BLIND COMPENSATION OF MEMORYLESS NONLINEAR DISTORTIONS IN SPARSE SIGNALS

Jiri Malek

Faculty of Mechatronics, Informatics, and Interdisciplinary Studies, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic.

jiri.malek@tul.cz

ABSTRACT

This paper focuses on blind compensation of memoryless nonlinear distortions in sparse signals. A nonlinear distortion tends to decrease the sparsity. We propose to compensate for the unknown distortion using a general monotonic function, such that the compensated signal is as sparse as possible.

Novel estimator for the compensation function is presented, which is able to compensate for both symmetric and asymmetric distortions. The selection of a sparsity measure is discussed with respect to compensation performance. The functionality of the method is evaluated in experiments with artificially distorted real-world speech signals.

Index Terms— Memoryless nonlinear distortion, blind compensation, sparse signal.

1. INTRODUCTION

Many analog amplifiers/transducers exhibit some kind of nonlinear behavior, causing the processed signal to be distorted to some extend. This is encountered, e.g., in the context of electro-chemical sensors or acquisition of audio signals.

In the digital domain, these distortions are modeled as nonlinear systems, which either have memory or are memoryless (instantaneous). There exist several distinct models to represent nonlinear systems with memory, e.g., the Voltera filters or the nonlinear auto-regressive moving average (NARMA) model [1]. The simpler memoryless nonlinearities are modeled as functions applied element-wise to the signal. Many complex distortions can be described as a sequence of an instantaneous nonlinearity and linear system(s). An example is the LTI-ZMNL-LTI model [2] (Linear Time Invariant - Zero Memory NonLinear, i.e. two linear subsystems separated by instantaneous nonlinearity) or the Hammerstein model [3] (an instantaneous nonlinearity followed by an all-pole linear system).

In the case, when some specific input training signals and the corresponding outputs are available, the retrieval of the original signal often proceeds through identification of the nonlinear system (review of methods is given in [4]) and compensation by its (approximate) inversion [5].

When the input signal is unavailable, the identification can be performed in a blind fashion, based on some strong prior knowledge concerning the statistical properties of the input. Blind identification of complete Voltera representation is difficult though and thus, the methods often resort to simplified models. Identification of LTI-ZMNL-LTI model was proposed for circularly symmetric Gaussian input in [2]. Hammerstein nonlinearities were identified, assuming long Gaussian processes as inputs, in [3].

Another group of approaches attempts to compensate the nonlinear system directly, without its prior identification. The direct compensation of memoryless nonlinearities, based on known/estimated cumulative distribution function (cdf) of the original signal, is performed by Histogram Equalization (HEQ) [6] techniques, which are used in various variants as preprocessing prior the automatic speech recognition.

However, much weaker assumptions about the unknown input are sufficient to perform the direct blind compensation of memoryless nonlinearities. Compensation method for band limited signals was presented in [7]. It was based on the fact that passing of signal through nonlinear system causes spectral spreading, i.e. a spectral content outside the original band appears in the output signals.

An alternative weak assumption is the sparsity of the original signal in some known domain, which is decreased by the applied distortion. Direct blind compensation based on this principle was proposed in [8], where a memoryless nonlinearities distorting artificial sparse signals were compensated by polynomial functions.

Current paper extends the idea of *blind compensation* for memoryless distortions based on sparsity recovery. Novel estimator for the compensation function is presented, which is able to compensate for both symmetric and asymmetric distortions. Suitable measure of sparsity is selected with respect to compensation performance. The functionality of the technique is demonstrated in experiments with artificially distorted real-world speech signals.

This work was supported by the Technology Agency of the Czech Republic (project no. TA01011142).

2. PROBLEM FORMULATION

Let the vector $\mathbf{s} \in \mathbb{R}^N$ represent discrete unknown signal consisting of N samples $s[n], n = 1..N$. Let the signal $\mathbf{s}$ be distorted by an unknown function $f : \mathbb{R} \rightarrow \mathbb{R}$, which is applied element-wise (as are all functions applied to vectors in this paper), such that

$$\mathbf{x} = f(\mathbf{s}), \tag{1}$$

gives the observed signal $\mathbf{x}$. The task considered in this work is the blind estimation of the original signal $\mathbf{s}$, based only on the samples of the distorted signal $\mathbf{x}$.

