ds}, \quad (10)$$

with $p(s)$ as the probability density of input signal, in (9), we have that the non-linear decorrelation is sufficient to provide independence [11].

In addition to that, we can establish an interesting relation between the ℓ_p norms and the Maximum Likelihood for generalized Gaussian distributions [12], as we show next.

When applied to the deconvolution problem, the Maximum Likelihood criterion is given by:

$$\begin{aligned} \mathbf{W}^* &= \arg\max_{\mathbf{w}} J_{ML}(\mathbf{W}) \\ &= \arg\max_{\mathbf{w}} \frac{1}{R}\sum_{r=1}^{R}\sum_{n=0}^{T-1} log\Big(p_S\big(\mathbf{w_n}\mathbf{x}(n)\big)\Big) + log\big(|det\ \mathbf{W}|\big), \end{aligned} \quad (11)$$

where $\mathbf{W}$ represents the convolution matrix associated with the filter $w(n)$, $\mathbf{w_n}$ it is the n-th line of $\mathbf{W}$, with $n = 1, 2, \cdots T$; $p_S(\cdot)$ is the probability distribution of $s(n)$; $\mathbf{x} = [x(0)\ x(1)\ \cdots\ x(T-1)]^T$; T is the total number of samples available and R denotes the number of realizations.

For the particular case in which we use a forward prediction error filter, the convolution matrix corresponds to a triangular matrix with unit diagonal, which leads to:

$$det\ (\mathbf{W}) = 1. \quad (12)$$

Applying (12) in (11), we have:

$$\begin{aligned}\mathbf{W}^* &= \arg\max_{\mathbf{w}} \ J_{ML}(\mathbf{W}) \\ &= \arg\max_{\mathbf{w}} \ \frac{1}{R}\sum_{r=1}^{R}\sum_{n=0}^{T-1} log\Big(p_S\big(\mathbf{w_n x}(n)\big)\Big).\end{aligned} \tag{13}$$

The Maximum Likelihood criterion provides an unbiased and efficient estimator, *i.e.*, an estimator that is able to achieve the Cramer-Rao bound, as more samples are used [13], [7]. However, this criterion requires the explicit knowledge of the signal probability distribution. Fortunately, many distributions of practical interest can be unified by the generalized Gaussian [12], given by:

$$p_S\big(s(n)\big) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-\left(\frac{|s(n)-\mu|}{\alpha}\right)^{\beta}} \tag{14}$$

where μ is the mean value of the distribution, $\Gamma(\cdot)$ is the Gamma function, α is the spread parameter and β is the shape parameter, both related by:

$$\alpha^\beta = \beta\mathbb{E}\big[|s(n)-\mu|^\beta\big]. \tag{15}$$

For $\beta = 1$, (14) corresponds to the Laplacian distribution; if $\beta = 2$, we have the classic Gaussian distribution. Finally, as $p \to \infty$, the generalized Gaussian approaches the uniform distribution.

Applying the Maximum Likelihood criterion, with the natural logarithm, to the generalized Gaussian distribution (with $\mu = 0$ for simplicity), we have:

$$\begin{aligned}J_{ML}(\mathbf{W}) &= \frac{1}{R}\sum_{r=1}^{R}\sum_{n=0}^{T-1} ln\Big(p_S\big(\mathbf{w_n x}(n)\big)\Big) \\ &= RTln\Big(\frac{\beta}{2\alpha\Gamma(1/\beta)}\Big) - \frac{1}{\alpha^\beta}\frac{1}{R}\sum_{r=1}^{R}\sum_{n=0}^{T-1}|\mathbf{w_n x}(n)|^\beta \\ &= RTln\Big(\frac{\beta}{2\alpha\Gamma(1/\beta)}\Big) - \frac{1}{\alpha^\beta}\frac{1}{R}\sum_{r=1}^{R}\sum_{n=0}^{T-1}|e(n)|^\beta \\ &= RTln\Big(\frac{\beta}{2\alpha\Gamma(1/\beta)}\Big) - \frac{1}{\alpha^\beta}\frac{1}{R}\sum_{r=1}^{R}||e(n)||_\beta^\beta.\end{aligned} \tag{16}$$

