

Cascade of Linear Predictors for Deconvolution of Non-Stationary Channels in Sparse and Antispase Scenarios

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

Abstract—This work deals with adaptive predictive deconvolution of non-stationary channels. In particular, we investigate the use of a cascade of linear predictors in the recovering of sparse and antispase signals. To do so, we first discuss the behavior of the ℓ_p Prediction Error Filter (PEF), with $p \neq 2$, showing that it can better attenuate the effects of non-minimum phase channels in comparison with the classical ℓ_2 PEF, although the ℓ_p PEF, with $p \neq 2$, still presents intrinsic limitations in compensating the channel distortions, due to its direct forward structure. Hence, the cascade of linear predictors, i.e., one forward filter followed by a backward one, emerges as a possible solution to circumvent the structure limitation addressed. We apply the proposed cascade structure in the deconvolution of non-stationary channels, with minimum, maximum-, mixed- and variable-phase responses, and also in noisy scenarios. From the simulation results we observed that, besides the duality relation between the ℓ_1 and ℓ_∞ norms, they present different algorithmic behavior: a cascade adjusted by minimizing the ℓ_1 norm of the error attains a fast convergence, enhancing the cascade tracking capacity, but is more sensitive to noise. Adjusting the cascade by minimizing the ℓ_4 norm of the error (a smooth approximation of the ℓ_∞ norm), on the other hand, leads to a filter more robust to noise, but presents slower convergence and tracking capability.

Index Terms—Predictive deconvolution, ℓ_p norms, cascade of linear predictors, non-stationary channels.

I. INTRODUCTION

The problem of deconvolution consists in recovering a signal of interest $s(n)$ that has been distorted by

The authors would like to thank CAPES Code 001 and CNPq (305621/2015-7) for financial support.

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

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

Digital Object Identifier: 10.14209/jcis.2021.10

a channel, with impulse response $h(n)$, generating the distorted signal $x(n)$ [1], as depicted in Fig. 1a.

The objective in this problem, as illustrated in Fig. 1b, is to design a deconvolution filter, with impulse response $w(n)$, able to compensate the distortions introduced by the channel and to produce an estimate of the original signal $\hat{s}(n)$.

When we consider the unsupervised version of the problem, we must adjust $w(n)$ without access of the samples of $s(n)$ and without the explicit knowledge of the channel. In this case, the deconvolution is carried out in order to recover a property of $s(n)$ that is no longer present in $x(n)$. This property is referred to as a *prior* information of $s(n)$ and its recovery is equivalent to recovering the signal of interest.

A possible *prior* information to be employed is the independence among the samples of $s(n)$ [1], [2], [3], [4]. Since we are modeling the channel as a time-invariant linear system, we have

$$x(n) = h(n) * s(n) = \sum_{k=0}^{L-1} h(k)s(n-k), \quad (1)$$

where $*$ denotes the discrete-time convolution and $L > 1$ is the channel order. Due to the above convolution operation, it is evident that the independence property is no longer verified in the signal $x(n)$ at the output of the channel when it presents a non-flat magnitude.

Once we have access to the distorted signal, we must adjust the deconvolution filter in order to eliminate the statistical redundancy in $x(n)$. One possible way to do this is using the prediction error filter (PEF) [1], [2]. We illustrate this filter structure in Fig. 2.

Classically, PEF parameters are adjusted in order to minimize the mean squared error (MSE). The obtained filter presents two interesting properties [1], [5]:

- 1) It is a whitening filter, i.e., a sufficient long PEF provides an error signal $e(n)$ so that

$$\mathbb{E}[e(n)e(n-l)] = 0, \text{ for } l > 0, \quad (2)$$



Fig. 1: (a) In the deconvolution problem, the original signal $s(n)$ is distorted by a channel with impulse response $h(n)$, originating the signal $x(n)$. (b) Once the signal of interest has been distorted, we can adjust another filter, with impulse response $w(n)$, to compensate the channel distortions and produce an estimate $\hat{s}(n)$.

where the operator $\mathbb{E}[\cdot]$ is the statistical mean and we are considering a zero mean signal [1].

- 2) It is a minimum-phase filter, *i.e.*, all of its zeros lay inside the unit circle [1].

From the two properties of the classical PEF we can note the main limitations of this structure optimized by means of the MSE criterion. Because it is a whitening filter, it can only perform decorrelation in $e(n)$ samples, which can be understood as a “weak independence”. And from the PEF phase response, we can observe that this filter is as suitable structure only for the deconvolution of minimum-phase channels, which is a very restrictive condition [6]. Therefore we must seek a criterion and a suitable structure able to promote independence and generic phase response.

In [6] and [7] we have shown that the ℓ_p ¹ PEF, with $p \neq 2$, performs the nonlinear decorrelation, *i.e.*, the error signal produced has the property

$$\mathbb{E}\left[f(e(n))g(e(n-k))\right] = 0, \quad k \geq 1, \quad (3)$$

for all odd, nonlinear functions $f(\cdot)$ and $g(\cdot)$ [1].

We also investigated in [7] the phase-response of the ℓ_p PEF, with $p \neq 2$, verifying that it can present a non-minimum phase response. However, we also observed that the ℓ_p PEF was not able to compensate the distortions of a channel with a generic phase response. This limitation of the ℓ_p PEF, with $p \neq 2$, is supported by the work of Knockaert [8], which shows that all of the zeros of this type of filter lay inside a circle of radius 2. This result shows that the ℓ_p PEF, with $p \neq 2$, is more general than the classical one, since it provides non-linear decorrelation and presents a non-minimum-phase response, but still has a restriction due to its forward linear structure.