The estimation is not possible without additional assumptions about $\mathbf{s}$ and $f(\cdot)$. Our key assumption is that signal $\mathbf{s}$ is *sparse* in some known domain. Further, we require the function $f(\cdot)$ to be *monotonic*, in order to be invertible.

3. PROPOSED METHOD

3.1. Basic outline

The estimation of $\mathbf{s}$ is performed via a monotonic compensating function $g : \mathbb{R} \rightarrow \mathbb{R}$, such that the estimate $\mathbf{y}$ given by

$$\mathbf{y} = g(\mathbf{x}) = g(f(\mathbf{s})) \tag{2}$$

is as close to $\mathbf{s}$ as possible. The blind estimation admits scaled versions of $\mathbf{s}$ as perfect solutions for the task (scaling ambiguity).

Our procedure is based on observation from [8] that a nonlinear distortion applied to signal $\mathbf{s}$ decreases the sparsity. This is known for speech in the frequency domain, where the nonlinearity causes formation of new harmonic frequencies in the spectrum [7]. Therefore, we propose to estimate $\mathbf{s}$ via application of function $g(\cdot)$, such that its output $\mathbf{y}$ is as sparse as possible.

The selection of transform domain where $\mathbf{s}$ is considered sparse does not influence structure of the proposed procedure. Therefore, we restrict ourselves in this paper to the Discrete Cosine Transform (DCT) domain, due to its convenient properties. The DCT tends to concentrate the signal energy in a few low frequency components [9] for a large class of signals (such as audio signals or images). Moreover, DCT is a real transform, which simplifies further optimization.

The implementation of the compensation procedure consists of three steps; 1) the selection of the structure of the compensation function $g(\cdot)$, 2) the selection of a suitable sparsity measure (denoted by $S(\cdot)$) and 3) the formulation of the optimization problem and the choice of an algorithm for its numerical solution.

3.2. Structure of the compensating function

A polynomial is a reasonable candidate for the compensation function $g(\cdot)$, because an arbitrary analytical function can be approximated by its power series expansion. Therefore, authors in [8] propose compensation function given as polynomial of the form

$$\mathbf{y} = g_p(\mathbf{x}) = \sum_{i=1}^{M} w_i \mathbf{x}^{2i-1}. \tag{3}$$

To ensure the monotonicity, only the odd order monomials are included in $g_p(\mathbf{x})$ and the weights w_i are required to be positive. These constraints restrict the flexibility of $g_p(\cdot)$ considerably.

We propose to utilize a more general class of $g(\cdot)$ given as

$$\mathbf{y} = g_s(\mathbf{x}) = \sum_{i=1}^{M} w_i b_i(\mathbf{x}), \tag{4}$$

where $b_i(\cdot)$ are some monotonic basis functions and $w_i \geq 0$. In order to allow $g_s(\cdot)$ compensate broad range of distortions, we suggest to select $b_i(\cdot)$ as inverse functions to known potential nonlinearities, supplemented by odd order monomials and root functions.

In order to allow compensation of asymmetric distortions, the structure of the compensation function can be extended to

$$\mathbf{y} = g_a(\mathbf{x}) = \sum_{i=1}^{M} w_{m,i} b_{m,i}(\mathbf{x}) + \sum_{i=1}^{M} w_{p,i} b_{p,i}(\mathbf{x}), \tag{5}$$

where $b_{m,i}(\cdot)$ and $b_{p,i}(\cdot)$ are bases derived as positive and negative parts of the domain of $b_i(\cdot)$, such that

$$b_{m,i}(\cdot) = \begin{cases} b_i(\cdot) & \text{if } b_i(\cdot) < 0 \\ 0 & \text{else,} \end{cases} \tag{6}$$

and

$$b_{p,i}(\cdot) = \begin{cases} b_i(\cdot) & \text{if } b_i(\cdot) > 0 \\ 0 & \text{else.} \end{cases} \tag{7}$$

Function $g_a(\cdot)$ is able to compensate independently the positive/negative intervals within the domain of asymmetric distortion $f(\cdot)$, allowing accurate estimation of $\mathbf{y}$.

