

Alternative Criteria for Predictive Blind Deconvolution

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

Abstract—Blind deconvolution is a major theme in signal processing and has been intensely investigated over the last decades. Among its several applications, we can mention the problem of seismic deconvolution and channel equalization in telecommunications. In these two cases, predictive techniques have been studied by different authors, and presented satisfactory results when some suitable conditions were fulfilled. In fact, the predictive deconvolution structure, when associated with the classical mean squared error criterion is only effective when the distortion system is minimum phase. In the case of nonminimum phase systems, it only provides magnitude equalization, but the phase response remains distorted. In order to overcome this problem, we present in this work some interesting results obtained with the use of ℓ_p norms, with $p \neq 2$, as optimization criteria.

First we demonstrate that the ℓ_p prediction error filter works as the Maximum Likelihood solution for blind deconvolution when the signal to be recovered has a generalized Gaussian distribution, with *i.i.d.* (identically and independently distributed) samples. From this, we show how the best p can be chosen according to the signal distribution. Then we further investigate the phase response of the ℓ_p filter, emphasizing its potential as well as some limitations in dealing with blind deconvolution, even for nonminimum phase distortion system. Finally, some performance simulations results are provided.

Index Terms—Predictive deconvolution, ℓ_p norms, nonminimum phase deconvolution.

I. INTRODUCTION

The problem of estimating future values of a time series from its present and past samples is a very challenging problem in the field of signal processing. Its origins goes back to the works of Kolmogorov [1], in the context of stationary discrete time signals, and Wiener [2], where the continuous time optimal predictor is formulated.

The objective in prediction consists in finding a mapping $F[\cdot]$ that, when applied to a set of past samples of a time series $\mathbf{x}(n-1) = [x(n-1) \ x(n-2) \ \cdots \ x(n-L)]^T$, results in an estimative to its present value $x(n)$ [3], [4]:

$$F[\mathbf{x}(n-1)] = \hat{x}(n). \tag{1}$$

The generic mapping in (1) may be associated to a particular filter structure, with a set of parameters to be adjusted. For the case of linear prediction, this mapping can be represented by a finite impulse response (FIR) filter, namely prediction filter, which input-output relation is given by:

$$\hat{x}(n) = \sum_{i=1}^L w_i x(n-i) = \mathbf{w}^T \mathbf{x}(n-1), \tag{2}$$

where w_1, w_2, \dots, w_L are the prediction coefficients, $x(n-i)$, for $i = 1, 2, \dots, L$, are the past samples of $x(n)$, and $\hat{x}(n)$ is its estimative. Moreover, we can define $\mathbf{w} = [w_1 \ w_2 \ \cdots \ w_L]^T$ as the vector of parameters to be adjusted. The difference between $x(n)$ and $\hat{x}(n)$ leads to the so called prediction error [5]:

$$e(n) = x(n) - \hat{x}(n) = x(n) - \sum_{i=1}^L w_i x(n-i). \tag{3}$$

Meanwhile (2) stands for the prediction filter, the input-output relation given by (3) stands for what we call the prediction error filter (PEF), which structure is illustrated by the diagram of Figure 1.

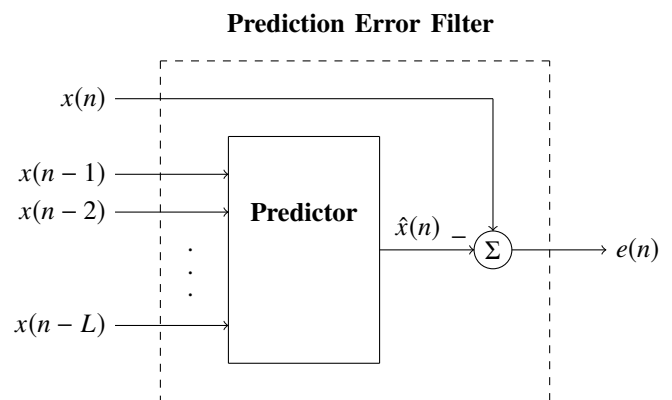


Fig. 1: The prediction filter consists in mapping the past samples of $x(n)$ in its estimative $\hat{x}(n)$. The prediction error filter stands for the mapping of $x(n)$ and its past samples in the prediction error $e(n)$.

The authors acknowledge CAPES and FAPESP (2017/13025-2) for financial support

Renan D. B. Brotto and João M. T. Romano are with the School of Electrical and Computer Engineering, State University of Campinas (UNICAMP), campus Campinas, São Paulo, Brazil (e-mail: {rbrotto, romano}@decom.fee.unicamp.br)

Kenji Nose-Filho is with Center of Engineering, Modeling and Applied Social Sciences, Federal University of ABC (UFABC), campus Santo André, São Paulo, Brazil (e-mail: kenji.nose@ufabc.edu.br)

Digital Object Identifier: 10.14209/jcis.2018.27

The problem of linear prediction, as described above, is of great relevance for the field of signal processing, and can be applied for both parametric and nonparametric models. It can be used to estimate the parameters of autoregressive (AR)

models [6], such as in the linear predictive coding of speech signals; in time series forecasting [7]; and in the deconvolution problem, which is the problem of interest in this work [3].

It is well known in the literature that the classical prediction error filter obtained through the mean squared error (MSE) criterion is only effective for the deconvolution of minimum phase systems, since this filter provides decorrelated output samples. Hence, for signals that have been distorted by non-minimum phase systems, the classical PEF equalizes only magnitude, but the phase response remains distorted.

To overcome this problem, we need to explore a stronger hypothesis: the independence between the samples. This can be attained by means of a nonlinear PEF structure adjusted by the MSE criterion, as proposed in [8], [9], or rather, maintaining the linear PEF structure but employing alternative criteria. We have investigated the latter solution in [10] and presented some preliminary results. This work introduced an approach based on ℓ_p PEF, showing that it works as a Maximum Likelihood solution for blind deconvolution when the signal to be recovered has a generalized Gaussian distribution.

The present paper extends the study in [10], investigating the zero location and phase response of the ℓ_p PEF, and presents several new results. Such results concern specially a deeper assessment about the zero location (and then, the phase response) of the ℓ_p PEF, when the value of p is adjusted according to the properties of sparsity and antisparcity [11] of the signal to be recovered. All the results have been obtained in a noise free scenario with real valued signals.

The paper is organized as follows: In Section II we revisit the principle and main properties of the classical MSE PEF. Then, in Section III, we introduce the problem of predictive deconvolution, emphasizing the limits of the MSE solution. In Section IV we demonstrate that the ℓ_p PEF works as the Maximum Likelihood solution for blind deconvolution when the signal to be recovered has a generalized Gaussian distribution; from this, we show how the best p can be chosen in according to the signal distribution. Then, in Section V, we provide a deeper study about the phase response of the ℓ_p filter, emphasizing its potential as well as some limitations in dealing with blind deconvolution, even for nonminimum phase distortion system. In Section VI we evaluate the performance of the ℓ_1 deconvolution for a sparse signal and the ℓ_∞ deconvolution for an antisparse signal, considering first and second order channels, with minimum and maximum phase responses. Finally, in Section VII we close the paper with some conclusion remarks and perspectives of future works.