One possible way to circumvent this limitation is using a cascade of linear predictors: a forward linear predictor followed by a backward one, as we did in [9]. The use of a series of deconvolution filters has been addressed in previous works as [10], [11], [12] and [13]. Compared

to the mentioned works, our approach here is to employ a filter decomposition that preserves the filter linearity, leading to a simpler parameter adjustment, and that is also suitable for non-stationary channels.

The present paper further studies the potentialities of the cascade of linear predictors by applying it distinctly in two well-defined scenarios: for the blind deconvolution of antisparse signals, we propose to implement an MFE (Mean Fourth Error) cascade and for the blind deconvolution of sparse sources the MAE (Mean Absolute Error) cascade. We define a sparse signal as a signal that concentrates most of its information in a few samples (in this case, zeros or very small values do not carry significant information). In this work, we have used a Bernoulli-Gaussian distribution to model such signal. An antisparse signal, in a dual manner of the sparse one, is a signal that spreads the information equally among its samples and therefore is suitably modeled by an uniform distribution. An example of this kind of signal is a BPSK, which assumes the values +1 and -1 with equal probability, and can be considered an antisparse signal for our purposes in this work.

The MFE criterion is used as a smooth approximation of the ℓ_∞ norm, allowing the implementation of a stochastic descendent algorithm and the handling of non-stationary channels. The proposed structure is adapted by the LMF (Least Mean Fourth) algorithm, which can be directly derived from the MFE criterion. For the MAE criterion, the structure is adapted by the sign-error LMS algorithm. Simulation results show that this improved solution accomplishes efficiently the blind deconvolution task for non-stationary and non-minimum-phase channels.

The present work is organized as follows: in Section II we detail our approach and present suitable algorithms to adjust the parameters of the predictors; in Sections III-A and III-B we present the simulation results. Finally, in Section IV, we state our conclusions as well as some perspectives of future works.

II. PROPOSED STRUCTURE AND ALGORITHMS

Our approach here is to employ a cascade of linear predictors, as depicted in Fig. 3.

¹Here, we use ℓ_p for the p-norm of a vector.

In Fig. 3, $A(z)$ denotes the forward linear predictor, with error signal given by:

$$e_f(n) = x(n) - \sum_{k=1}^K a_k x(n-k), \quad (4)$$

where a_1, \dots, a_K are the parameters of the forward prediction coefficients.

In its turn, $B(z)$ denotes the backward predictor filter, which has an error signal given by:

$$e_b(n) = x(n-K) - \sum_{k=1}^K b_k x(n-k+1), \quad (5)$$

where b_1, \dots, b_K are the backward prediction coefficients.

Considering the structure in Fig. 3a and the equations (4) and (5), we can write the error signal $e(n)$ as

$$e(n) = y_1(n-K) - \sum_{k=1}^K b_k y_1(n-k+1), \quad (6)$$

where $y_1(n) = x(n) - \sum_{k=1}^K a_k x(n-k)$.

Due to the commutative property (Fig. 3b), we can also write $e(n)$ as follows:

$$e(n) = y_2(n) - \sum_{k=1}^K a_k y_2(n-k), \quad (7)$$

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

One of our interests in this work is to perform unsupervised deconvolution in order to recover telecommunication signals, which usually follow a uniform distribution [14], and, hence, present an antisparsity structure [15]. Therefore, as we discussed in [7], the most suitable measure for this property is the ℓ_∞ norm, which is equivalent to the Maximum Likelihood estimator for a uniform distribution. However, we will employ here the ℓ_4 norm as a smooth approximation for the ℓ_∞ one.

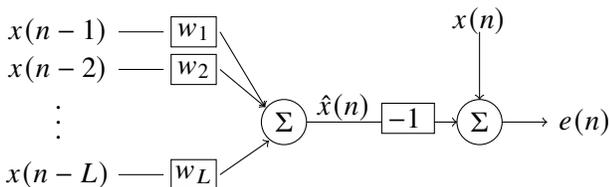


Fig. 2: Scheme of the predictor and the prediction error filters.

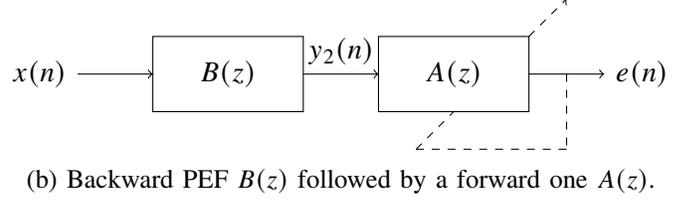
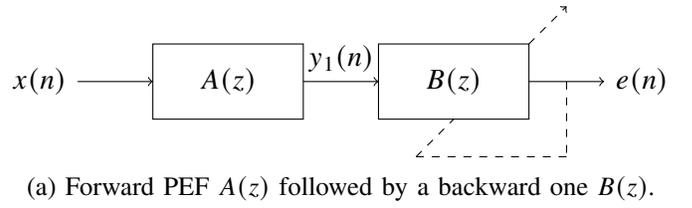


Fig. 3: The predictive cascade structures.