3.3. Selection of sparsity measure

An optimal sparsity measure should strongly reflect the nonuniformity in distribution of signal energy and should be computationally simple. To this purpose, the ℓ_0 norm is often stated. However, the ℓ_0 norm is severely discontinuous function (which optimization is restricted to combinatorial search) and is sensitive to noise. To overcome these properties, the *smoothed ℓ_0 norm* was proposed as a sparsity measure in [8], given as

$$S_{\ell_0}(\mathbf{Y}) = N - \sum_{k=1}^{N} a(Y[k], \sigma_a), \tag{8}$$

where $\mathbf{Y}$ is the DCT of $\mathbf{y}$, $Y[k], k = 1..N$ are its elements and $a(Y[k], \sigma_a)$ is a Gaussian kernel of zero mean and standard deviation σ_a. The free parameter σ_a determines the smoothness of $S_{\ell_0}(\cdot)$. For large values of σ_a, the measure is smoother, but it becomes worse approximation of ℓ_0 norm. The optimal value of σ_a depends on measured data and is difficult to estimate.

To avoid the dependence on the free parameter and in order to evaluate the influence of sparsity measure on the compensation performance, we propose two alternative functions $S(\cdot)$. These functions are not smooth, but their optimization is much simpler compared to ℓ_0 norm (can be performed, e.g., through simple subgradient method [10] belonging to convex optimization). The frequently utilized ℓ_1 *norm* is given as

$$S_{\ell_1}(\mathbf{Y}) = \sum_{k=1}^{N} |Y[k]| \,. \tag{9}$$

The *Hoyer measure* (proposed in [11]) is a normalized ratio of ℓ_1 and ℓ_2 norms, given by

$$S_H(\mathbf{Y}) = \left(\sqrt{N} - \frac{\sum_{k=1}^{N} |Y[k]|}{\sqrt{\sum_{k=1}^{N} Y[k]^2}}\right) \left(\sqrt{N} - 1\right)^{-1}. \tag{10}$$

This measure equals 1 if $\mathbf{Y}$ contains only a single non-zero component and equals 0 if all components are equal up to signs. Since the Hoyer measure grows with increasing sparsity, it should be negated prior the minimization.

3.4. Optimization Problem

The optimization problem of finding the sparsest $\mathbf{Y}$ can be formulated as

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & S(\mathbf{Y}) = S\left(\sum_{i=1}^{M} w_i \mathbf{V}_i\right) \\ \text{subject to} & \|\mathbf{Y}\|_2 = 1 \\ & w_i \geq 0,\ i = 1..M, \end{array} \tag{11}$$

where $\mathbf{V}_i$ are the DCT transforms of signals $\mathbf{v}_i = b_i(\mathbf{x})$ and $\mathbf{w}$ is the vector composed of weights $w_i, i = 1..M$. The equality constraint prevents trivial solution and the inequality constraint ensures the monotonicity of $g_s(\cdot)$.

For the case of asymmetrical compensation function, the optimization problem needs to be modified as

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & S(\mathbf{Y}) = S\left(\sum_{i=1}^{M} w_{m,i}\mathbf{V}_{m,i} + \sum_{i=1}^{M} w_{p,i}\mathbf{V}_{p,i}\right) \\ \text{subject to} & \left\|\sum_{i=1}^{M} w_{m,i}\mathbf{V}_{m,i}\right\|_2 = 1 \\ & \left\|\sum_{i=1}^{M} w_{p,i}\mathbf{V}_{p,i}\right\|_2 = 1 \\ & w_{m,i} \geq 0,\ w_{p,i} \geq 0,\ i = 1..M, \end{array} \tag{12}$$

where $\mathbf{V}_{m,i}$ and $\mathbf{V}_{p,i}$ are DCT transforms of signals $\mathbf{v}_{m,i} = b_{m,i}(\mathbf{x})$ and $\mathbf{v}_{p,i} = b_{p,i}(\mathbf{x})$, respectively. The ℓ_2 norm of each set of basis functions needs to be constrained separately in order to prevent trivial solutions on positive/negative interval of the $g_a(\cdot)$ domain of definition.

Table 1. Distortions considered in the experiments

Abbreviation	$f(z)$
Pow	$f(z) = z^5$
Tanh	$f(z) = \tanh(5z)$
αTanh	$f(z) = \alpha \tanh(5z)$
Pw3Pw7	$f(z) = \begin{cases} z^3 & \text{if } z < 0 \\ z^7 & \text{else} \end{cases}$

Signals $\mathbf{V}_i$ differ much in energy, which complicates the minimization. It is convenient to normalize the energy of $\mathbf{V}_i$ prior the optimization.