II. THE MEAN SQUARED PREDICTION ERROR FILTER REVISITED

The prediction coefficients $w_i, i = 1, \dots, L$ are adjusted in order to minimize a given cost function $J(\mathbf{w})$, which, classically, is chosen as the mean squared error criterion:

$$\begin{aligned} J(\mathbf{w}) &= \mathbb{E} [e^2(n)] = \mathbb{E} \left[\left(x(n) - \mathbf{w}^T \mathbf{x}(n-1) \right)^2 \right] = \\ &= \mathbb{E} [x(n)x(n)] - \mathbb{E} [x(n)\mathbf{x}^T(n-1)] \mathbf{w} + \\ &- \mathbf{w}^T \mathbb{E} [\mathbf{x}(n-1)x(n)] + \\ &+ \mathbf{w}^T \mathbb{E} [\mathbf{x}(n-1)\mathbf{x}^T(n-1)] \mathbf{w}, \end{aligned} \quad (4)$$

where $\mathbb{E}[\cdot]$ stands for the expectation operator.

Besides the signal variance $\sigma_x^2 = \mathbb{E}[x(n)x(n)]$, other two terms completely characterize $J(\mathbf{w})$: the autocorrelation matrix $\mathbf{R} = \mathbb{E} [\mathbf{x}(n-1)\mathbf{x}^T(n-1)]$, and the autocorrelation vector $\mathbf{p} = \mathbb{E} [\mathbf{x}(n-1)x(n)]$. Rewriting (4) in terms of σ_x^2 , \mathbf{R} , and \mathbf{p} results in:

$$J(\mathbf{w}) = \sigma_x^2 - \mathbf{p}^T \mathbf{w} - \mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R} \mathbf{w}. \quad (5)$$

The cost function in (5) is a quadratic function of the adjustable parameters and describes an elliptic paraboloid with one single minimum point. To determine this point, we set the gradient equal to the null vector:

$$\nabla J(\mathbf{w}) = 2\mathbf{R}\mathbf{w} - 2\mathbf{p} = \mathbf{0} \rightarrow \mathbf{R}\mathbf{w} = \mathbf{p}. \quad (6)$$

The solution of (6) is known in the literature as Wiener's solution [3] and is given by:

$$\mathbf{w}^* = \mathbf{R}^{-1} \mathbf{p}. \quad (7)$$

Now, in order to introduce the relationship between linear prediction and deconvolution, let us consider two properties of the mean squared prediction error filter, given by (7).

Due to the equivalence between MSE minimization and the orthogonalization procedure, the error produced by the Wiener filter is uncorrelated with its input [3]:

$$\mathbb{E} [e(n)x(n-i)] = 0, \quad 1 \leq i \leq L. \quad (8)$$

Since the prediction error $e(n)$ is a linear combination of the past samples of $x(n)$, we have, from (3) and (8):

$$\mathbb{E} [e(n)e(n-k)] = 0, \quad \forall k > 1, \quad (9)$$

which means that, for a sufficient number of coefficients L , the PEF tends to generate an uncorrelated (white) output error.

Another property of the prediction error filter is related to its phase response: in fact, it is possible to show that all of its zeros lay inside the unit circle, so that the PEF is a minimum phase filter [5].

III. PREDICTIVE DECONVOLUTION AND BLIND EQUALIZATION OF MINIMUM PHASE SYSTEMS

These two properties of the mean squared prediction error filter lead us to an interesting result: the possibility of removing, or attenuating, the effects of linear and time invariant systems. This problem is known as the predictive deconvolution [3]. The first application of predictive deconvolution goes back to Robinson's work [12] in the context of seismic deconvolution.

In seismic data acquisition, an artificially and controlled wave is emitted and, due to the difference between the acoustic

impedance of the different earth layers, part of its energy is reflected back to the surface. The reflected signals are captured by sensors, originating the so called seismic traces. Therefore, the seismic trace corresponds to the convolution between the seismic wavelet and the earth impulse response.

From convolution commutative property [13] we can consider the earth impulse response as an input signal $s(n)$ and the seismic wavelet as a system impulse response $h(n)$, as shown in Figure 2:

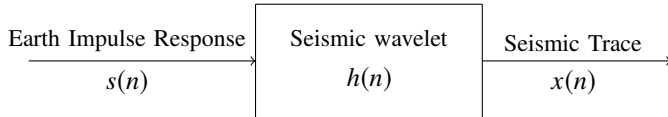


Fig. 2: Simple model of the seismic data acquisition. The seismic trace $x(n)$ can be seen as the convolution of the seismic wavelet impulse response $h(n)$ with the earth impulse response $s(n)$.

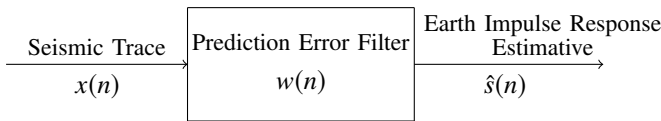


Fig. 3: Simple model of seismic deconvolution. The prediction error filter is applied in order to remove the effects introduced by the seismic wavelet, providing an estimative of the earth impulse response.

So, the output seismic trace is given by:

$$x(n) = s(n) * h(n). \quad (10)$$

Robinson's approach consists in applying a prediction error filter to the signal $x(n)$ in order to remove, or attenuate, the effects introduced by the seismic wavelet. Thus, the recovery of the earth impulse response is based in two hypotheses [12]:

- 1) The earth impulse response behaves like a white noise (i.e. all of its samples are uncorrelated);
- 2) The seismic wavelet corresponds to the impulse response of a minimum phase system.

Under these two simplifying hypothesis, it is possible to show that a sufficient length prediction error filter is able to remove the effects introduced by the seismic wavelet, as depicted in Figure 3.

Another related problem of interest is that of *Channel Equalization*, in which the main goal is to retrieve the signal transmitted through a communication channel, usually modeled by an FIR filter. The received signal is given by the convolution of the transmitted signal with the impulse response of the channel. In this case, a prediction error filter can be used to perform deconvolution if the transmitted signal is composed of a sequence of uncorrelated samples and the channel is minimum phase.

An important difference between the above mentioned problems, seismic deconvolution and channel equalization, is related to the characteristics of the input signals. In seismic deconvolution, the samples of the earth impulse response can

be modeled by *i.i.d.* sparse and continuous random variables [14]. On the other hand, the samples of the transmitted signal in the channel equalization problem are usually modeled by *i.i.d.* discrete and uniformly distributed random variable [15]. However, in both cases, the seismic wavelet and the communication channel may be given by a nonminimum phase system [16], [17].

In the case of nonminimum phase systems, the mean squared prediction error filter is not able to remove nor attenuate the effects of the seismic wavelet or the communication channel. It provides only amplitude equalization, i.e., it acts as a whitening filter only. In other words, to explore the decorrelation property of the signal to be recovered is not sufficient to provide blind deconvolution when the distortion system is nonminimum phase. To overcome this limitation, we aim to search for a cost function that could explore a stronger property of the desired signal $s(n)$: the *i.i.d.* hypothesis. In fact, while correlation provides only second order statistic information, exploring independence implies dealing with all the statistic behavior of the signal, so that it could be properly recovered, as explained next.

IV. THE ℓ_p PEF AS AN ALTERNATIVE FOR BLIND DECONVOLUTION

As posed in the previous sections, the MSE criterion is given by the equation (4) and deals with the second order statistic behavior of the signals. Once such information is not always available in practical scenarios, an alternative consists in using temporal means of the prediction error signal, as follows:

$$J_{LS}(\mathbf{w}) = \sum_{n=0}^{N-1} \lambda(n) e^2(n), \quad (11)$$

where $\lambda(n)$ is a weight factor that controls the degree of relevance of the error sample at time instant n , $e(n)$ is the error signal and \mathbf{w} are the filter coefficients.