Once the parameters α and β fixed, we have the following optimization problem:

$$\begin{aligned}\arg\max_{\mathbf{w}} \ J_{ML}(\mathbf{W}) &= \arg\max_{\mathbf{w}} \ -\frac{1}{\alpha^\beta}\frac{1}{R}\sum_{r=1}^{R}||e(n)||_\beta^\beta \\ &= \arg\min_{\mathbf{w}} \ \frac{1}{R}\sum_{r=1}^{R}||e(n)||_\beta^\beta.\end{aligned} \tag{17}$$

Equation (17) shows that maximizing the Likelihood criterion is equivalent to minimize the sample average of the ℓ_p norm, for $p = \beta$, of the prediction error $e(n)$. From this relation between p and β we have:

- For super Gaussian distributions [14], we choose $1 \le p < 2$, with $p = 1$ for the Laplacian distribution.
- For a Gaussian distribution, we adopt $p = 2$.
- For sub Gaussian distributions [14], we use $p > 2$, with $p \to \infty$ for the uniform distribution.

In [9] we provide a rather complete study on ℓ_p criteria for predictive deconvolution, which shows that the ℓ_p forward prediction error filters are able to perform the blind deconvolution of some non-minimum-phase systems. However, it still presents performance limitations according to the positions of the zeros of the channel. In a way, our result meets that of the work of Knockaert [15], which shows that the ℓ_p forward prediction error filters, with $p \neq 2$, have all their zeros inside a circle with radius 2 in the complex plane. In short, despite the value of p in the ℓ_p criterion, there is a limitation in using the ℓ_p forward prediction error filters which is intrinsically connected to its linear structure. Hence, together with the optimization criterion, the choice of a suitable structure for the deconvolution filter is also an important step in its project.

In this work, and inspired by some well-known works in both seismic deconvolution [16], [17], and channel equalization [18], [19], we propose the use of a cascade of linear forward and backward predictors. Comparing to the above mentioned works, we have two goals in mind: first, we aim to preserve the linearity of the filters, instead of using the non-linear approach in [18], [19], in order to have a relatively simple parameter optimization; second, we look for an adequate adaptive procedure to make the scheme effective in a non-stationary scenario.

IV. Cascade predictors and deconvolution

Fig.2 illustrates the proposed approach where, due to the commutative property, the forward predictor can be written in terms of the backward error and vice-verse. Since our interest concerns communications signals, which are generally assumed to present an independent and identically distribution, the ℓ_p criterion to be employed in the structure optimization must consider $p \to \infty$. We adopt $p = 4$, which corresponds to the MFE criterion, as a smooth approximation of the ℓ_∞ norm. This allows applying the LMF adaptive algorithm, in order to track non-stationary behaviors. The updating procedure can be simply derived from the scheme in Fig. 2, as shown in the following.

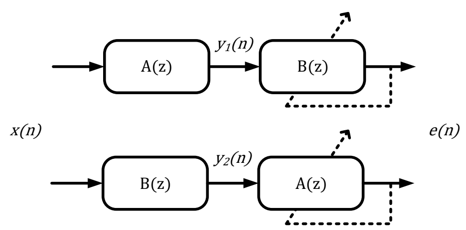


Fig. 2: Scheme of the proposed adaptive procedure.

$$e(n) = y_1(n-K) - \sum_{k=1}^{K} b_k y_1(n-k+1), \tag{18}$$

where $y_1(n) = x(n) - \sum_{k=1}^{K} a_k x(n-k)$, or, due to the commutative property:

$$e(n) = y_2(n) - \sum_{k=1}^{K} a_k y_2(n-k), \tag{19}$$

where $y_2(n) = x(n-K) - \sum_{k=1}^{K} b_k x(n-k+1)$.

