RESTORATION OF NONLINEARLY DISTORTED AUDIO WITH THE APPLICATION TO OLD MOTION-PICTURES

TAMÁS B. BAKÓ, BALÁZS BANK AND TAMÁS DABÓCZI

Budapest University of Technology and Economics (BUTE),
Department of Measurement and Information Technology,
Budapest, Hungary
{bako,bank,daboczi}@mit.bme.hu

In this paper a robust and efficient method is presented for restoration of nonlinearly distorted movie soundtracks. The method is based on the a priori knowledge of the shape of the nonlinear function, which is assumed to be a static nonlinearity. The original undistorted signal is modeled by a set of harmonically related sinusoids. This signal is led through the model of the nonlinear function. The parameters of the model are determined by minimizing the difference of the modeled distorted signal and the signal of the movie soundtrack. Tikhonov regularization is applied to avoid noise intensification during the restoration process. The proposed method has been successfully applied to old motion-picture audio tracks.

INTRODUCTION

Old movies often suffer from poor sound quality, mostly due to nonlinear distortion of the audio signal. In the professional (35 mm) films, the sound is optically recorded. Nowadays the transversal recording technique is used, where the sound information is carried by the width of the sound-stripe (left of Fig. 1). This is advantageous, because the development conditions and the strength of the recording light have little influence on the recorded sound. However, until the 1950's, the intensity recording was used (right of Fig. 1). Here, the sound-information is carried by the darkness (density) of the sound stripe. The density-characteristics of the film is a static nonlinear function of the intensity (Fig. 2). At high signal level of sound, or at wrong working point of the density curve, the recorded sound can be strongly distorted. This distortion sometimes even makes the sound recording incomprehensible. If the distortion is unacceptable, the distorted signal should be post processed to reconstruct the original sound. The signal reconstruction is a difficult task, since the exact nonlinear characteristics is not known and the distorted signal is superimposed by wide-band noise.

Most of prior works dealing with compensation of nonlinear systems use Volterra-kernels to describe the nonlinearity. These methods can handle a wide range of nonlinearities ([1], [2], [3] and [4]).

Somewhat different method is used for nonlinear compensation in [5]. There, a static nonlinear system (a cathode ray tube) is described. The static nonlinearity is approximated by a polynomial.

In [6], [7] and [4] the proposed algorithms are made specifically for sound-restoration. In [6], a histogram equalization technique is used to estimate a static nonlinear transfer function in the case of human speech.

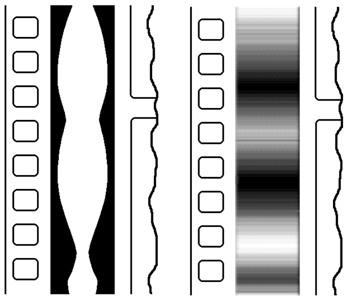

Figure. 1: The shape of the sound band on movie-films created with transversal (left) and variable density (right) method.

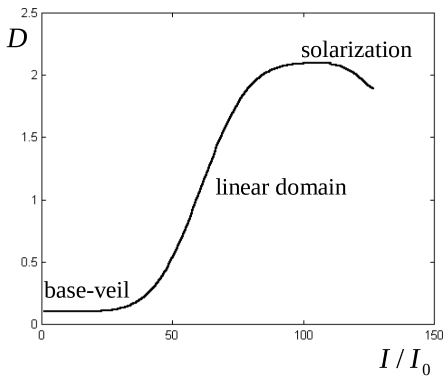

Figure 2: The density characteristics of the film.

In [7], the nonlinear function is assumed to be given and an iterative technique is used for restoration. In [4], a statistic-model based reconstruction method is proposed. As most of these previous works state, the effects of noise in the proposed algorithms are not clearly described. Further work is needed to establish the effects of noise in [1] and [2]. In [6], the algorithm is not applicable, when the histogram of the noise is markedly different from that of speech. In [3] and in [5], a prefiltering technique was used, where the effects of noise are smaller, therefore the effects of noise were not handled. However, in the case of movies only postfiltering technique can be used to reconstruct the observed, noisy and nonlinearly distorted signals, because the signal is already recorded on the optical track. The methods in [7] and [4] are able to handle the effect of noise, but they are iterative algorithms and they require intensive computational time. The optimal number of iterations in these works is not given.
In our paper a new method is proposed, which is a noniterative post-processing technique that works on static nonlinearities and takes the effect of noise into account. The method consists of two main steps. First is the identification of the static nonlinearity. The shape of the nonlinear function is given, since the density function of the film is known. However, at a given film-roll the working point and the amplification of the devices after the nonlinearity are not known. These parameters have to be identified from the recorded signal. This problem will be discussed in Section I.
The second step of the post processing technique is the restoration of the distorted signal. The nonlinearly distorted signal is corrupted by noise. In this case the exact inverse of the nonlinearity may not be optimal for reconstruction, because the noise is amplified during the reconstruction process and the noise level in the reconstructed signal can exceed the original sound level. An optimal characteristics is needed, which makes a trade-off between the distorted, and the undistorted but noisy signal. In Section II, these problems and the computation of the optimal (regularized) characteristics will be shown. Section III shows a simulation example. Conclusions are given in Section IV.