Considering stationary and ergodic processes [18], the criteria given by (4) and (11) are equivalent for practical purposes. So, minimizing the MSE criterion is equivalent to minimize the *least squares* (LS) criterion in (11).

A particular version of (11) is obtained when we use $\lambda(n) = \frac{1}{N}$, for all n :

$$J_{LS}(\mathbf{w}) = \frac{1}{N} \sum_{n=0}^{N-1} e^2(n). \quad (12)$$

From an optimization point of view, the scaling factor $1/N$ does not change the value for the optimal filter \mathbf{w}^* and so we have an equivalence between the two optimization problems:

$$\arg \min_{\mathbf{w}} \frac{1}{N} \sum_{n=0}^{N-1} e^2(n) = \arg \min_{\mathbf{w}} \sum_{n=0}^{N-1} e^2(n). \quad (13)$$

Considering the definition of ℓ_p norm [19]

$$\|e\|_p = \left(\sum_{n=0}^{N-1} |e(n)|^p \right)^{1/p}, \quad (14)$$

we have that the right hand side in equation (13) corresponds to the square of ℓ_2 norm of the prediction error. From the previous analysis, we have that, for our purposes here, the MSE criterion is equivalent to the ℓ_2 norm.

As said in the previous section, the ℓ_2 norm (or MSE criterion) can only provide linear decorrelation, what can be seen as a weak independence, and is efficient when dealing only with minimum phase systems. From this we can verify two restrictions of the ℓ_2 norm: one with respect to a signal property (linear decorrelation), and the second as a system restriction (minimum phase).

Our objective is to investigate how ℓ_p norms, with $p \neq 2$, can deal with these restrictions. For this, let us analyze these norms from two different perspectives:

- 1) ℓ_p norm as a Maximum Likelihood estimator for generalized Gaussian distribution, in order to take advantage of the signal probability distribution as well as sparsity properties;
- 2) ℓ_p PEF, with $p \neq 2$, as a nonminimum phase system, showing the potentiality to deal with the minimum phase restriction.

In the next subsections we show the development of these two perspectives of the ℓ_p prediction error filters, as summarized in Figure 4:

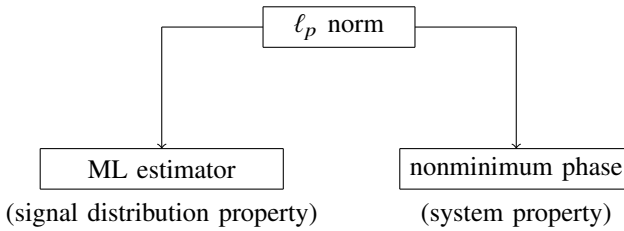


Fig. 4: Summary of the two perspectives of ℓ_p PEF: Maximum Likelihood estimators and nonminimum phase response.

A. ℓ_p norms and Maximum Likelihood Criterion

The Maximum Likelihood (ML) criterion can be viewed as a particular case of the Infomax principle [20]. It was derived in the context of neural networks and used in the Blind Source Separation problem [21], [3], which can be seen as an extension of the deconvolution problem. To obtain the ML criterion, we just need to set the nonlinearities $f_i(\cdot)$ as the cumulative distribution functions of the sources [3]. Cardoso's work (1998) [22] presents a deep study (in the context of BSS) of this criterion, given by:

$$\max_{\mathbf{W}} J_{Infomax}(\mathbf{W}) = \mathbb{E} \left[\sum_{i=1}^M \log(p_{s_i}(\mathbf{w}_i \mathbf{x})) \right] + \log(|\det(\mathbf{W})|), \quad (15)$$

where $p_{s_i}(\cdot)$ is the probability distribution function of the i -th source, for $i = 1, 2, \dots, M$.

By applying this criterion into the deconvolution task, the following optimization problem holds:

$$\max_{\mathbf{W}} J_{ML} = \max_{\mathbf{W}} \mathbb{E}_N \left[\sum_{i=1}^N \log(p_s(\mathbf{w}_i \mathbf{x})) \right], \quad (16)$$

where \mathbb{E}_N represents the temporal mean of the N available data; \mathbf{W} is the convolution matrix associated with the filter $w(n)$; \mathbf{w}_i is its i -th row; $p_s(\cdot)$ is the probability distribution of the input signal $s(n)$; and $\mathbf{x} = [x(0) \ x(1) \ \dots \ x(N-1)]^T$.

The Maximum Likelihood criterion provides an unbiased and efficient estimator, i.e., its variance reaches the Crammer-Rao limit as the number of samples goes to infinity [3], [23]. However, it requires the explicit knowledge of the input signal distribution function. Fortunately, many distributions with real applications can be unified in the generalized Gaussian distribution [24], given by:

$$p_S(s) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-\left(\frac{|s-\mu|}{\alpha}\right)^\beta}, \quad (17)$$

where α is the dispersion parameter and β is the shape parameter, both related by:

$$\alpha^\beta = \beta \mathbb{E}[|s|^\beta]. \quad (18)$$

For $\beta = 1$, (17) corresponds to the Laplacian distribution; if $\beta = 2$, we have the classical Gaussian distribution; and as $\beta \rightarrow \infty$, we obtain the uniform distribution.

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

$$\begin{aligned} J_{ML}(\mathbf{W}) &= \mathbb{E}_N \left[\sum_{i=1}^N \ln(p_s(\mathbf{w}_i \mathbf{x})) \right] = \\ &= \mathbb{E}_N \left[\sum_{i=1}^N \ln\left(\frac{\beta}{2\alpha\Gamma(1/\beta)}\right) - \frac{|\mathbf{w}_i \mathbf{x}|^\beta}{\alpha^\beta} \right] = \\ &= N \ln\left(\frac{\beta}{2\alpha\Gamma(1/\beta)}\right) - \frac{1}{\alpha^\beta} \mathbb{E}_N \left[\sum_{i=1}^N |\mathbf{w}_i \mathbf{x}|^\beta \right] = \\ &= N \ln\left(\frac{\beta}{2\alpha\Gamma(1/\beta)}\right) - \frac{1}{\alpha^\beta} \mathbb{E}_N \left[\|e\|_\beta^\beta \right]. \end{aligned} \quad (19)$$

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

$$\max_{\mathbf{W}} J_{ML}(\mathbf{W}) = \max_{\mathbf{W}} -\mathbb{E}_N \left[\|e\|_\beta^\beta \right] = \min_{\mathbf{W}} \mathbb{E}_N \left[\|e\|_\beta^\beta \right]. \quad (20)$$

Equation (20) shows that maximizing the Likelihood criterion is equivalent to minimize the temporal mean of the associated ℓ_p norm. Also, this equation establishes a relation between the signal distribution and the more suitable value for p , as summarized next:

- For super Gaussian distributions [21], we choose $1 \leq p < 2$, with $p = 1$ for Laplacian distribution.
- For a Gaussian distribution, we choose $p = 2$;
- For sub Gaussian distributions [21], we adopt $p > 2$, with $p \rightarrow \infty$ for uniform distribution.

