

BLIND COMPENSATION OF NONLINEAR DISTORTIONS VIA SPARSITY RECOVERY

Leonardo T. Duarte¹, Ricardo Suyama², Romis Attux¹, João Marcos T. Romano¹, Christian Jutten^{3*}

1-DSPCom Lab, University of Campinas (UNICAMP), Brazil

2-Engineering, Modeling and Applied Social Sciences Center (CECS), UFABC, Santo André, Brazil

3-GIPSA-lab, UMR CNRS 5216, Université Joseph Fourier, France

leonardo.duarte@fca.unicamp.br, ricardo.suyama@ufabc.edu.br, attux@dca.fee.unicamp.br

romano@dmo.fee.unicamp.br, christian.jutten@gipsa-lab.grenoble-inp.fr

ABSTRACT

In this work, we address the problem of compensating a nonlinear memoryless system in a blind fashion, i.e., without considering a set of training points. Our proposal works with the assumption that the input signal admits a sparse representation in a transformed domain that should be known in advance. By assuming that the nonlinear distortion function makes the observed signal less sparse (this is observed in frequency transforms), the proposed method aims at estimating the original signal via a sparsity recovery procedure. Our approach is based on an approximation of the ℓ_0 -norm and on the use of polynomial functions as compensating structures. In order to assess the viability of the developed method, we perform a representative set of experiments on synthetic data.

Index Terms— Blind compensation, nonlinear distortion, sparse signals.

1. INTRODUCTION

Nonlinear distortions arise in many applications. In some cases, this problem stems from the fact that the sensors used to acquire the desired signal are based on nonlinear transducer mechanisms. For instance, in electrochemical sensors, the conversion of chemical energy into electrical energy is clearly a nonlinear process [1]. Another common source of nonlinear distortion can be found in the amplifying stages that follow the signal acquisition. This situation often arises in audio signal processing and satellite communications [2].

When a set of training points is available, it is possible to counterbalance a nonlinear distortion either by adjusting a compensating function in a supervised fashion [3] or by carrying out an identification of the nonlinear function [4]. Conversely, when it is impossible or demanding to obtain training points, one must consider a blind (or unsupervised) framework. The resulting problem in this case is challenging, and can be solved only if additional prior information is taken into account.

In order to develop a blind framework for dealing with nonlinear distortions, we introduce in this work a novel approach which is based on the assumption that the input signal admits a sparse representation [5]. The cornerstone of our idea is the assumption that, as a consequence of the nonlinear distortion, the observed signal is usually less sparse than the original one. This is observed, for instance, in frequency transforms. Thus, we propose to adjust a compensating function so that the estimated input signal be as sparse as possible. To implement our proposal, we consider an approximation of the ℓ_0 -norm¹ as a measure of sparsity, and employ compensating devices parametrized by polynomial functions.

In a certain sense, this proposal can be regarded as a generalization of the idea firstly introduced by Landau and Miranker [6] and recently adapted to unsupervised scenarios [7, 8]. Indeed, the main assumption considered in these works is that the input signal is bandlimited. Since a nonlinear function tends to spread the spectrum of the observed signal, the basic idea in this case is to restore a bandlimited signal. Note that assuming the signal to be bandlimited is somewhat equivalent to considering that the desired signal presents a special pattern of sparsity — in fact, such signals can be classified as block-sparse. In our proposal, though, the input signal does not necessarily have to present a block-sparse character.

The paper is organized as follows. In Section 2, we introduce the problem of blind compensation of nonlinear functions. Then, in Section 3, we show how nonlinear distortions can be mitigated via a sparsity minimization procedure. In Section 4, we provide some numerical experiments to attest the validity of our proposal. Finally, our conclusions are presented in Section 5.

2. PROBLEM STATEMENT

Let the vector $\mathbf{s} \in \mathbb{R}^N$ represent an one-dimensional discrete, where N is the sample size. The element-wise function $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ models the effects introduced on \mathbf{s} by a nonlinear

¹Strictly speaking, the ℓ_0 -norm is not a mathematical norm [5]. However, we keep this nomenclature since the term “ ℓ_0 -norm” is commonly used.