From (6), we can express the Mean Fourth Error (MFE) by:

$$E[|e^4(n)|] = E \left[\left| \left(y_1(n-K) - \sum_{k=1}^K b_k y_1(n-k+1) \right) \right|^4 \right], \quad (8)$$

or, from (7), by:

$$E[|e^4(n)|] = E \left[\left| \left(y_2(n) - \sum_{k=1}^K a_k y_2(n-k) \right) \right|^4 \right]. \quad (9)$$

Since we are using a smooth cost function, we can derive a gradient adaptive algorithm to optimize the cascade parameters. Hence, the update rule for the LMF [16] is simply given by:

$$\begin{aligned} a_i(n) &= a_i(n-1) + 4\mu(e^3(n-1))y_2(n-1-i), \\ b_i(n) &= b_i(n-1) + 4\mu(e^3(n-1))y_1(n-i), \end{aligned} \quad (10)$$

where μ is the learning rate. For simplicity, we considered real signals in the derivation of (10) but these adaptation rules can be easily extended for complex signals [9].

Our second objective here is with respect to sparse signals, as those typically found in seismic deconvolution [17]. For this case, we must minimize the Mean Absolute Error (MAE), which can be written as:

$$E[|e(n)|] = E \left[\left| \left(y_1(n-K) - \sum_{k=1}^K b_k y_1(n-k+1) \right) \right| \right], \quad (11)$$

or, from (7),

$$E[|e(n)|] = E \left[\left| \left(y_2(n) - \sum_{k=1}^K a_k y_2(n-k) \right) \right| \right]. \quad (12)$$

For the minimization of this cost function, we can adapt our structure by means of the sign-error LMS [16]:

$$\begin{aligned} a_i(n) &= a_i(n-1) + \mu \text{sign}(e(n-1))y_2(n-1-i), \\ b_i(n) &= b_i(n-1) + \mu \text{sign}(e(n-1))y_1(n-i). \end{aligned} \quad (13)$$

In the following sections, we will present our simulations results of unsupervised deconvolution of non-stationary channels with a cascade of linear predictors, considering the recovery of antisparse (Section III-A) and sparse signals (Section III-B).

III. SIMULATION RESULTS

A. Antisparse Deconvolution

In this section, we consider a BPSK input signal [14], the samples of which assume the values ± 1 . We consider a second order non-stationary channel, the real-valued zeros of which evolve with the following update rule:

$$\begin{aligned} z_1(n) &= z_1(0) + n\lambda_1, \\ z_2(n) &= z_2(0) + n\lambda_2, \end{aligned} \quad (14)$$

with

$$\begin{aligned} \lambda_1 &= \frac{(z_1(T-1) - z_1(0))}{T}, \\ \lambda_2 &= \frac{(z_2(T-1) - z_2(0))}{T}, \end{aligned} \quad (15)$$

where T denotes the total of available samples. In our first experiment, we considered:

$$\begin{aligned} z_1(0) &= 0.1 \rightarrow z_1(T-1) = 0.5 \\ z_2(0) &= 0.5 \rightarrow z_2(T-1) = 0.9 \end{aligned}$$

To provide a stability condition for the algorithm, we considered a stationary channel for the 1.10^4 first and last samples, *i.e.*, we have a channel with zeros at $z_1 = 0.1$ and $z_2 = 0.5$ for the initial samples and a channel with zeros at $z_1 = 0.5$ and $z_2 = 0.9$ for the final ones. For the 5.10^4 intermediary samples, we have a non-stationary channel, with its zeros updated as (14) and (15). The objective of this initialization is to show that once the proposed method has converged to a solution of a stationary channel, it can track the solution to a non-stationary one.

With respect to the cascade, we used 5 taps for the forward predictor and 5 taps for the backward one, both numbers determined by preliminaries studies. It is important to highlight that our choice for the length of the forward and backward filters assumes a symmetry for the channel, *i.e.*, it presents approximately the same number of zeros inside and outside the unit circle. We adjusted the structure by means of the LMF algorithm.

To measure the cascade performance in the equalization task, we have adopted the Intersymbol Interference Rate (ISI) of the combined response. We define the ISI level, in dB scale, as follows:

$$\text{ISI}_{\text{dB}} = 10 \log_{10} \frac{\sum_i |g_i|^2 - \max_i |g_i|^2}{\max_i |g_i|^2}, \quad (16)$$

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

Ideally, the combined response would be a delayed impulse, indicating the perfect channel inversion. Plugging $g(n) = \delta(n-d)$ into (16) we obtain an ISI level of $-\infty$. Therefore, the lesser the ISI obtained, the better is the cascade performance in the equalization.

We show the ISI level evolution for the combined response of our first experiment in Fig. 4a, considering the mean performance over 200 Monte Carlo Simulations. We considered four values for the learning rate, *i.e.*, $\mu \in \{1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}\}$.

For comparison purposes, we have also obtained the Wiener solution for the channel, which is the best inverse filter of the channel obtained through the minimization of the MSE criterion by having access to the channel coefficients.

For each delay k in the reference signal, we compute

$$\mathbf{w}_k = (\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}\delta(n-k), \quad (17)$$

where \mathbf{H} is the convolution matrix and δ is a vector with only the k^{th} entry unitary and all the other are zero. Therefore, we choose the filter that provides the minimal MSE:

$$\|w_k(n) * h(n) - \delta(n-k)\|_2 \quad (18)$$

We used the Wiener solution as a baseline, that represents the best result (in terms of the MSE criterion, with a supervised algorithm) that we could obtain in terms of Zero Forcing solution. In a noiseless case, the closer the performance of a filter to the Wiener solution, the better.