From these results we observe two dual potential scenarios for predictive deconvolution:

- 1) the case where the desired signal is super Gaussian (sparsity property to be explored) and an ℓ_p predictor may be used with $1 \leq p \leq 2$ and

- 2) the case where the desired signal is sub Gaussian (anti-sparsity property [11] to be explored) and it is suitable to use $p > 2$.

B. ℓ_p norms and the PEF phase response

The location of the PEF zeros has been investigated in a number of works, especially during the 70s and 80s [25]–[27]. Usually, in the context of application of these works, the classical MSE (or ℓ_2) criterion was employed.

In order to get phase information in the deconvolution problem, alternative approaches have explored the use of ℓ_p norms, with $p \neq 2$, as a criterion to optimize the PEF coefficients. To the best of our knowledge, the work by Scargle [28] derived the first results on the location of ℓ_p PEF’s zeros, for $p \neq 2$. The work of Bednar [29] showed an application in seismic predictive deconvolution. The work of Knockaert [30] showed that all ℓ_p predictors (with $p \neq 2$) are unstable, in the sense that some of its zeros are outside the unit circle. Knockaert’s paper also claimed that such zeros cannot go outside of a circle with radius 2. Hence, Knockaert’s works shows the potentiality of ℓ_p (with $p \neq 2$) PEF, since the predictor can put some zeros outside the unit circle, but also shows a restriction: the filter zeros cannot be located in an arbitrary point of the complex plane.

These instigating results motivate us to perform a deeper study about the ℓ_p PEF zeros location, analyzing how flexible the filter phase response can be and how much it impacts the deconvolution performance. Next, we show the results obtained from this study in our context of application.

V. ASSESSMENT ON THE BEHAVIOR OF THE ℓ_p PEF PHASE RESPONSE

In order to study the ℓ_p PEF phase response, this section has a twofold objective.

First, we obtain a set of simulations results. Then, we assess more deeply the behavior of the PEF zeros by approximating a situation in which we have a one pole distortion system. The proposed FIR approximation allows us to study the location of the zeros of the ℓ_1 and ℓ_∞ PEFs, according to this pole position.

A. Duality between ℓ_1 and ℓ_∞ filters

With the optimization criteria chosen as the most suitable for each case, our objective here is to study the location of the ℓ_1 and ℓ_∞ PEF zeros. For comparison, we have also considered the supervised case. In this situation, we perform ℓ_1 and ℓ_∞ minimization of the error signal, which is produced by having access to the signal we are aiming to recover, possibly with a delay:

$$e(n)_{\text{supervised}} = s(n - d) - y(n), \tag{21}$$

with $y(n) = w(n) * x(n)$ as the filter output signal and d as the delay in the reference signal.

The sparse signal is generated accordingly to a Bernoulli-Gaussian distribution: we first generate a Bernoulli sequence with $Pr\{s(n) = 1\} = 0.02$; then, we multiply each element of

this sequence by a random variable $N(0, 1)$ (note that we have a different realization of this variable for each sample of the sequence). For the uniform signal, we have chosen a binary alphabet $(+1, -1)$.

The experiments consider three typical cases of channel:

- Minimum phase: $H_1(z) = 1 + 0.5z^{-1} + 0.3z^{-2}$;
- Mixed phase: $H_2(z) = 0.5 + 1z^{-1} + 0.3z^{-2}$;
- Maximum phase: $H_3(z) = 1 + 1.5z^{-1} + 1.8z^{-2}$.

For the unsupervised case, we used a PEF with one zero (and so, just one adjustable coefficient, p_1), i.e., $P(z) = 1 - p_1z^{-1}$. In the supervised scenario, we have also employed a deconvolution filter with one zero (but two adjustable coefficients, w_0 and w_1): $W(z) = w_0 + w_1z^{-1}$. After the optimization, we have normalized the filter in w_0 , just for comparison with the corresponding PEF.

A very important issue for the supervised case is the optimal delay in the reference signal: zero delay for the minimum phase, one sample for the mixed phase and three samples for the maximum phase channel.

To minimize the ℓ_1 norm, the filter coefficients were adjusted by the error signal LMS [31], while the ℓ_∞ PEFs coefficients were minimized with the populational meta heuristic of differential evolution [32], described in the Appendix of this paper.

The obtained results are depicted in Table I, where $P_1(z)$ and $P_\infty(z)$ denote the PEFs obtained by minimizing the ℓ_1 and ℓ_∞ norms, respectively, and $W_1(z)$ and $W_\infty(z)$ denote the filters obtained by minimizing the error ℓ_1 and ℓ_∞ norms, respectively, in the supervised case:

TABLE I: Deconvolution filters obtained in sparse and antisparsity deconvolution for both supervised and unsupervised scenarios.

Supervised	Unsupervised
Minimum phase channel	
$W_1(z) = 1 - 0.4997z^{-1}$	$P_1(z) = 1 - 0.5z^{-1}$
$W_\infty(z) = 1 - 0.5000z^{-1}$	$P_\infty(z) = 1 - 0.5z^{-1}$
Mixed phase channel	
$W_1(z) = 1 - 0.2993z^{-1}$	$P_1(z) = 1 - 0.3z^{-1}$
$W_\infty(z) = 1 - 0.2997z^{-1}$	$P_\infty(z) = 1 - 0.3z^{-1}$
Maximum phase channel	
$W_1(z) = 1 - 1.4969z^{-1}$	$P_1(z) = 1 - 1.2z^{-1}$
$W_\infty(z) = 1 - 1.4885z^{-1}$	$P_\infty(z) = 1 - 1.2z^{-1}$

From the Table I we have two interesting results:

- Minimizing the ℓ_1 norm to retrieve a sparse signal leads to the same filters obtained by minimizing the ℓ_∞ norm for deconvolution of an antisparsity signal, for all channels considered, in both supervised and unsupervised scenarios. This result confirms the duality relationships involving the ℓ_p norms and the signal distributions.
- Comparing the supervised and unsupervised filters, adjusted by both ℓ_1 and ℓ_∞ norms, we have the same zeros for the minimum (zero at 0.5) and mixed (zero at 0.3) phase channels. On the other hand, for maximum phase response, the filters obtained are slightly different: the supervised one put its zero at 1.5 (for both criteria), while the PEF put its zero at 1.2 (for ℓ_1 and ℓ_∞). From this difference we can see a limitation of the ℓ_p PEF, with

$p = 1$ or $p = \infty$, but also, we can see the potential of such a structure, since it positioned a zero outside the unit circle. This potential is deeply explored in the following.

B. ℓ_p prediction error filter zeros

In order to proceed with our study, let us first consider a very simple case of distortion system that can be easily deconvolved by an MSE PEF: a first order AR channel

$$H(z) = \frac{1}{1 - az^{-1}}. \tag{22}$$

Deconvolution is clearly attained by applying a single zero ℓ_p PEF:

$$W(z) = 1 - w_1z^{-1}. \tag{23}$$

Clearly, $W(z)$ will put its zero on the same position of the channel pole, i.e., $w_1 = a$, to eliminate all system distortion.

It is worth observing that the channel in (22) can be arbitrarily approximated by an FIR filter as follows:

$$\frac{1}{1 - az^{-1}} \approx 1 + az^{-1} + a^2z^{-2} + \dots + a^Kz^{-K} = \sum_{i=0}^K (az^{-1})^i. \tag{24}$$

This kind of approximation is particularly suitable in our study since it allows to simulate the effect of a pole outside the unit circle, i.e., the case where $|a| > 1$ in (22).