The minimization in (11) is a nonlinear programing task. The experiments presented in the following section were solved by an active-set algorithm provided by function `fmincon` in Matlab®. Although the `fmincon` solver is intended primarily for smooth objective functions, it is known to be effective on some non-smooth problems as well [12].

As an implementation-simple alternative to complex active-set algorithm, the constrained subgradient method [10] can be utilized. This method is intended for solving convex minimization problems and is proved to converge even for non-differentiable objective functions. The main advantage of the subgradient method resides in its implementation simplicity, but the convergence rate is slower compared to the active-set algorithm.

4. EXPERIMENTAL EVALUATION

In the following section, we present several experiments with artificially distorted real-world *speech signal* to demonstrate the functionality of the proposed method. Speech is known to be sparse in the DCT domain. We utilize a lecture recorded for streaming purposes at our university. The recording is sampled at 16kHz, lectured in Czech by female speaker and is approximately 90 minutes long. The signal is captured by a close-talk microphone. The common background noise of a lecture hall is present in the recording.

Within all following experiments, the speech signal is normalized to range [-1,1] and a synthetic distortion $f(\cdot)$ from in Table 1 is applied to it.

Signal frame of length $0.1\,$s is sufficient for the estimation of $g(\cdot)$. The distorted recording is divided into such frames and the compensation is performed and evaluated in each frame separately, in order to show the average properties of the method. Frames with low energy are discarded, in order to avoid estimation in segments without speech. We use 1500 frames in the experiments. The parameter α for the αTanh distortion is selected in each frame independently, such that $\mathbf{x}$ in that frame has unitary ℓ_1 norm.

We utilize two different sets of basis functions. The first, denoted by B1, consists of odd order monomials ($b(z) = z^c, c \in \{1,3,5,7\}$), odd order root functions ($b(z) = \sqrt[c]{z}, c \in \{3,5,7\}$) and the inverse hyperbolic tangent (denoted as aTanh), included as inverse for soft clipping nonlinearity Tanh. The other set, denoted as B2, consists of the same monomials and root functions, but misses the aTanh basis.

The proposed compensation structures/sparsity measures are compared to proposal from [8], i.e. purely polynomial structure $g_p(\cdot)$ with smoothed ℓ_0 norm ($\sigma_a = 0.01$, $M = 4$). Moreover, the proposed procedure is compared to a well-known technique for compensation of memoryless nonlinear distortions - Histogram equalization (HEQ) [6]. The technique is not blind, since it assumes the knowledge of the cdf of the original signal. In our case, this cdf was estimated using another lecture (of length 90 minutes) given by the same female speaker in the same room as the lecture to-be compensated. The cdf of the distorted signal was estimated using the whole distorted recording as well.

As a measure of compensation performance we utilize the Signal to Distortion Ratio [dB], given by

$$\mathrm{SDR} = 10\log\left(\frac{\mathbf{s}^T\mathbf{s}}{\min_\beta[(\mathbf{s}-\beta\mathbf{y})^T(\mathbf{s}-\beta\mathbf{y})]}\right), \qquad (13)$$

where β is a parameter compensating the scaling ambiguity. The results are stated in the means of *SDR improvement*, i.e. the difference of SDRs between the compensated signal and the distorted signal.

4.1. Sparsity measures and compensation structures

In this experiment, we compare the performance of the considered sparsity measures and structures of the compensation function. The results are summarized in Tables 2 and 3.

The best results are often achieved using the $S_{\ell_0}(\cdot)$ measure, with varying value of σ_a, however. The optimal value of σ_a depends on the applied compensation structure $g(\cdot)$ and the encountered distortion. Values $\sigma_a \in [0.01, 5\cdot 10^{-4}]$ seem to be adequate for speech signal compensation, outside this interval the performance drops rapidly. The results of the non-smooth $S_{\ell_1}(\cdot)$ measure are comparable to $S_{\ell_0}(\cdot)$, without dependence on the free parameter.

The compensation of Tanh distortion is affected by scaling of $\mathbf{x}$. It stems from the fact that, unlike to monomial/root distortions, atanh(z/α) is not a scaled version of atanh(z), i.e. the basis ensuring perfect compensation for αTanh is not available in the B1 set. Nevertheless, the presence of aTanh in B1 improves the compensation performance of αTanh distortion compared to compensation via B2 or $g_p(\cdot)$ from [8], where aTanh is missing. This demonstrates the advantage of the inclusion of the inverses to known potential nonlinearities in $g_s(\cdot)$. Moreover, $g_p(\cdot)$ is completely unable to compensate the polynomial nonlinearities (such as Pow and Pw3Pw7).