In this work, we have considered a 6-tap filter and a 11-tap one, and, after each new update of the channel zeros, we obtained the channel convolution matrix \mathbf{H} and solved (17) for the best delay in the reference signal. These two Wiener solutions give us the best performance that can be achieved using 6 taps, as if only one of the cascade filters were employed, and the best performance for 11-tap filters, in which case both parts of the cascade are in use.

From Fig. 4a we can see that as we increment the learning rate, the cascade tracking capacity gets better, as we can observe from the proximity of the continuous

lines (which represent the combined response for different values of μ) and the dashed ones (representing the Wiener solutions), especially in the initial samples. As the cascade parameters converge ($n > 3 \cdot 10^4$), the ISI lines associated with different μ values become closer, except for $\mu = 1 \cdot 10^{-4}$.

For the second experiment, we considered a maximum-phase channel, with its zeros specified by:

$$\begin{aligned} z_1(0) &= 10 \rightarrow z_1(T-1) = 2, \\ z_2(0) &= 2 \rightarrow z_2(T-1) = 1.1. \end{aligned}$$

This choice of zeros gives us an equivalency between the minimum- and maximum-phase channels in terms of amplitude distortion [18].

Again we considered a 5-tap forward predictor and a 5-tap backward predictor for the cascade, with the same learning rates as before, and we show the results for the maximum-phase scenario in Fig. 4b.

As we observed for the minimum-phase channel, the tracking capacity is very close related with the learning rate: a bigger learning rate leads to a cascade with a faster convergence rate. It is also interesting to note the similarities from Fig.s 4a and 4b, due to the relationship between the zeros of the channels considered.

In our third experiment with non-stationary channels, we considered a mixed-phase one, with its zeros linearly evolving as

$$\begin{aligned} z_1(0) &= 0.2 \rightarrow z_1(T-1) = 0.8, \\ z_2(0) &= 3.0 \rightarrow z_2(T-1) = 3.6. \end{aligned}$$

We kept the other parameters, *i.e.*, number of cascade coefficients, learning rates and number Monte Carlo simulations, as in the previous simulations and we present our results in Fig. 4c.

For the mixed-phase channel, we verify that for the initial samples ($n < 1 \cdot 10^4$) the best performance was attained by the cascade adjusted with $\mu = 1 \cdot 10^{-3}$; for the next samples (up to $n = 4 \cdot 10^4$) the best performance was achieved with the biggest learning rate; finally, in the final samples ($n > 4 \cdot 10^4$), the cascade adjusted with $\mu = 5 \cdot 10^{-4}$, $1 \cdot 10^{-3}$, $5 \cdot 10^{-3}$ had similar performances.

We can also see that the ISI lines associated with the cascade are distant from the one associated with the 10-tap Wiener solution, represented by the black dashed line. This result suggests that the cascade needs a slower transition for the channel zeros when it has a mixed-phase channel than when the channel has minimum- and maximum-phase responses. Conversely, for this kind of channel we need a faster adaptation rule for the cascade.

Also, we considered a variable-phase channel, *i.e.* a channel that starts with a maximum-phase response, passes to a mixed-phase and, finally, turns into a

minimum-phase one, with its two zeros varying in the intervals specified below

$$\begin{aligned} z_1(0) &= 3.0 \rightarrow z_1(T-1) = 0.1, \\ z_2(0) &= 5.0 \rightarrow z_2(T-1) = 0.7. \end{aligned}$$

We repeated the same parameters as in the first three experiments and we present our simulation results in Fig. 4d.

For this fourth channel, we observe that for $1 \cdot 10^4 < n < 3 \cdot 10^4$ the cascades adjusted with the biggest values of μ presented better performances, relating again the learning rate with the tracking capacity. In this particular channel we can observe some ISI peaks, due to the proximity of the channel zeros to the unit circle. When we adjusted the cascade with the smallest learning rate ($\mu = 1 \cdot 10^{-4}$), the proposed filter was able to track the first observed peak; for the other μ values considered, we can verify an ISI peak with a delay. For the second peak, we observe in Fig. 4d that all the μ values considered lead to cascades able to track this peak, again with some delay. For the last samples, $n > 4 \cdot 10^4$, we verify again the effect of the learning rate: the bigger the value of this parameter, the better was the performance of the associated cascade.

Finally, we evaluated the cascade performance under a noisy scenario. Here we considered the mixed-phase channel, with an additive Gaussian noise. We present our simulation results in Fig. 5a, for SNR = 20 dB, and Fig. 5b, for SNR = 10 dB. When we are dealing with a low level of noise (SNR \geq 20 dB), the ℓ_4 cascade performance is little affected, and the tracking performance is almost the same of the noiseless case. On the other hand, as the noise level increases, the cascade performance gets poorer and its tracking capacity gets lower, as we can observe from Fig. 5b, for all the learning rates considered.