We can rewrite equation (24) in a more compact form:

$$H(z) = \sum_{i=0}^K (az^{-1})^i = \frac{1 - (az^{-1})^{K+1}}{1 - az^{-1}}. \tag{25}$$

Equation (25) shows that the channel zeros correspond to the $K + 1$ complex roots [33], except the one at a . Therefore, the channel tries to approximate a pole by a crown of complex and symmetric zeros of magnitude a , except at $z = a$. When the complete inversion is achieved, the zeros of the combined response $H(z)W(z)$ (channel and filter) in the complex plane will form a complete crown with radius a .

To illustrate it, let us consider a channel of order $K = 15$ (for symmetry purposes), adopting the ℓ_2 norm as optimization criterion and adjusting the filter coefficients by differential evolution¹, with a population of 30 individuals in 800 generations. In this first case, we have used $a = 0.7$ (so the channel is minimum phase) and we have adopted an input signal with a Gaussian distribution $N(0, 1)$. The result of this example is shown if Figure 5:

¹When we use ℓ_2 norm, we have a convex optimization problem, with a closed solution and so it is not necessary the use of a heuristic. But the differential evolution has shown a faster convergence when compared to the LMS algorithm, which, alongside the simplicity of this meta heuristic, justifies its application to this problem.

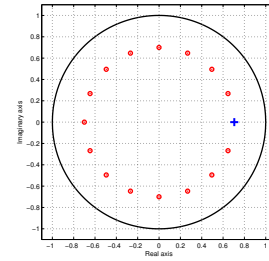


Fig. 5: Diagram of the channel (red circles) and filter (blue cross) zeros, for a Gaussian signal and $a = 0.7$. We can see a crown of zeros, which indicates the complete inversion of the channel.

In Figure 5 we can see the crown of zeros, which indicates that the obtained filter completely inverts the channel, providing magnitude and phase equalization.

Repeating the experiment for $a = 1.2$ (maximum phase channel) we have the following result:

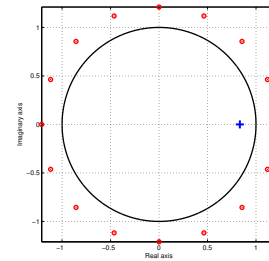


Fig. 6: Diagram of the channel (red circles) and filter (blue cross) zeros, for a Gaussian signal and $a = 1.2$. In this case, the filter cannot put its zero outside the unit circle, elucidating the minimum phase restriction.

As we can see in Figure 6, the filter cannot put its zero outside the unit circle. Therefore, deconvolution is not achieved in this case: the filter performed only a decorrelation between the signal samples, but not independence. This is equivalent to provide magnitude equalization only.

Next, we will repeat the same experiment for ℓ_1 and ℓ_∞ filters, again with a $K = 15$ and a first order filter $W(z) = 1 - w_1z^{-1}$.

For the input signal distribution we consider 3 cases:

- 1) Gaussian distribution, $N(0, 1)$;
- 2) Bernoulli-Gaussian, $Pr(s(n) = 1) = 0.02$ e $N(0, 1)$;
- 3) Uniform distribution, with support $[-1, 1]$.

In all cases, the filter coefficients were adjusted by differential evolution, with a population of 40 individuals in 2000 generations. We have also considered 2000 samples for the involved signals. The results for the ℓ_1 filter are depicted in Figures 7, 8 and 9:

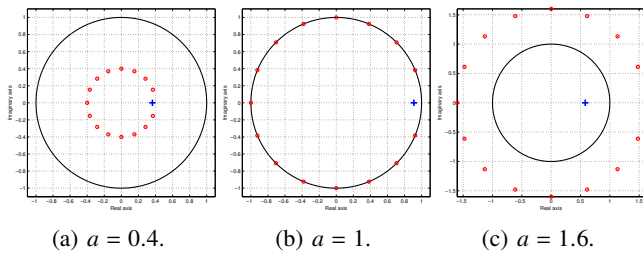


Fig. 7: Zero diagram for a ℓ_1 filter and a Gaussian signal. For this type of signal, the ℓ_1 PEF zero remained inside the unit circle.

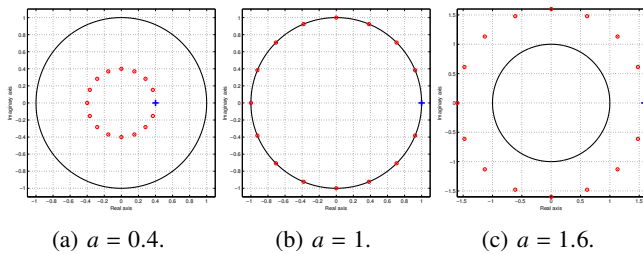


Fig. 8: Zero diagram for a ℓ_1 filter and a Bernoulli-Gaussian signal. For this type of signal (i.e. a sparse one), the filter put its zero outside the unit circle, showing its potential to perform nonminimum phase deconvolution.

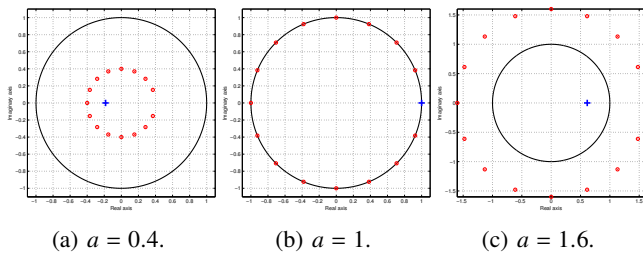


Fig. 9: Zero diagram for a ℓ_1 filter and a uniform signal. Another case where the filter zero remained inside the unit circle.

As we can see, the ℓ_1 filter places a zero outside the unit circle for the Bernoulli-Gaussian signal (8c), which has a sparse structure to be explored. For this kind of signal, deconvolution was well succeeded even in maximum phase cases and the filter provides both magnitude and phase equalization.

However, in both cases of Figures 7 (Gaussian distribution) and 9 (uniform distribution), where there the input signal is not sparse, the filter zero remains inside the unit circle, showing the same limitation of the ℓ_2 case.

Repeating the experiment (i.e., one coefficient for the filter, a $K = 15$ order channel, optimization carried out by differential evolution with 40 individuals, in 2000 generations and signals with 2000 samples) for ℓ_∞ filter, we have the results in Figures 10, 11 and 12:

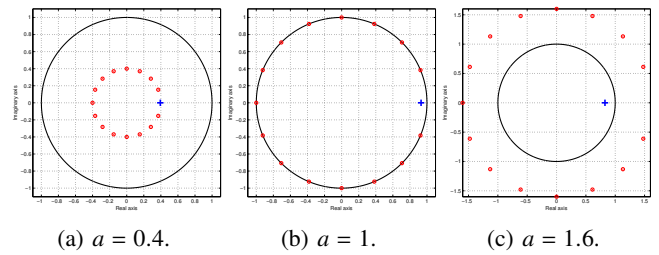


Fig. 10: Zero diagram for a ℓ_∞ filter and a Gaussian signal. As occurred with the ℓ_1 case, the filter cannot put its zero outside the unit circle.