From (18), we can express the MFE criterion by:

$$E[|e^4(n)|] = E\left[\left|\left(y_1(n-K) - \sum_{k=1}^{K} b_k y_1(n-k+1)\right)\right|^4\right], \tag{20}$$

or, from (19), by:

$$E[|e^4(n)|] = E\left[\left|\left(y_2(n) - \sum_{k=1}^{K} a_k y_2(n-k)\right)\right|^4\right]. \tag{21}$$

Hence, the update rule for the LMF [20] is simply given by:

$$\begin{aligned} a_i(n) &= a_i(n-1) + \\ &\quad \mu e(n)e^*(n)\Big(y_2(n-1-i)e^*(n) + y_2^*(n-1-i)e(n)\Big), \\ b_i(n) &= b_i(n-1) + \\ &\quad \mu e(n)e^*(n)\Big(y_1(n-i)e^*(n) + y_1^*(n-i)e(n)\Big), \end{aligned} \tag{22}$$

or, for real input symbols:

$$\begin{aligned} a_i(n) &= a_i(n-1) + \mu(e^3(n-1))y_2(n-1-i), \\ b_i(n) &= b_i(n-1) + \mu(e^3(n-1))y_1(n-i), \end{aligned} \tag{23}$$

where μ is the step-size.

Now we present our simulation results.

V. Results

First, we consider the communication channel switching over three different situations: minimum-phase to mixed-phase to maximum-phase responses. The minimum-phase transfer function is given by $H_1(z) = 1.8 - 1.2z^{-1} + 1z^{-2}$; the maximum-phase response corresponds to $H_3(z) = H_1(1/z)z^{-2}$; and the mixed-phase response corresponds to $H_2(z) = H_1(z)H_3(z)$. In all situations the channels are normalized by their ℓ_∞ norm and fed, each one, by 30.000 samples generated by an i.i.d. discrete uniformly distributed random variable with two symbols $[-1, 1]$. For the predictive equalizer we consider a 5-tap forward and 5-tap backward PEF, to be adapted by the LMF algorithm.

Besides the good performance of the algorithm, the first simulation indicates the relevance of a certain degree of a priori knowledge about the phase behavior of the channel, which is related to the distribution of their zeros inside and outside the unit circle. Having this in mind, we adopt as benchmarks the 6-tap and 11-tap best delay Wiener filters. The 11-tap one has been attained only for the mixed-phase channel. This is because, for the minimum-phase, only the 5-tap forward prediction structure intervenes in the equalization, the same way that only the 5-tap backward prediction error affects the equalization of the maximum-phase channel. So, the 6-tap Wiener filter represents the best solution to be attained with a 5- tap forward PEF and a 5-tap backward PEF respectively. Moreover, the 11-tap Wiener filter represents the best result to be attained for the cascade structure for general mixed phase channels.

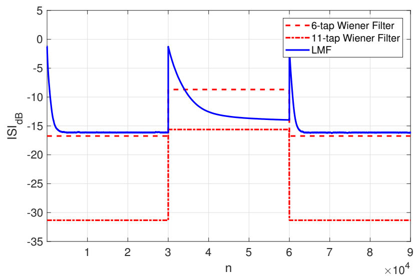


Fig. 3: ISI for the LMF algorithm ($\mu = 7.5e^{-4}$ - blue) and for the 6-tap and 11-tap best delay Wiener filter (red).

In order to illustrate the above discussion, Fig. 3 depicts the degree of intersymbol interference (ISI) along the whole 90.000 iterations of the algorithm. Such parameter is given by

$$\text{ISI}_{\text{dB}} = 10\log_{10}\frac{\sum_i |g_i|^2 - \max_i |g_i|^2}{\max_i |g_i|^2}, \tag{24}$$

where g_i are the coefficients of the combined response $G(z) = H(z)A(z)B(z)$.