As expected, the ability of symmetric compensation function $g_s(\cdot)$ (regardless to available bases) to compensate asymmetric distortion Pw3Pw7 is limited.

Table 2. Achieved SDR improvement averaged over 1500 frames, using the compensation structure $g_s(\cdot)$. Unless otherwise stated, basis function set B1 is used.

Measure	Pow	Tanh	αTanh	αTanh (B2)	Pw3Pw7
$S_{\ell_0}(1\cdot 10^{-2})$	44.43	49.05	16.45	9.51	8.98
$S_{\ell_0}(5\cdot 10^{-3})$	42.68	47.69	17.61	9.55	8.64
$S_{\ell_0}(1\cdot 10^{-3})$	44.01	41.01	17.09	8.64	6.92
$S_{\ell_0}(5\cdot 10^{-4})$	47.20	43.74	15.78	8.08	5.99
Hoyer	39.89	40.02	16.57	8.91	8.86
S_{ℓ_1}	49.89	42.72	16.66	9.05	8.87
HEQ	14.86	9.25	9.25	9.25	13.48
Proposal in [8]	-	9.50	9.50	9.50	-

The compensation of symmetric distortions via $g_a(\cdot)$ achieves lower SDR values compared to compensation via $g_s(\cdot)$. The difference is most evident in cases, when the basis for perfect compensation is available in the basis set. On the other hand, the utilization of function $g_a(\cdot)$ leads to significant performance improvement compared to $g_s(\cdot)$, when compensating asymmetric nonlinearity.

The proposed procedure outperforms HEQ in scenarios, when the compensation structure $g(\cdot)$ respects the distortion, i.e. asymmetric nonlinearity is compensated by $g_a(\cdot)$. HEQ achieves comparable or better results in situations, when the set of basis functions does not contain a basis allowing perfect compensation (distortion αTanh compensated via basis set B2).

Regarding computational load, HEQ is less demanding compared to the proposed approach. For the unoptimized Matlab® implementations (executed on computer with 3.4GHz quad-core processor), the compensation of 1500 frames via HEQ required 22s. This includes the estimation of the cdfs, i.e. the most demanding phase of HEQ. The time required by the proposed procedure depends on the number of basis functions M, compensation structure $g(\cdot)$ and the initialization of $\mathbf{w}$. Considering initialization $w_i = 1, i = 1..M$, for $g_s(\cdot)$ the computation took about 31s when $M = 4$ and about 66s for $M = 8$. For $g_a(\cdot)$ the computation took about 42s when $M = 4$ and about 210s for $M = 8$.

4.2. Preprocessing prior to the automatic transcription of speech

The proposed method can be utilized e.g. as a part of preprocessing for nonlinearly distorted speech prior automatic speech recognition (ASR). Although the memoryless model by itself may not be general enough for all complex nonlinearities encountered in audio, the presented approach could be used for the compensation of the instantaneous nonlinear subsystem in the LTI-ZMNL-LTI or Hammerstein model.

In this experiment, we improve the recognition accuracy

Table 3. Achieved SDR improvement averaged over 1500 frames, using the compensation structure $g_a(\cdot)$. Unless otherwise stated, basis function set B1 is used.

Measure	Pow	Tanh	αTanh	αTanh (B2)	Pw3Pw7
$S_{\ell_0}(1 \cdot 10^{-2})$	24.59	15.96	14.20	8.46	26.44
$S_{\ell_0}(5 \cdot 10^{-3})$	25.86	16.86	14.75	8.22	27.49
$S_{\ell_0}(1 \cdot 10^{-3})$	27.16	17.81	13.84	7.17	28.73
$S_{\ell_0}(5 \cdot 10^{-4})$	24.16	18.26	12.96	6.53	24.91
Hoyer	24.36	15.65	13.86	7.91	27.13
S_{ℓ_1}	26.64	16.68	13.97	7.95	28.58
HEQ	14.86	9.25	9.25	9.25	13.48
Proposal in [8]	-	9.50	9.50	9.50	-

for artificially distorted speech submitted to the large vocabulary continuous speech recognition (LVCSR) system.

The utilized recognition system for Czech language [13] uses an acoustic model based on tied-state context dependent Hidden Markov Models of Czech phonemes and several non-speech events. The lexicon contains 504000 items.