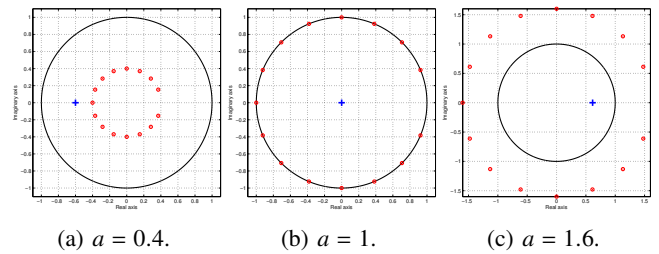


Fig. 11: Zero diagram for a ℓ_∞ filter and a Bernoulli-Gaussian signal. Again, the filter zero remained inside the unit circle, as a consequence of the signal distribution.

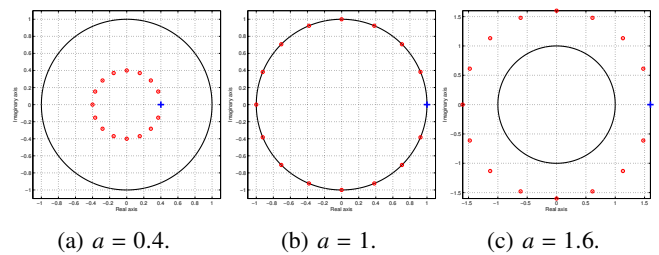


Fig. 12: Zero diagram for a ℓ_∞ filter and a uniform signal. When we use an antisparse signal, the ℓ_∞ PEF put its zero outside the unit circle.

From these results, we can see that the ℓ_∞ filter places a zero outside the unit circle (Figure 12c) and performs magnitude and phase deconvolution for maximum phase channels. The key feature for this is that the input signal has a uniform (antisparse) distribution, which is the suitable one for the ℓ_∞ norm. For the other two distributions, no zero was positioned outside the unit circle, as seen in Figures 10 and 11, and only magnitude equalization is provided, as in the ℓ_2 case.

Therefore, we can see that the ℓ_p PEFs have the potential to perform blind deconvolution in nonminimum phase channels, as long as the input signals have compatible structures (like sparsity or antisparse). Exploring such characteristics, the ℓ_p norm can go further the decorrelation condition and towards to independence, as desired.

To close our study, in the next section we show some performance results in the deconvolution of first and second order channels, using ℓ_1 and ℓ_∞ PEFs and comparing them with the classical ℓ_2 PEF.

VI. RESULTS ON THE PERFORMANCE OF THE ℓ_p PEF IN BLIND DECONVOLUTION

Now let us evaluate the performance of the ℓ_1 deconvolution for a sparse signal, and the ℓ_∞ deconvolution for an antisparse signal. Our results consider first and second order channels, of both minimum and maximum phases.

For the first order case, we have varied the channel (real) zero from 0.1 to 1.5. For the second order channel we have used complex-conjugated zeros in polar coordinates:

$$z_1 = re^{j\theta}, z_2 = re^{-j\theta}. \quad (26)$$

In this case, we have varied r in the interval $[0.1, 1.5]$ and θ in $[\frac{\pi}{2}, \pi]$.

In all cases, we used a $L = 10$ order filter, with coefficients adjusted by differential evolution; for the ℓ_1 filter we have used a population of 20 individuals in 2000 generations and for the ℓ_∞ filter we used 30 individuals in 4000 generations.

To decide whether deconvolution was well succeeded, we considered the inter symbol interference (ISI), in decibel scale (dB), of the global response (i.e. $g(n) = h(n) * w(n)$):

$$ISI_{dB}(g(n)) = 10 \log_{10} \frac{\sum_i |g_i^2| - \max_i |g_i|^2}{\max_i |g_i|^2}. \quad (27)$$

When $ISI_{dB}(g(n)) \leq -5$ dB, the channel was considered equalized. For a better visualization, we will show our results in a zero diagram in complex plane: the red circles in the diagram represent the channel zero (or complex-conjugated zeros) for which deconvolution was well succeeded (accordingly to (27)). For example, if the channel with a zero at 0.2 was equalized, we will have a red circle at 0.2 in the complex plane.

Figure 13 presents the results for sparse deconvolution, using ℓ_1 and ℓ_2 PEF:

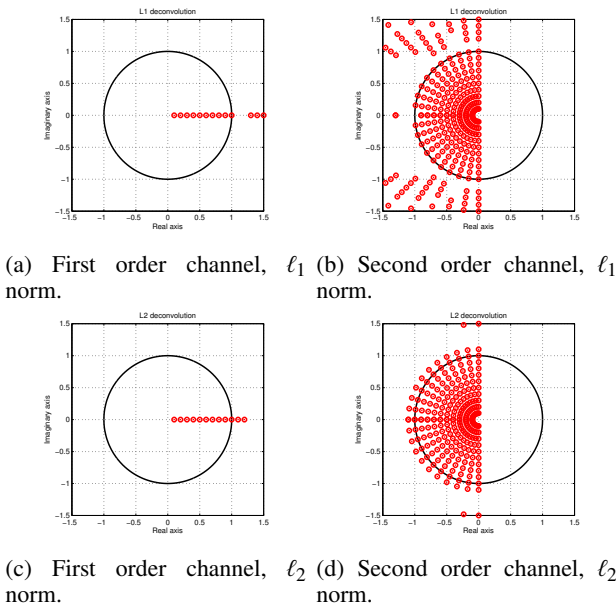


Fig. 13: Sparse deconvolution for first and second order channels, using ℓ_1 and ℓ_2 norm. The red circles represent the equalized channel zeros.

Figure 13 shows the zeros of the equalized channels, i.e., channels that achieve the threshold set for $ISI_{dB}(g(n))$, when using ℓ_1 and ℓ_2 PEF for the retrieval of a sparse signal. As Figures 13a and 13b show, ℓ_1 PEF performs deconvolution in maximum phase channels, but not in all of them (ideally, we must have all the left half of the complex plane marked, indicating a well succeeded deconvolution for all channels). This result shows both the potential and limitation of this criterion. On the other hand, the ℓ_2 PEF performs deconvolution essentially in minimum phase channels (although we can see some maximum phase channels near the unit circle), as expected due to the restriction of the ℓ_2 norm.

We obtained very similar results for the ℓ_∞ deconvolution, with antisparse signal, as shown in Figure 14:

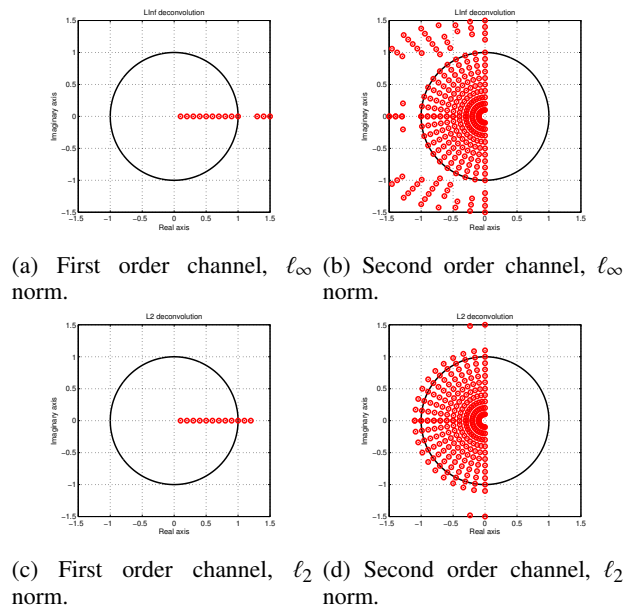


Fig. 14: Antisparse deconvolution for first and second order channels, using ℓ_∞ and ℓ_2 norm. The red circles represent the equalized channel zeros.