The ISI is presented in red for the two considered benchmarks, *i.e.*, the 6-tap and the 11-tap Wiener filters. The blue curve corresponds to the cascade structure, adapted by the LMF algorithm, for 100 Monte Carlo (MC) simulations. As we can see, for the minimum and maximum phase channels, the cascade structure almost reaches the best performance to be attained with a 6-tap best delay Wiener filter. For the mixed-phase channel, its performance is between the 6 and 11-tap best delay Wiener filters.

In our simulations, we adopted the same number of taps for the forward and backward PEF in the cascade. By doing so, we are assuming, even implicitly, that the mixed-phase channel has the same number of zeros inside and outside the unit circle. In our next experiment, we evaluate how this symmetry interferes in the cascade performance.

To do so, let us now consider the channel switching over the three following mixed-phase transfer functions: $H_1(z)$, with zeros $z_1 = 0.5e^{j\pi/4}$, $z_2 = z_1^*$, $z_3 = 0.7$ and $z_4 = 1.5$; $H_2(z)$, with zeros $z_1 = 0.5e^{j\pi/4}$, $z_2 = z_1^*$, $z_3 = 1.5e^{j\pi/3}$ and $z_4 = z_3^*$ and $H_3(z)$ with zeros at $z_1 = 0.5$, $z_2 = 1.5e^{j\pi/4}$, $z_3 = z_2^*$ and $z_2 = 1.5$ (all of them normalized by their ℓ_∞ norm). We show the results in Figure 4.

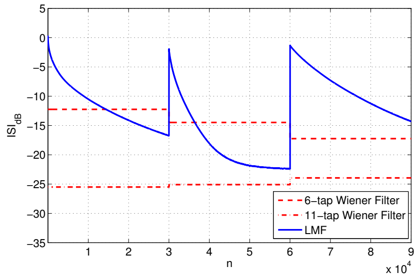
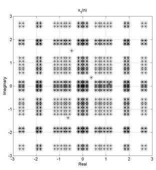


Fig. 4: ISI for the LMF algorithm ($\mu = 7.5e^{-4}$ - blue) and for the 6-tap and 11-tap best delay Wiener filter (red) for non-stationary mixed-phase channels (variable zeros arrangement).

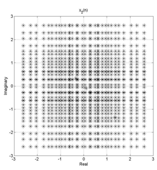
From the above result, we can see that the arrangement of the zeros has a considerable effect on the cascade performance. When we have more zeros inside ($H_1(z)$) or outside ($H_3(z)$) the unit circle, the cascade is close to the 6-tap Wiener filter. For the case in which the channel has the same number of zeros inside and outside the unit circle, and therefore match to our assumption for the number of forward and backward taps, the result is between the 6-tap and 11-tap Wiener filter.

Finally, in order to illustrate the behavior of the proposed technique with noisy channels and multilevel modulation the next presented result considers 4-QAM with symbols $\pm 1 \pm j$. We considered again the same switching channels of the previous simulation (normalized by ℓ_∞ norm), with the addition of a white Gaussian noise, with zero mean, and a signal-to-noise ratio of 30 dB.

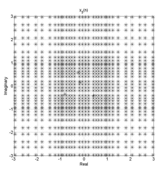
Figure 5 shows the outputs $x_i(n)$ of the corresponding channel $H_i(z)$, for $i = 1, 2, 3$ (again with 100 MC simulations). The effect of the severe distortions is clear, rendering the system to a closed-eye condition [21].



(a) $x_1(n)$



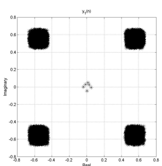
(b) $x_2(n)$



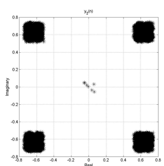
(c) $x_3(n)$

Fig. 5: Channel outputs

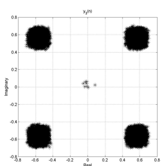
Figure 6 presents the recovered signals $e_i(n)$, $i = 1, 2, 3$, after 30.000, 60.000 and 90.000 iterations, respectively.