We distort the considered lecture by function $f(z) = \tanh(bz)$ simulating soft clipping encountered during recording and submit the signal to the transcription engine. Parameter b influences the degree of distortion. Subsequently, we compensate the signal with the presented approach and let it be transcribed again. The compensation is performed via function $g_s(\cdot)$, either with the basis set B1 or B2. The weights $\mathbf{w}$ are given as an average weight vector estimated over 200 frames of $\mathbf{x}$ with the highest energy. The cdf of the distorted signal, required by HEQ, was estimated using the whole distorted recording.

The results, stated by means of recognition accuracy (ACC), are shown in Figure 1. The best performance was achieved using the basis set B1, when the recognition performance for compensated data approaches the original value 74.88%. The results achieved by both B1 and HEQ are independent from the distortion degree. For the basis set B2, the achieved accuracy is lower compared to B1 and comparable or lower to HEQ. However, even when the basis aTanh allowing perfect compensation is unavailable, the presented approach is able to compensate the distortion significantly. The compensation performance of basis set B2 deteriorates with higher degrees of distortion, because the monomial/root bases alone cannot entirely invert the highly nonlinear distortion.

5. CONCLUSIONS

We proposed an approach for blind compensation of sparse signals affected by a memoryless nonlinear distortion. The functionality of the technique is shown in experiments with real-world speech signals distorted by artificial nonlinearities. Achieved compensation is demonstrated through significant

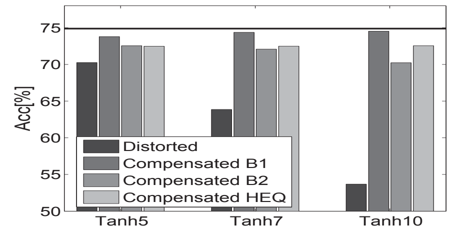

Fig. 1. Automatic recognition accuracy for experiment described in Section 4.2. The bold black line indicates recognition accuracy for undistorted lecture (74.88%).

improvement of the Signal to Distortion Ratio and of the automatic transcription accuracy.

Further work needs to address a generalization towards compensation of distortions with memory, which are encountered often e.g. in audio applications. Here, the present method can be used to compensate the zero-memory nonlinear subsystem within the LTI-ZMNL-LTI model or the Hammerstein model.

6. REFERENCES

[1] Tommy W. S. Chow, Hong-Zhou Tan, and Yong Fang, *Nonlinear System Representation*, John Wiley and Sons, Inc., 2001.

[2] S. Prakriya and D. Hatzinakos, "Blind identification of nonlinear models using higher order spectral analysis," in *Acoustics, Speech, and Signal Processing, 1995 International Conference on*, 1995, vol. 3, pp. 1601 –1604.

[3] G. Picard and O. Cappe, "Blind identification of hammerstein nonlinear distortion models," in *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.*, 2003, pp. 17 – 20.

[4] Hong-Zhou Tan and T. Aboulnasr, "Tom-based blind identification of nonlinear volterra systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 55, no. 1, pp. 300 – 310, 2006.

[5] J. Tsimbinos, *Identification and Compensation of Nonlinear Distortion*, Ph.D. thesis, University of South Australia, 1995.

[6] De la Torre A., Peinado A., Segura J.C., Perez Cordoba J.L., Benitez C., and Rubio A., "Histogram equalization of speech representation for robust speech recognition," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 3, pp. 355–366, 2005.

[7] K. Dogancay, "Blind compensation of nonlinear distortion for bandlimited signals," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 52, no. 9, pp. 1872 – 1882, 2005.

[8] L. T. Duarte, R. Suyama, R. Attux, Romano J. M. T., and Jutten Ch., "Blind compensation of nonlinear distortions via sparsity recovery," in *EUSIPCO 2012*, 2012, pp. 2362–2366.

[9] K. H. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, 1990.

[10] Bertsekas D., *Nonlinear Programming*, Athena Scientific, 1999.

[11] P. O. Hoyer and P. Dayan, "Non-negative matrix factorization with sparseness constraints," *Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.

[12] MathWorks Inc., "Optimization toolbox help - choosing a solver," `http://www.mathworks.com/help`, 2013.

[13] P. Cerva et al., "Browsing, indexing and automatic transcription of lectures for distance learning," in *MMSP'12*, Banff, AB, Canada, September 2012, pp. 198–202.