B. Sparse Deconvolution

Now we present our simulation results with sparse signals and must adopt an ℓ_p norm with $1 \leq p < 2$. For the sake of the simplicity of algorithm, we use the ℓ_1 norm, which gives raise to the following adaptation rules

$$\begin{aligned} a_i(n) &= a_i(n-1) + \mu \text{sign}(e(n-1))y_2(n-1-i), \\ b_i(n) &= b_i(n-1) + \mu \text{sign}(e(n-1))y_1(n-i). \end{aligned} \quad (19)$$

We modeled the sparse signal by means of the Bernoulli-Gaussian distribution: first we generate a Bernoulli sequence [19], with a probability p_1 of getting the value 1; then we multiply the sequence obtained

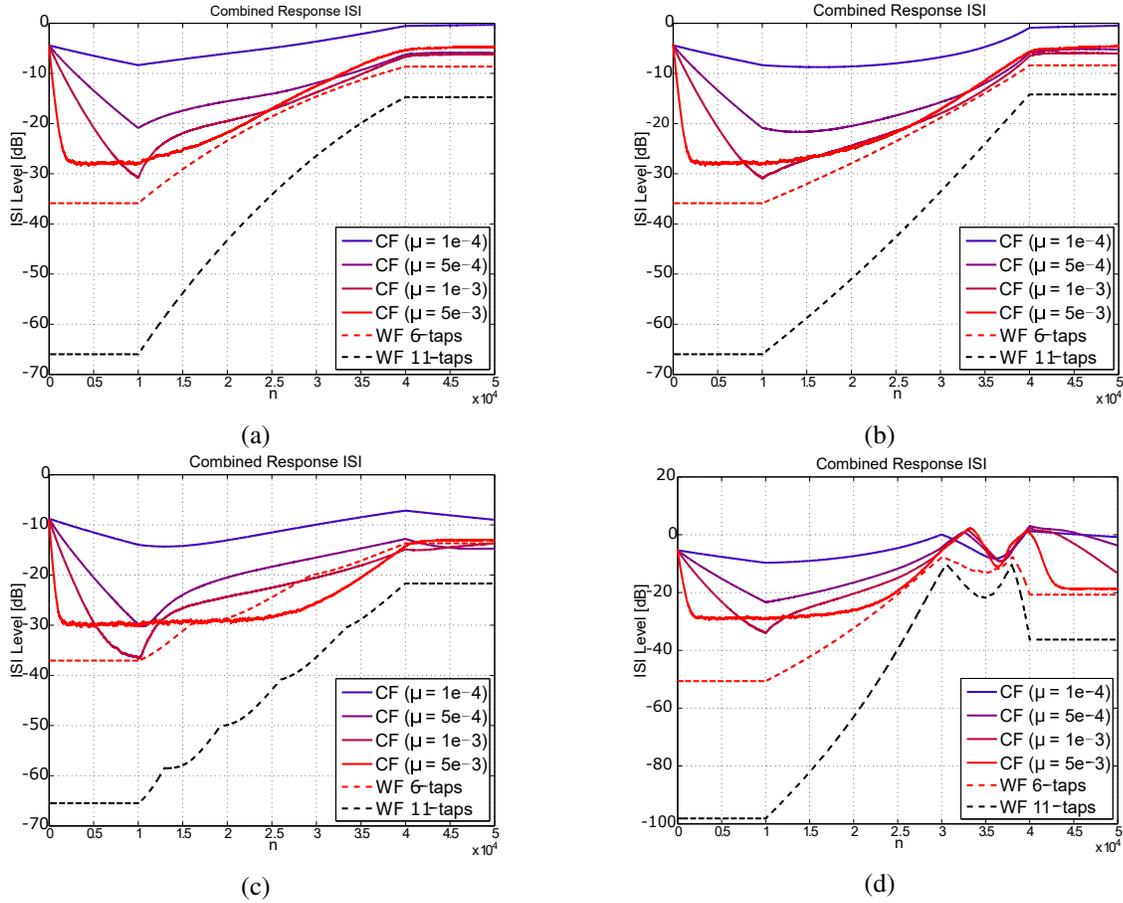


Fig. 4: Deconvolution of antispase signals. Combined response ISI of the cascade filter (CF) and Wiener Filter (WF), in dB scale, considering 200 Monte Carlo simulations for: (a) the minimum phase system, (b) the maximum phase system, (c) the mixed phase-system and (d) the variable phase system.

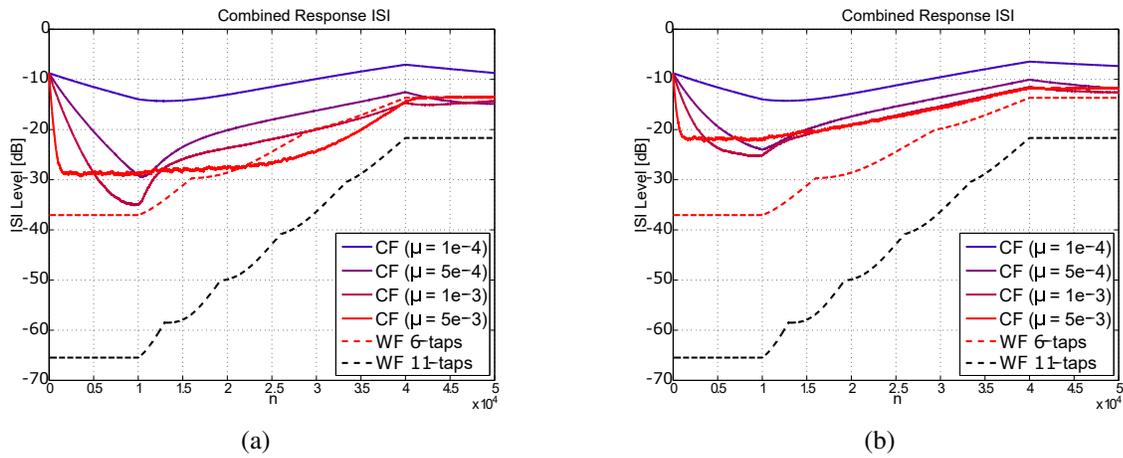


Fig. 5: Antispase deconvolution in a noisy scenario. We considered the same mixed-phase channel, 200 Monte Carlo simulations and an SNR of (a) 20 db and (b) 10 dB.