As illustrated by Figures 14a and 14b, the ℓ_∞ PEF also performs blind deconvolution in some nonminimum phase channels. For the ℓ_2 filter with an antisparse signal we have the same limitations as before. Comparing Figures 13a and 13b with 14a and 14b, respectively, we can see again the duality between the ℓ_1 and ℓ_∞ filters, since the channels equalized in both cases are very similar.

Another interesting result is that there is a gap near the unit circle (radius 1.1) between the channels equalized by ℓ_2 filters and those equalized by ℓ_1/ℓ_∞ ones: at the point that the ℓ_2 is no longer effective (i.e., from where the red circles are no longer present), we can apply the ℓ_1/ℓ_∞ , depending on signal structure, to carry out the deconvolution. This suggests that these criteria can be used together, in a cascaded hybrid-norm filter: the first filter could be adjusted by ℓ_2 norm, dealing with the minimum phase component of the system, and the second one by ℓ_1/ℓ_∞ norm, dealing with the maximum phase component left.

Finally, we analyze the $ISI(g(n))$ evolution while the filter's coefficients are adjusted. For this, we considered two particular channels: a first order one, with a zero at 1.5 and a second order one, with $r = 1.3$ and $\theta = 7\pi/10$. For each channel, we have used a ℓ_1 PEF for sparse deconvolution and a ℓ_∞ PEF for the antisparse case. We show these results in Figure 15:

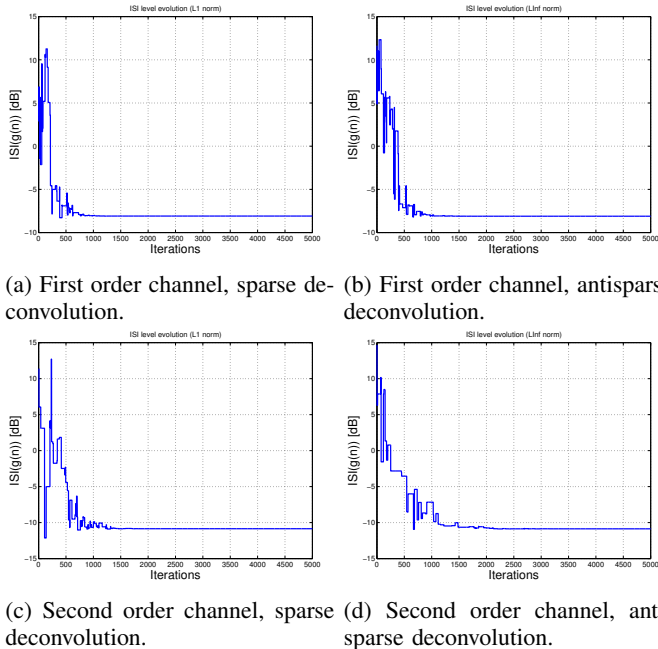


Fig. 15: $ISI(g(n))$ evolution during the filter adaptation, for first and second order channels, in sparse and antisparse deconvolution.

For the first order channel, we can see that the $ISI(g(n))$ level reaches its steady value in 1500 iterations, for both sparse and antisparse cases. In the second order channel, we have a constant $ISI(g(n))$ level in 1500 iterations for the sparse scenario, and in 2000 iterations for the antisparse one.

As we can see in Figure 15, the Differential Evolution approach has a fast converge rate, and once the steady value is reached, we do not observe the typical misadjustments of a gradient approach.

VII. CONCLUSION

This work discussed the application of ℓ_p norms in unsupervised deconvolution problem, using a predictive structure. We have shown that the ℓ_p criterion corresponds to the Maximum Likelihood estimator for generalized Gaussian distributions. Supported by this relation, we have applied a ℓ_1 PEF in order to retrieve a sparse signal and a ℓ_∞ PEF for an antisparse one. Our experiments illustrated that both filters could place some of its zeros outside the unit circle, providing nonminimum phase responses.

We also evaluated ℓ_1 and ℓ_∞ deconvolution, for first and second order channels, with minimum and maximum phase responses. We compared the results of these two experiments with the classical ℓ_2 deconvolution. For a suitable input signal, i.e., a sparse signal for the ℓ_1 case and an antisparse signal for

the ℓ_∞ , the proposed alternative criteria have shown a superior performance when compared to the classical ℓ_2 deconvolution, especially for nonminimum phase distorting systems.

Above all, the present work confirms the potential of ℓ_p predictors in overcoming the minimum phase restriction, which open interesting perspectives in different applications of unsupervised deconvolution. In this sense, we are particularly interested in testing the proposed approach in more realistic models for telecommunications channels and for geophysics. As far as theoretical investigations are concerned, future works will focus on the application of the ℓ_p norms alongside with alternative filter structures, as well as in the extension to other challenging problems, like Blind Source Separation and multichannel equalization.

APPENDIX

DESCRIPTION OF DIFFERENTIAL EVOLUTION HEURISTIC

In this appendix we describe the meta heuristic of differential evolution, which has been employed in the present work. The differential evolution meta heuristic uses a population of P individuals, formed by K -dimensional vectors, at each generation G . In our application, each individual represents a prediction error filter with $K + 1$ coefficients (i.e. a individual $\mathbf{w}_{i,G} = [w_{i1} \ w_{i2} \ \dots \ w_{iK}]$ is associated with the filter $\mathbf{w}(\mathbf{w}_{i,G}) = [1 \ -w_{i1} \ -w_{i2} \ \dots \ -w_{iK}]$).

First, we randomly initialize the population in a way to better explore the search space. In cases where one has some information about the solution, like promising regions or even partial solutions, it can be used in the initialization process; otherwise, the population must be uniformly initialized, as done here: for each individual in the population, we generate K independent and uniformly distributed random variables in the interval $[-1, 1]$, for each candidate set of filter coefficients.

Differential evolution generates new individuals by a mutation process. This processes is performed by the addition of weighted differences between two vector to a third one, as shown next:

$$\mathbf{t}_{i,G+1} = \mathbf{w}_{r_1,G} + F(\mathbf{w}_{r_3,G} - \mathbf{w}_{r_2,G}), \quad (28)$$

where $\mathbf{t}_{i,G+1}$ is the new generated vector; $\mathbf{w}_{r_l,G}$, with $l = 1, 2, 3$, are individuals present in the population at generation G ; and r_1, r_2, r_3 are mutually different indexes in the interval $[1, P]$ and also different from i . F is a real constant in the range $[0, 2]$, which determines the step towards the difference vector $\mathbf{w}_{r_3,G} - \mathbf{w}_{r_2,G}$.

Once the mutation is done, the mutated vectors are combined with the target vectors, generating the called trial vectors, accordingly with the following rule:

$$z_{ji,G+1} = \begin{cases} t_{ji,G+1}, & \text{if } r_j \leq CR \text{ or } j = l_i, \\ w_{ji,G}, & \text{if } r_j > CR \text{ and } j \neq l_i, \end{cases}$$

where $\mathbf{z}_{i,G+1} = [z_{1i} \ z_{2i} \ \dots \ z_{ji} \ \dots \ z_{Pi}]$ is a trial vector, with $j = 1, 2, \dots, D$; r_j is a random variable uniformly distributed in the interval $[0, 1]$; $CR \in [0, 1]$ its a real constant defined accordingly to the application and defines the crossover rate; and l_i is a randomly chosen index in the interval $[1, P]$,

which guarantees that $\mathbf{z}_{i,G+1}$ receives at least one component of $\mathbf{t}_{i,G+1}$.