*C. Jutten is also with Institut Universitaire de France.

memoryless system, being the i -th element of the observed signal, $\mathbf{x} \in \mathbb{R}^N$, given by

$$x_i = f(s_i). \quad (1)$$

In our notation, the overall mapping is denoted by $\mathbf{x} = f(\mathbf{s})$. The problem addressed in this work is to estimate \mathbf{s} in an unsupervised fashion, that is, based only on the observed signal \mathbf{x} . As is the rule in unsupervised methods, scaling ambiguities are accepted, so scaled versions of \mathbf{s} are also considered perfect solutions to the problem.

In order to obtain estimates of the input signal, we consider a compensating function $g : \mathbb{R}^N \rightarrow \mathbb{R}^N$ that should be adjusted so $\mathbf{y} = g(\mathbf{x})$ be as close as possible to \mathbf{s} . We assume that both f and g are monotonic functions with non-vanishing derivative. Otherwise, the problem becomes very difficult to solve, requiring strong prior information.

3. BLIND COMPENSATION OF NONLINEAR DISTORTIONS VIA SPARSITY RECOVERY

3.1. The basic idea

In order to introduce the motivations of our proposal, let us present an example. Consider the discrete signal \mathbf{s} , illustrated in the left side of Figure 1. This signal can be sparsely represented in a transformed domain, being the forward transform represented by the orthogonal matrix Φ , so that $\mathbf{s}_c = \Phi \mathbf{s}$ corresponds to the transform coefficients. In our example, we are considering the discrete cosine transform (DCT).

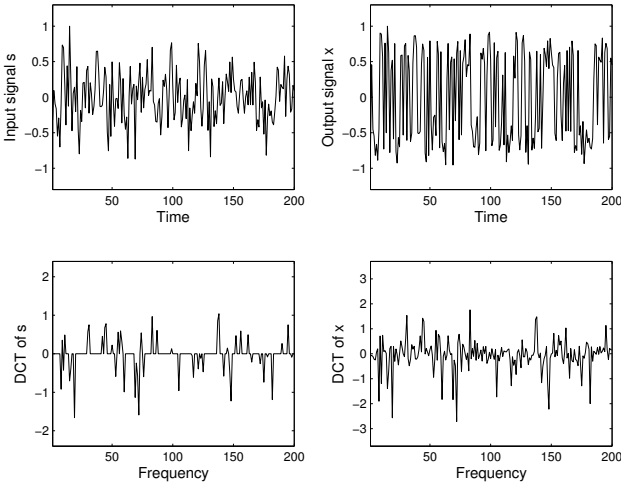


Fig. 1. Effects of a cubic root function on the DCT of the observations. The input and output signals, and their DCTs, are shown in the left and right sides of the figure, respectively.

After imposing the input signal to a cubic root function, i.e., $x_i = f(s_i)$, we obtained the signal depicted in the right side of Figure 1. The key point here is that the DCT of \mathbf{x} is clearly less sparse than the one of the input signal. This is

analogous to the spectral spreading phenomenon that arises when bandlimited signals are submitted to nonlinear functions [6, 7]. Therefore, a natural criterion to adjust the compensating function g is to recovery a sparse signal. We discuss in the sequel how this can be carried out in practice.

3.2. Implementation aspects

A first point that we must deal with to implement our idea concerns the compensating nonlinear function g . As mentioned before, this function must be monotonic. Otherwise, there is a risk of having a compensating function able to generate null elements in the transformed domain.

Another important issue here is to choose a compensating structure that, without violating the constraint of strict monotonicity, be as flexible as possible. In this spirit, we consider polynomial functions, so the estimated signal is given by:

$$y_i = \sum_{j=1}^{N_p} w_j x_i^{2j-1}. \quad (2)$$

Note that, to ensure monotonicity, only the odd terms are considered. Moreover, a non-negativeness constraint must be imposed to the coefficients w_j . In vector notation, the outputs of the polynomial compensating are given by

$$\mathbf{y} = \mathbf{X}^T \mathbf{w}, \quad (3)$$

where

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ x_1^3 & x_2^3 & \cdots & x_N^3 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{2N_p-1} & x_2^{2N_p-1} & \cdots & x_N^{2N_p-1} \end{bmatrix}. \quad (4)$$