1. IDENTIFICATION OF THE NONLINEARITY

At film-rolls, the sound can be distorted due to the static nonlinear density-function of the film. We describe the nonlinear function by the following equation:

$$y(t) = G_1 \cdot \Phi(G_2 \cdot x(t) + O_2) + O_1 , \qquad (1)$$

where $x(t)$ is the original, undistorted signal, $y(t)$ is the distorted signal, and $\Phi(\,)$ refers to the density function that is assumed to be known. G_2 and O_2 are the amplification and offset before, G_1 and O_1 are the amplification and offset after the nonlinearity. These are produced by the recording and playing amplifiers, respectively. The amplification and the offset are assumed to be constant in the case of a particular film-roll. The observed signal is corrupted by noise:

$$o(t) = y(t) + n(t) , \qquad (2)$$

where $o(t)$ is the observed signal and $n(t)$ is the noise, which is assumed to be wide-band and zero mean.
For reconstructing $x(t)$, the values of G_1, O_1 and O_2 have to be determined. Note that G_2 is not important, because this parameter only adjusts the volume of the original sound. The reconstruction is difficult, since only the observed signal, $o(t)$, is known. However, the recorded signal is mainly human voice and this can be used as an a priori information.
If the recorded signal is periodic, it can be written as a sum of harmonically related sinusoids:

$$s(t) = \sum_i a_i \cdot \sin(i \cdot 2\pi f_0 \cdot t + \phi_i) , \qquad (3)$$

where $s(t)$ stands for the periodic signal, f_0 for the fundamental frequency of the periodical signal, and a_i and ϕ_i are the amplitude and phase of the i-th sinusoid. In (3), we assume, that $s(t)$ has no DC component. If the signal, $s(t)$, is led through a static nonlinear system, a different periodic signal arises:

$$\begin{aligned} u(t) &= G_1 \cdot \Phi(G_2 \cdot s(t) + O_2) + O_1 = \\ &= \sum_j b_j \cdot \sin(j \cdot 2\pi f_0 \cdot t + \varphi_j) + b_0 \end{aligned} . \qquad (4)$$

Eq. (3) and (4) form a common transformation, which assigns a $u(t)$ signal to every value of the unknown parameters:

$$u(t) = T\left(\underline{v}(f_0, t)\right) , \qquad (5)$$

where $\underline{v}(f_0, t)$ is the set of the unknown variables:

$$\underline{v}(f_0, t) = \{G_1, O_1, O_2, a_1 ... a_N, \phi_1 ... \phi_N\} . \qquad (6)$$

The unknown variables can be obtained, if the $T(\,)$ transformation is invertible. A sufficient condition is, when the number of a and ϕ parameters are limited, $s(t)$ has no DC component, and $\Phi(\,)$ is a strictly monotonic nonlinear function.
In the case of movie-soundtracks, these conditions are usually fulfilled. The uttered vowels in the movie contain periodic parts, which are ideal for the identification. The sound has no DC component, or if it is removed, it does not affect the sound-quality. The

recording is bandlimited, and the density-function of the film is a strictly monotonic function.

The signal of movie soundtracks is corrupted by wide-band noise. In this case, the problem is ill-posed, because the observed samples can exceed the limits of the output domain of the nonlinear function. A solution for $\underline{v}(f_0,t)$ can be found by minimizing the following form:

$$Cost = \int_{t_1}^{t_2} \left(u(t) - T\left(\hat{\underline{v}}(f_0,t)\right)\right)^2 dt \,. \tag{7}$$

Least mean squares minimization is optimal in the presence of white Gaussian noise, but it was found to be robust for the colored noise of film rolls, too.