Finally, we have the selection step. If the trial vector's fitness value is better than the fitness of the associated target value, the first takes the place of the latter in the next generation. Here, we have adopted the following fitness function:

$$fitness(\mathbf{w}_{i,G}) = \frac{1}{\|\mathbf{y}\|_p^p + 1}, \quad (29)$$

where $\mathbf{y} = [y(0) \ y(1) \ \dots \ y(N - 1)]$ is the signal generated at the output of the filter associated with $\mathbf{w}_{i,G}$.

For the ℓ_∞ norm case, we have used the *min max* estimator, so the fitness function is given by:

$$fitness(\mathbf{w}_{i,G}) = \frac{1}{\max(|\mathbf{y}|) + 1}. \quad (30)$$

All the process described above (mutation, crossover and selection) are repeated until the maximum number of generations is achieved. Once the process is concluded, we take the best individual (i.e. the one with the highest fitness value) as the solution to our problem.

REFERENCES

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

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

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

[4] K. Nose-Filho, "Adaptive sparse linear prediction: a promising tool for blind deconvolution," *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016, DOI:10.1109/SAM.2016.7569713.

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

[6] M. Najim, *Modeling, Estimation and Optimal Filtering in Signal Processing*. John Wiley and Sons, Inc., 2008.

[7] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Wiley, 2008.

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

[9] J.-M. Brossier, *Signal et communication numérique: égalisation et synchronisation*. Hermes, 1997.

[10] 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.

[11] C. Elvira, P. Chainais, and N. Dobigeon, "Bayesian antisparsity coding," *IEEE Transaction on Signal Processing*, vol. 65, pp. 1660–1672, 2015, DOI:https://doi.org/https://doi.org/10.1109/TSP.2016.2645543.

[12] E. A. Robinson, "Predictive decomposition of time series with applications to seismic exploration," Ph.D. dissertation, Massachusetts Institute of Technology, 1954.

[13] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and systems*. Prentice Hall, 1996.

[14] K. Nose-Filho, "Desconvolução e separação cega de sinais esparsos e aplicações em sísmica de reflexão," Ph.D. dissertation, Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, 2015.

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

[16] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 385–399, 1980, DOI:10.1109/TAC.1980.1102343.

[17] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Transactions on Information Theory*, vol. 36, pp. 312–321, 1990, DOI:10.1109/18.52478.

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

[19] J. R. Rice and J. S. White, "Norms for smoothing and estimation," *SIAM Review*, vol. 6, no. 3, pp. 243–256, 1964. [Online]. Available: <http://www.jstor.org/stable/2027590>

[20] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995, DOI:https://doi.org/10.1162/neco.1995.7.6.1129.

[21] P. Comon and C. Jutten, Eds., *Handbook of blind source separation*. Academic Press, 2010.

[22] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proceedings of IEEE*, vol. 86, pp. 2009–2025, 1998, DOI:https://doi.org/10.1109/5.720250.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[24] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, n.7, pp. 685–694, 2005, DOI:https://doi.org/10.1080/02664760500079464.

[25] J. Makhoul, "Linear prediction: a tutorial review," *Proceedings of the IEEE*, vol. 63, pp. 561–580, 1975, DOI:https://doi.org/10.1109/PROC.1975.9792.

[26] G. Cybenko, "Locations of zeros of predictor polynomials," *IEEE Transactions on Automatic Control*, vol. Ac-27, pp. 235–237, 1982, DOI:https://doi.org/10.1109/TAC.1982.1102870.

[27] J. M. T. Romano and M. Bellanger, "Zeros and poles of linear prediction digital filter," in *Proceedings of the 1986 EUSIPCO*, 1986.

[28] J. D. Scargle, "Absolute value optimization to estimate phase properties of stochastic time series," *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 140–143, 1977, DOI:https://doi.org/10.1109/TIT.1977.1055668.

[29] T. W. J. Bednar, R. Yarlagadda, " ℓ_1 deconvolution and its application to seismic signal processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 6, pp. 1655–1658, 1986, DOI:https://doi.org/10.1109/TASSP.1986.1165004.

[30] 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, DOI:https://doi.org/10.1109/18.149499.

[31] S. Haykin, *Adaptive filter theory*. Prentice Hall, 2001.

[32] J. A. L. K. Price, R. M. Storn, *Differential Evolution: A Practical Approach to Global Optimization*. Natural Computing Series, Springer, 2005.

[33] J. W. Brown and R. V. Churchill, *Complex variables and applications*. McGraw-Hill, 2009.



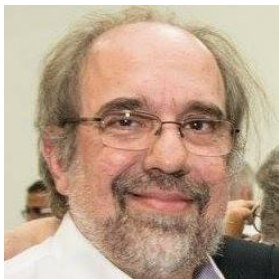
Renan Del Buono Brotto was born in Serra Negra, Brazil, in 1993. He obtained the B.S. degree in Computer Engineering from State University of Campinas, Brazil, in 2016. He is currently pursuing the Master degree at the same university in the Digital Signal Processing

for Communications Laboratory. His research interests are unsupervised signal processing, artificial intelligence and applications in communications and geophysics.



Kenji Nose-Filho received his B.S. (2008) and M.S. (2011) in Electrical Engineering at the São Paulo State University (UNESP). In 2013/2014 he made a six months stage at the GIPSA-Lab. In 2015 he received the Ph.D. degree from the University of Campinas (UNICAMP). From 2015 to 2017, he

made a post-doc at the University of Campinas (UNICAMP). His main research interest is in digital signal processing and machine learning, with applications in geophysics, time series forecasting, and image processing. He has been an assistant professor at the Engineering, Modeling and Applied Social Sciences Center at the Federal University of ABC since 2017.



João Marcos Travassos was born in Rio de Janeiro on April 12, 1960. He obtained the BSc and MSc degrees in Electrical Engineering at the University of Campinas (UNICAMP), in 1981 and 1984 respectively. He received his PhD from the University of Paris-XI in 1987.

Since 1988, he is at the School of Electrical and Computer Engineering (FEEC) of UNICAMP, where he is currently Professor. He was Visiting Professor at the CNAM/Paris and the Université René Descartes - Paris.

Prof. Romano has been Head of the Communications Department of FEEC/UNICAMP, Associate Coordinator of Graduate Program in Electrical Engineering, Associate Director of FEEC, and Director of Partnerships and Collaborative Projects of the Innovation Agency of the university (INOVA-UNICAMP). Currently, he is the director of the School of Electrical and Computer Engineering at UNICAMP.

Prof. Romano coordinates the research activities of the Signal Processing for Communications (DSPCom) Lab, where he works in the areas of adaptive filtering, unsupervised learning, blind source separation and signal processing applications in communications and geophysics. He has a CNPq research grant, level 1B.

Prof. Romano has been Vice-President (1996-2000) and President (2000-2004) of the Brazilian Telecommunications Society (SBrT), and is an Emeritus Member of SBrT. He has been the Editor-in-Chief of the SBrT/IEEE JCIS from 2011 to 2014 and he is currently a member of SBrT's Deliberative Council.