Having defined the structure of the compensating function, it is necessary to choose a sparsity criterion to guide the adaptation of w_j . A possible choice in this case is to minimize the ℓ_0 -norm of $\mathbf{y}_c = \Phi \mathbf{y}$ [5], denoted $\|\mathbf{y}_c\|_0$, which simply counts the number of non-null elements of the signal in the transformed domain characterized by Φ — this matrix could be, for instance, the DCT matrix or other matrix associated with other frequency transform. In a practical situation, though, the use of a criterion based on the ℓ_0 -norm is limited, because the sparse signal often presents many elements that are close to zero, but not necessarily null. Therefore, one must consider practical approximations of the ℓ_0 -norm. In the present work, we make use of the smoothed version of the ℓ_0 -norm [9], which, for a signal \mathbf{z} of size N , is given by:

$$S_{\ell_0}(\mathbf{z}) = N - \sum_{i=1}^N k(z_i, \sigma), \quad (5)$$

where $k(z_i, \sigma)$ is Gaussian kernel of zero mean and standard deviation σ . Therefore, such measure approaches to the ℓ_0 -norm as σ tends to zero. The definition of σ should be done

based on the observation of how close to zero the low-energy elements of \mathbf{z} are.

Based on the elements described so far, our idea can be implemented by the following optimization problem:

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && S_{\ell_0}(\mathbf{y}_c) = S_{\ell_0}(\Phi \mathbf{X}^T \mathbf{w}) \\ & \text{subject to} && \mathbf{w}^T \mathbf{X} \mathbf{X}^T \mathbf{w} = 1 \\ & && w_i \geq 0, i = 1, \dots, N_p. \end{aligned} \quad (6)$$

As discussed earlier, the inequality constraints are required in order that monotonic compensating functions be obtained. The equality constraint is set to avoid trivial solutions, in which all the elements of \mathbf{y}_c are null or very close to zero — note that such a constraint corresponds to simply fixing the ℓ_2 -norm of \mathbf{y} to one.

The problem described above is an example of a nonlinear programming problem with equality and inequality constraints. There are many techniques to solve it. In the present work, this problem is tackled by an active-set algorithm [10] implemented by the function *fmincon*, available in the software Matlab.

3.3. Theoretical aspects

A very important question concerns the conditions for which the proposed idea is valid. To illustrate this point, we here provide some theoretical elements by considering a simple situation. Our analysis is based on the ℓ_0 -norm.

Let us assume that the global mapping between the input signal and the estimated one, $\mathbf{y} = g(f(\mathbf{s}))$, can be represented by $\mathbf{y} = \beta \mathbf{s} + \gamma h(\mathbf{s})$, where $\beta \in \mathbb{R}^*$, $\gamma \in \mathbb{R}$. This is the case, for instance, when the distorting function is given by $x_i = \sqrt[3]{s_i}$, and the compensating function by $y_i = \beta x_i^3 + \gamma x_i$. In this situation, the residual function h is given by $h = \sqrt[3]{s_i}$.

The implementation of our idea is achieved by maximizing the sparsity of $\mathbf{y}_c = \Phi \mathbf{y}$, which, in our analysis, corresponds to minimizing $\|\mathbf{y}_c\|_0$. In the sequel, we search for a sufficient condition assuring that the minimization of $\|\mathbf{y}_c\|_0$ necessarily leads to a perfect compensation, which corresponds to $\gamma = 0$.

Firstly, let us represent \mathbf{y}_c as follows

$$\mathbf{y}_c = \Phi(\beta \mathbf{s} + \alpha h(\mathbf{s})) = \beta \mathbf{s}_c + \gamma \Phi h(\mathbf{s}). \quad (7)$$

Although the ℓ_0 -norm is not a true norm, it satisfies the triangle inequality, and, consequently, the reverse triangle inequality. Moreover, the ℓ_0 -norm is scale invariant, and, thus, $\|\beta \mathbf{s}_c\|_0 = \|\mathbf{s}_c\|_0$. With these observations in mind, the following lower bound is obtained:

$$\|\mathbf{y}_c\|_0 \geq \left| \|\mathbf{s}_c\|_0 - \|\gamma \Phi h(\mathbf{s})\|_0 \right|. \quad (8)$$

Let $\|\mathbf{s}_c\|_0 = K$ and $\|\Phi h(\mathbf{s})\|_0 = L$. When $\gamma \neq 0$, one has $\|\gamma \Phi h(\mathbf{s})\|_0 = \|\Phi h(\mathbf{s})\|_0 = L$ and, thus,

$$\|\mathbf{y}_c\|_0 \geq \left| K - L \right|. \quad (9)$$

The key point here is that if $K < L/2$, then, from Equation (9), it asserts that $\|\mathbf{y}_c\|_0 > K$. Conversely, if $\gamma = 0$ (perfect compensation), then $\|\mathbf{y}_c\|_0 = K$. To sum up, the validity of our proposal is assured if the nonlinear function h decreases the sparsity of \mathbf{y} by, at least, a factor of two.