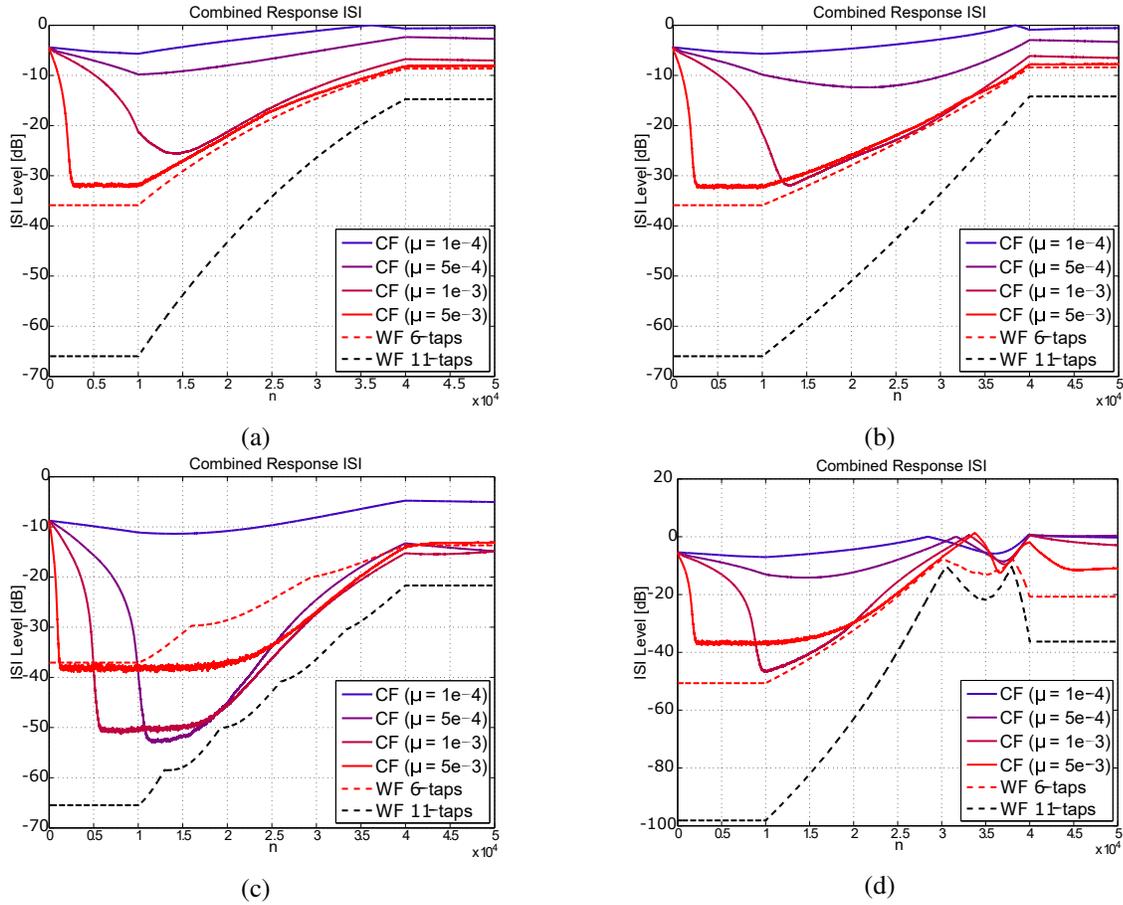


Fig. 6: Deconvolution of sparse signals. Combined response ISI of the cascade filter (CF) and Wiener Filter (WF), in dB scale, considering 200 Monte Carlo simulations for: (a) the minimum phase system, (b) the maximum phase system, (c) the mixed phase-system and (d) the variable phase system.