(a) $e_1(n)$



(b) $e_2(n)$



(c) $e_3(n)$

Fig. 6: Recovered signals

As we can see, the recovered signals are formed by four well defined clouds, *i.e.*, in a opened-eye condition. These clouds are centered at points different from $\pm 1 \pm j$, which indicates the necessity of a gain stage after the cascade [19]. Besides, we can see a few points around the complex plane origin, due to filtering transient effects.

VI. Conclusions

In this paper we proposed an alternative solution for predictive blind deconvolution of non-minimum-phase and non-stationary channels driven i.i.d. discrete signals. The proposed solution is based on the cascade of forward and backward prediction error filters, the parameters of which are optimized by the MFE. In order to deal with a non-stationary scenario, such optimization is carried out by means of the LMF adaptive algorithm. The presented results confirmed the performance of the proposed solution in the blind deconvolution of non-stationary and non-minimum-phase channels, with results close to the ones attained with the best delay Wiener filters. The presented analysis and results have shown that a cascade of linear ℓ_p PEFs, associated to a simple adaptive procedure, is able to achieve blind deconvolution in cases in which classical solutions fail. This is a stimulating issue, since it opens interesting perspectives both in practical and theoretical points of view.

References

[1] A. N. Kolmogorov, "Sur l'interpolation et l'extrapolation de suites stationaires," *Comptes Rendus de l'Académie des Sciences*, vol. 207, pp. 2043–2045, 1939.

[2] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series*. New York: Wiley, 1949.

[3] J. Makhoul, "Linear prediction: a tutorial review," *Proceedings of the IEEE*, vol. 63, pp. 561–580, 1975.

[4] M. Bellanger, *Adaptive Filters and Signal Analysis*. Dekker, 1988.

[5] P. P. Vaidyanathan, *The theory of linear prediction*. Morgan & Claypool Publishers, 2008.

[6] E. A. Robinson, *Predictive decomposition of time series with applications to seismic exploration*. PhD thesis, Massachusetts Institute of Technology, 1954.

[7] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, *Unsupervised signal processing: channel equalization and source separation*. CRC Press, 2011.

[8] R. D. B. Brotto, K. Nose-Filho, and J. M. T. Romano, "Critérios alternativos para desconvolução não supervisionada," *8º Simpósio de Instrumentação e Imagens Médicas/ 7º Simpósio de Processamento de Sinais*, 2017.

[9] R. D. B. Brotto, K. Nose-Filho, and J. M. T. Romano, "Alternative criteria for predictive blind deconvolution," *Journal of Communication and Information Systems*, vol. 33(1), pp. 271–281, 2018.

[10] S. M. Kay, *Intuitive Probability and Random Processes Using MATLAB*. Springer, 1ª ed., 2006.

[11] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2001.

[12] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, n.7, pp. 685–694, 2005.

[13] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Europe, 2002.

[14] P. Comon and C. Jutten, eds., *Handbook of blind source seapartion*. Academic Press, 2010.

[15] L. Knockaert, "Stability of linear predictors and numerical range shift operators in normal spaces," *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1483–1486, 1992.

[16] M. J. Porsani, P. E. M. de Melo, and T. J. Ulrych, "Highly resolved deconvolution via a sparsity norm," *Journal of Seismic Exploration*, vol. 12, pp. 127–140, 2003.

[17] Y. Zhang and J. Claerbout, "A new bidirectional deconvolution method that overcomes the minimum phase assumption," tech. rep., SEP-142, 2011.

[18] O. Macchi, C. A. Faria, and J. M. T. Romano, "Egalisation adaptive autodidacte par rétroprédiction et prédiction," in *Proceddings of the Colloque GRETSI*, 1993.

[19] C. A. F. da Rocha, *Técnicas prditivas para equalização autodidata*. PhD thesis, Faculdade de engenharia elétrica e de computação, Universidade de Campinas, 1996.

[20] A. H. Sayed, *Fundamentals of adaptive filtering*. Wiley, 2003.

[21] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill Education, 5th ed., 2007.