4. RESULTS

We consider a set of experiments in order to assess the performance of our proposal. In our tests, the input signal was generated as follows:

$$\mathbf{s} = \Phi \mathbf{s}_c = \Phi(\mathbf{s}_{c_S} + \alpha \mathbf{r}), \quad (10)$$

where \mathbf{s}_{c_S} correspond to realizations of a Bernoulli-Gaussian process, i.e., each element of \mathbf{s}_{c_S} has a probability P of being non-zero, and, when active, it is obtained from a zero-mean Gaussian distribution of unity variance. The elements of \mathbf{r} are also obtained from a standard Gaussian distribution. This vector, which is weighted by the coefficient α , is used to model the small residual energy typical of practical sparse signals. Therefore, when $\alpha = 0$ and P is small, \mathbf{s}_c becomes a sparse signal in the sense of the ℓ_0 -norm. Finally, in our tests, we consider that Φ corresponds to the DCT matrix. Note, however, that our approach is valid for other transformed domains for which the nonlinear distortion decreases the sparsity of the original signal, as discussed in Section 3.3.

In some tests, we consider a noisy generative model, in which the observation is given by:

$$\mathbf{x} = \mathbf{f}(\mathbf{s}) + \mathbf{n}, \quad (11)$$

where \mathbf{n} follows a Gaussian distribution. Finally, as performance index, we adopted the signal-to-interference ratio (SIR), which is given by:

$$\text{SIR} = 10 \log(\mathbf{s}^T \mathbf{s}) - 10 \log((\mathbf{s} - \mathbf{y}^*)^T (\mathbf{s} - \mathbf{y}^*)),$$

where \mathbf{y}^* denotes the recovered signal after performing a normalization² with respect to \mathbf{s} . This normalization is necessary, since we admit scale ambiguities.

4.1. Case in which perfect inversion is possible

Let us first consider the case in which the perfect compensation of the distorting function is possible. In this context, we address the compensation of $f(s_i) = \sqrt[3]{s_i}$ via the polynomial function (2). The sparse signal was generated according (10), considering $P = 0.2$, $\alpha = 0$ (ideal sparse signal), and $N = 1000$ samples.

The effect of the nonlinear distortion can be observed in Figure 2(a), which shows the mapping between s_i and x_i . We applied the proposed approach considering a polynomial of

²This normalization is given by $\mathbf{y}^* = k\mathbf{y}$, where $k = \arg \min_{k^*} (\mathbf{s} - k^* \mathbf{y})^T (\mathbf{s} - k^* \mathbf{y})$.

degree 7, i.e., $N_p = 4$ (see Equation (2)), and $\sigma = 0.01$. By observing Figure 2(b), which shows the mapping between s_i and y_i , one can readily note that our proposal led to a perfect compensation of the nonlinear distortion in this case.

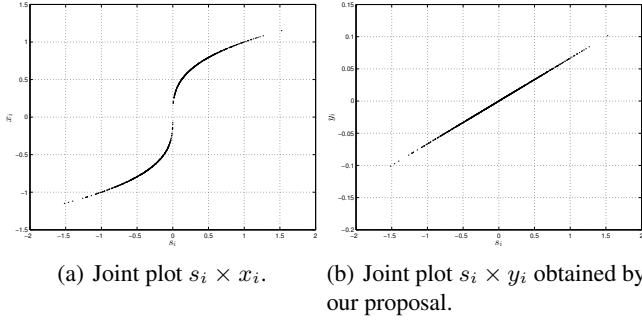


Fig. 2. Compensation when a perfect inversion is possible.

4.2. Case in which perfect inversion is not possible

Although the experiment presented in the last section is very useful in illustrating the performance of our approach, it is not always possible to achieve a perfect compensation in practical situations. For instance, we here consider a new situation in which $f(s_i) = \tanh(3s_i)$; $N = 1000$ samples of the input signal was generated considering $P = 0.2$ and $\alpha = 0$. In Figure 3(a), the nonlinear distortion is illustrated by the resulting mapping between s_i and x_i .