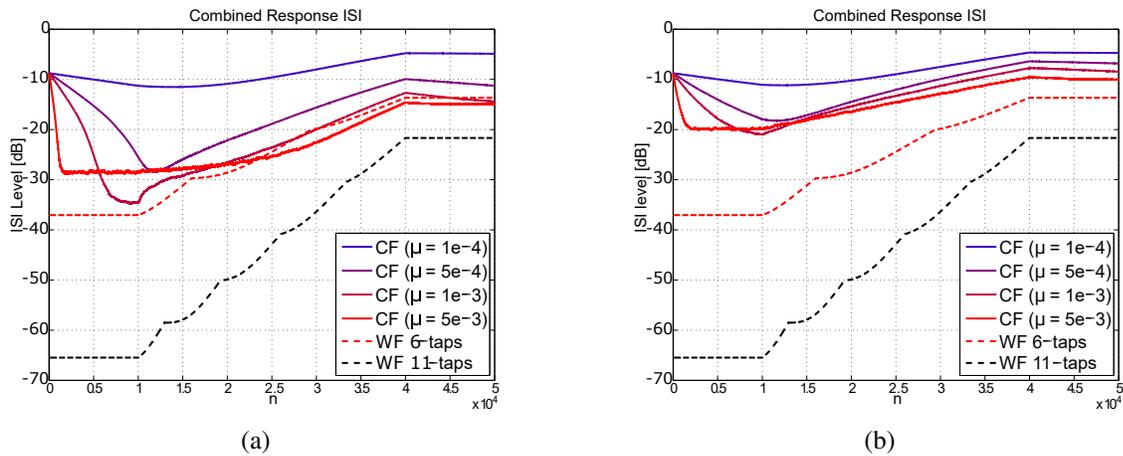


Fig. 7: Sparse deconvolution in a noisy scenario. We considered the same mixed-phase channel, 200 Monte Carlo simulations and an SNR of (a) 20 db and (b) 10 dB.

by another random sequence, with values drawn from a normal distribution, with zero mean and unit variance. The parameter p_1 allows us to control the degree of sparsity of the signal. In our simulations, we used $p_1 = 0.1$.

As we did for the antisparse case, we used 5 taps for the forward component and 5 taps for the backward one, both structures adjusted by (19). We considered the same channels and learning rates and measured the performance by means of the ISI level.

We present our simulation results of deconvolution with the minimum-phase channel ($z_1(0) = 0.1$, $z_2(0) = 0.5$, $z_1(T-1) = 0.5$, and $z_2(T-1) = 0.9$) in Fig. 6a.

In this first experiment, the result is very close to the one obtained for the antisparse signal: the larger the values for the learning rate, *i.e.*, $\mu = 1 \times 10^{-3}$ and $\mu = 5 \times 10^{-3}$, the better the results, since the ISI levels obtained with this parameters are close to the 5-tap Wiener solution. For small values of μ , the tracking capacity of the cascade is reduced and, for $\mu = 1 \times 10^{-4}$ and $\mu = 5 \times 10^{-4}$, the performance of the sparse deconvolution is worse than the performance of the antisparse deconvolution with the same parameters.

Considering a maximum-phase channel ($z_1(0) = 10$, $z_2(0) = 2$, $z_1(T-1) = 2$, and $z_2(T-1) = 1.1$), we present our deconvolution results in Fig. 6b. For this second experiment with sparse signals, the performance is also very similar to the one of the antisparse case. Again, we got the best performances, in the ISI level sense, with the larger values of the learning rate.

For a mixed-phase channel ($z_1(0) = 0.2$, $z_2(0) = 3$, $z_1(T-1) = 0.8$, and $z_2(T-1) = 3.6$) sparse deconvolution, we present our simulation results in Fig. 6c. For this channel, the best results were achieved when we adopted $\mu = 5 \cdot 10^{-4}$ and $\mu = 1 \cdot 10^{-3}$. Comparing the results in Fig. 6c with the ones in Fig. 4c, we observe that for the sparse case, with the adaptation rule given by (10), the ℓ_1 cascade tracked better the channel than the ℓ_4 cascade.

Finally, we considered again a variable phase channel ($z_1(0) = 3$, $z_2(0) = 5$, $z_1(T-1) = 0.1$, and $z_2(T-1) = 0.7$) and present our simulation results in Fig. 6d.

For this case, we see that the performance of the ℓ_1 cascade is very similar to the performance of the ℓ_4 cascade. For this scenario we observe that the ℓ_1 cascade can also replicate the observed ISI peaks, but with a delay. As we verified with the ℓ_4 cascade, when we have the channel zeros close to the unit circle, we need to speed up the cascade adaptation.

As we did for the antisparse deconvolution, we also considered here a noisy scenario, as depicted in Fig. 7. The cascade adaptation based on the ℓ_1 norm was more

susceptible to the additive noise, even for an SNR = 20 dB, and the performance is very different than the one obtained in the noiseless scenario. In addition, as we increase the noise level, the tracking ability is reduced.

This result reinforces our previous comment about the algorithmic differences of the cascade adaptation based on the ℓ_1 and ℓ_4 norms: using the ℓ_1 norm, the cascade parameters present a fast convergence rate, which improves the filter tracking capacity, but is more susceptible to noise. On the other hand, using the ℓ_4 norm to adjust the parameters, leads to a cascade with a slower convergence and a lower tracking capacity, but is more robust with respect to noise.

IV. CONCLUSION

In this work we extended our previous results in blind deconvolution using a cascade of linear predictors. We considered here the deconvolution of antisparse signals, by means of the ℓ_4 cascade, and sparse signals, using the ℓ_1 cascade, both for non-stationary channels.

We observed that for minimum- and maximum-phase channels, both structures were able to track the channel evolution. For the mixed-phase one, the ℓ_4 cascade presented an inferior performance than the ℓ_1 cascade, but in the presence of noise we observe the opposite. This particular result shows that besides the norms duality, different optimization criteria lead to different algorithmic behavior: using the ℓ_1 norm favors a fast convergence, increasing the tracking capacity, but is more susceptible to noise; using the ℓ_4 norm yields a slower convergence and less tracking capacity, but is more robust with respect to noise.