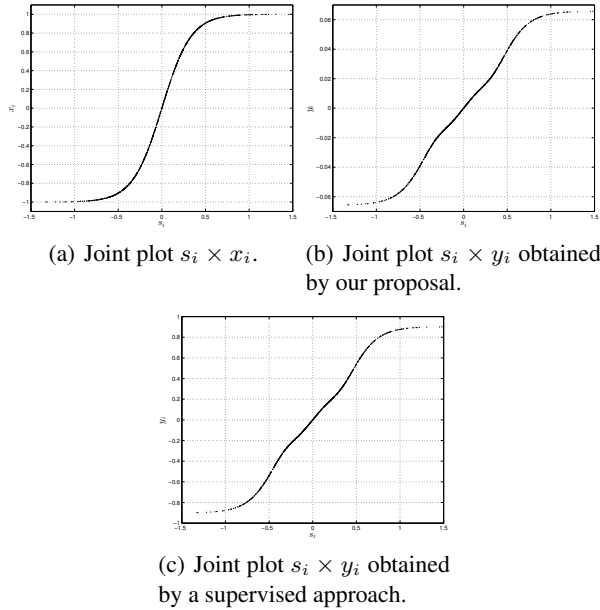


Fig. 3. Case in which a perfect inversion is not possible.

In this situation, we tested our method considering a polynomial of degree 9 ($N_p = 5$), and $\sigma = 0.1$ — we observed very close results considering other values of σ in the range

[0.001, 0.1]. As can be observed in Figure 3(b), our method was able to mitigate the effects of the nonlinear distortion; this is specially clear for high amplitudes. Concerning the signal-to-interference ratio, the proposal increased the original value of SIR= 9.4dB obtained between the input and the observed signals, to SIR= 17.1dB. For matter of comparison, we also considered a supervised approach, in which the polynomial coefficients were adjusted in order to minimize the mean-squared error (MSE) between the input signal s and the estimated signal y . The obtained input signal-estimated signal mapping is shown in Figure 3(c). Note that, as in the result obtained by our approach, there is still a residual non-linear distortion (SIR= 17.4dB). This reveals that the SIR obtained with the proposed method was actually very close to the structural limit inherent to the chosen nonlinear compensating device, thus attesting that an expressive performance level was reached.

4.2.1. Influence of residual elements

So far, we have considered that the input signal is perfectly sparse, having many null elements — this situation corresponds to assume $\alpha = 0$ in the signal model described in (10). However, as we mentioned before, actual sparse signals contain small coefficients that are not necessarily zero. Motivated by this observation, we here conduct an experiment to analyze the effects of these residual elements on the performance of our proposal. To accomplish this task, we performed a set of simulations in which the parameter α were varied from 0 to 1. In these tests, we considered $N_p = 4$ and the following different bandwidths: $\sigma = 0.1$ and $\sigma = 0.01$. The nonlinear distortion was set $f(s_i) = \tanh(1.5s_i)$ and $N = 1000$ samples of the input signal were generated considering $P = 0.2$ — in this case, we performed a variance normalization of the input signal to obtain fair comparisons between different values of α .

As can be seen in Figure 4, which shows the average SIR obtained considering 100 trials for each α , the performance of the proposed approach becomes worse as α increases. This is expected, since the input signal tends to be less sparse as α increases, violating the basic assumption underlying our proposal. Another interesting point here is the influence of σ . A higher value of σ , in this case, leads to a better performance as α increases. Conversely, for α close to zero, one can note that defining a smaller σ leads to a better performance, very close, in fact, to that associated with the supervised solution. Finally, note that, as expected, the performance of the supervised solution does not depend on the degree of sparsity of the input signal.