Finally, for a variable phase channel, both structures were not able to track the channel behavior, mostly when the channel zeros were close to the unit circle. This result makes patent the necessity to investigate methods to speed up the cascade adjustment.

REFERENCES

- [1] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, *Unsupervised signal processing: channel equalization and source separation*. CRC Press, 2011.
- [2] E. A. Robinson, *Predictive decomposition of time series with applications to seismic exploration*. PhD thesis, Massachusetts Institute of Technology, 1954.
- [3] 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.
- [4] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2001.
- [5] P. P. Vaidyanathan, *The theory of linear prediction*. Morgan & Claypool Publishers, 2008.

[6] K. Nose-Filho and J. M. T. Romano, "Adaptive sparse linear prediction: A promising tool for blind deconvolution," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5, 2016. DOI:10.1109/SAM.2016.7569713.

[7] R. Brotto, K. Nose-Filho, and J. M. T. Romano, "Alternative criteria for predictive blind deconvolution," *JCIS*, vol. 33, no. 1, pp. 271–281, 2018. DOI:10.14209/jcis.2018.27.

[8] 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.

[9] R. Brotto, K. Nose-Filho, R. Attux, and J. M. T. Romano, "Adaptive linear predictors in cascade for blind deconvolution of non-stationary and non-minimum-phase channels," *Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBRT 2019. Petrópolis - RJ.*, 2019. DOI:10.14209/SBRT.2019.1570558960.

[10] 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.

[11] Liqing Zhang, A. Cichocki, and S. Amari, "Multichannel blind deconvolution of nonminimum-phase systems using filter decomposition," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1430–1442, 2004. DOI:10.1109/TSP.2004.826185.

[12] Y. Z. A. Guitton and J. Claerbout, "A new bidirectional deconvolution method that overcomes the minimum phase assumption," tech. rep., SEP-142, 2011. DOI:https://doi.org/10.3997/2214-4609.20149156.

[13] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: can you skip the training period?," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 921–930, 1998. DOI:10.1109/26.701319.

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

[15] C. Elvira, P. Chainais, and N. Dobigeon, "Bayesian antisparse coding," *IEEE Transaction on Signal Processing*, vol. 65, pp. 1660–1672, 2017. DOI:https://doi.org/https://doi.org/10.1109/TSP.2016.2645543.

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

[17] K. Nose-Filho, A. K. Takahata, R. Lopes, and J. M. T. Romano, "Improving sparse multichannel blind deconvolution with correlated seismic data: Foundations and further results," *IEEE Signal Processing Magazine*, vol. 35, no. 2, pp. 41–50, 2018. DOI:10.1109/MSP.2017.2783379.

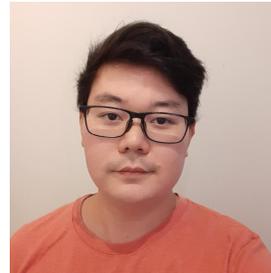
[18] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Pearson, 3rd ed., 2009.

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



Renan Del Buono Brotto was born in Serra Negra, Brazil, in 1993. He obtained the B.S. degree in Computer Engineering in 2016, and his M.Sc degree in 2019, in the area of Unsupervised Signal Processing, both from the University of Campinas, Brazil. He is currently pursuing the

PhD degree at the same university in the Digital Signal Processing for Communications Laboratory. His research interests are Unsupervised Signal Processing, Blind Source Separation, Unsupervised Deconvolution and Fairness in Machine Learning.



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 research interests in signal processing are adaptive filtering, predictive deconvolution, blind source separation, machine learning, and optimization methods, with applications in reflection seismic, time-series analysis and forecasting, and image processing. He is currently an assistant professor at the Federal University of ABC (UFABC), member of the Scientific Initiation Program of the UFABC, vice coordinator of the Undergraduate Information Engineering Course and chair of the IEEE Signal Processing South Brazil Section Chapter.



Romis Attux was born in Goiânia, in 1978. He received the titles of Electrical Engineer (1999), Master in Electrical Engineering (2001) and Doctor in Electrical Engineering (2005) from the University of Campinas. Presently, he is an associate professor at the same university. His main research interests are adaptive filtering, machine learning, brain-computer interfaces, dynamical systems / chaos and ethical aspects of AI.

João M. T. Romano received the B.S. and M.S. degree in Electrical Engineering from the University of Campinas (Unicamp), Brazil, in 1981 and 1984 respectively. In 1987 he received the Ph.D. degree from the University of Paris-XI. In 1988 he joined the Communications

Department of the School of Electrical and Computer Engineering (FEEC), Unicamp, where he is currently Professor. He is recipient of the “Zeferino Vaz Academic Achievements Prize” (Unicamp, 2008), the “25 Years of the Brazilian Telecommunication Society Prize” (SBrT, 2008) for research and teaching contribution in the field and the “Unicamp Award for University Extension Activities” (2020).



He is the head of the Signal Processing for Communications Laboratory and his research interests concern signal processing and machine learning with emphasis in blind deconvolution, source separation and adaptive filtering. Since

1988, he is a recipient of a Research Fellowship from CNPq-Brazil. He was the President of the Brazilian Communications Society (SBrT) from 2000 to 2004. He is a Senior Member of IEEE and an Emeritus Member of SBrT.

Prof. Romano was the Director of the School of Electrical and Computer Engineering from 2015 to 2019. He is now the Vice-President for Research of the University of Campinas