4.2.2. Performance in the presence of noise

A last experiment is performed with the aim of investigating the impact of noise on the proposed method. We consider the same scenario of the first experiment described in the present

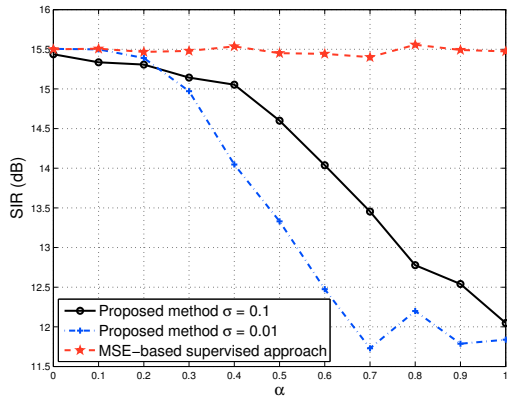


Fig. 4. Influence of α on the compensation.

section ($N = 1000$, $P = 0.2$, $\alpha = 0$, $N_p = 5$ and nonlinear function $f(s_i) = \tanh(3s_i)$). Noise was added according model (11). In Figure 5, we plot the evolution of the SIR as the signal-to-noise ratio (SNR) increases — each point corresponds to the average obtained in 100 trials. A first point that should be stressed here is that our proposal (with both $\sigma = 0.1$ and $\sigma = 0.01$) provided SIRs very close to the ones obtained by the MSE-based supervised solution. It is also interesting to stress the effects of the noise, which in nonlinear systems can be very harmful. In this example, for instance, we note a rapid performance degradation for SNRs smaller than approximately 30dB.

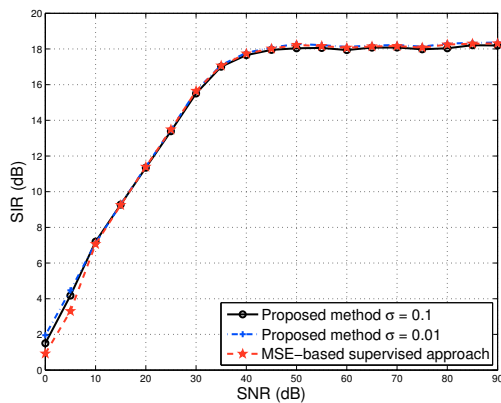


Fig. 5. Influence of noise on the compensation.

5. CONCLUSIONS

In this work, we proposed a new framework for blindly compensating nonlinear distortions. The groundwork of our proposal is the observation that, after applying a nonlinearity, the sparsity of the input signal in a given frequency transformed domain is usually lost. Therefore, our approach aims at adapting a compensating function by maximizing a sparsity mea-

sure, which in our work, was obtained from an approximation of the ℓ_0 -norm. The viability of our proposal was illustrated by means of a set of experiments considering synthetic data.

There are several points that should be addressed in future works. A first one concerns the derivation of precise conditions indicating when the proposed idea is valid. Although we provided some elements considering a simple scenario, a thorough investigation on this challenging subject would be quite helpful. Another topics that deserve further investigation are: i) the utilization of criteria built on other measures of sparsity, and ii) a study on more flexible structures (such as splines and monotonic neural networks) to be used in the compensating device. Finally, we intend to address the application of our approach to real-word data. We are currently testing it to process data acquired by chemical sensor arrays.

6. REFERENCES

- [1] P. Gründler, *Chemical sensors: an introduction for scientists and engineers*, Springer, 2007.
- [2] T. T. Ha, *Digital satellite communications*, McGraw-Hill, 1990.
- [3] M. Ibnkahla, “Applications of neural networks to digital communications: a survey,” *Signal Processing*, vol. 80, pp. 1185–1215, 2000.
- [4] J. Tsimbinos and K.V. Lever, “Nonlinear system compensation based on orthogonal polynomial inverses,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 48, no. 4, pp. 406–417, April 2001.
- [5] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*, Springer, 2010.
- [6] H. J. Landau and W. L. Miranker, “The recovery of distorted band-limited signals,” *Journal of Mathematical Analysis and Applications*, vol. 2, pp. 97–104, 1961.
- [7] K. Dogancay, “Blind compensation of nonlinear distortion for bandlimited signals,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 9, pp. 1872–1882, 2005.
- [8] L. T. Duarte, R. Suyama, R. R. F. Attux, B. Rivet, C. Jutten, and J. M. T. Romano, “Source separation of base-band signals in post-nonlinear mixtures,” in *Proceedings of the IEEE MLSP 2009*, 2009.
- [9] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, “A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm,” *IEEE Transactions on Signal Processing*, vol. 57, pp. 289–301, 2009.
- [10] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, 1984.