The cost-function can be minimized by Monte-Carlo method. It is still a question, how many sinusoids should be used to describe the original signal $s(t)$. This can be estimated from the graph of the optimal cost versus the number of sinusoids. If the $s(t)$ signal is undermodeled, the cost will be high and will quickly decrease for higher number of sine signals. On the contrary, if the periodic signal is overmodeled, the use of higher number of sinusoidal signals will not change the optimal cost drastically. Hence, by finding this turning point, the number of sinusoids can be choosen. The use of eight sinusoidal signals has been found appropriate to give good results.

2. THE OPTIMAL INVERSE CHARACTERISTICS

2.1. Model of the reconstruction process

The signal model of the reconstruction process can be seen in Fig. 3, where $\Phi(x)$ denotes the nonlinear function of the measurement system, $x(t)$ denotes the input and $y(t)$ refers to the distorted output of the system where $y(t) = \Phi(x(t))$. The observation, $o(t)$, is disturbed by additive measurement noise, $n(t)$. $\mathrm{K}(o)$ is the inverse nonlinear function and $\hat{x}(t)$ is the estimate of the input signal.

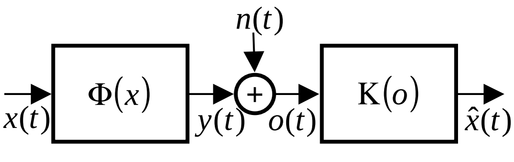

Figure 3: Signal model of the reconstruction process.

The mathematical analysis of such a model is difficult, because the nonlinear equations cannot be analytically solved. At a given working point, with small alterations of $x(t)$, $o(t)$ can be approximated by the first two elements of the Taylor polynomial, which will be a linear approximation:

$$o = o_0 + \Delta o \approx \Phi(x_0) + \left.\frac{d\Phi(x)}{dx}\right|_{x=x_0} \cdot \Delta x \,. \tag{8}$$

So the alteration of $o(t)$ is:

$$\Delta o(t) \approx \frac{d\Phi(x(t))}{dx(t)} \cdot \Delta x(t) \,. \tag{9}$$

The reconstruction process for the alteration is shown in Fig. 4:

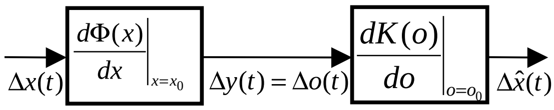

Figure 4: Signal model of the reconstruction process for small signal changes.

This model is already linear and can be applied in each point of the original characteristics. The noise in this model is not additive, but affects the working point, hence it changes the amplification of the second module. If the amplification of the second module is not the reciprocal of the first module, it causes differences between Δx and $\Delta \hat{x}$.

If the exact inverse of the original nonlinearity is used for reconstruction, the noise may be highly amplified. The noise amplification in case of small noise amplitudes is revealed, when the nonlinearity is described by Taylor polynomials:

$$\hat{x}_0 = \Phi^{-1}\left(y_0 + n\right) \approx x_0 + \frac{1}{\left.\frac{d\Phi(x)}{dx}\right|_{x=x_0}} \cdot n \tag{10}$$

The noise will be amplified, if the derivative of the nonlinear function $\Phi(x)$ is smaller than unity.

2.2. Regularized compensation

To optimize the model, first a measure for the quality of the estimate has to be defined. We define the best solution as the minimum of the following equation:

$$Cost = \left(\left\| \Delta x - \Delta \hat{x} \right\|\right), \tag{11}$$

where $\|a\|$ is the norm of a.The solution of this error criterion is not applicable directly, because it requires the knowledge of Δx. A solution to such an ill-posed problem was originally proposed by Tikhonov [8], who created a method to solve ill-posed integral equations with regularization operators. The error criterion used in Tikhonov's method is an extension of the output error criterion. One possible form is:

$$Cost = \left(\|\Delta o - \Delta \hat{o}\| + \lambda \|\Delta \hat{x}\| \right), \tag{12}$$

where $\hat{o}$ is the estimated output, computed from $\Delta\hat{x}$, which is led through the copy of the first module in Fig. 4. λ is a regularization parameter that reduces the errors caused by noise. λ cannot be too high, because it produces a distortion in the estimation. Its optimal value has to be found. In practice the l_2 norm is used, which minimizes the energy of the error. In the case of sampled signals this takes the form:

$$Cost = \left(\sum_i (\Delta o_i - \Delta \hat{o}_i)^2 + \lambda \sum_i \Delta \hat{x}_i^2 \right) \tag{13}$$

The minimum of this equation is at the point where $\frac{\partial}{\partial \Delta \hat{x}_i} = 0$. In this case the solution for $\Delta\hat{x}/\Delta o$ is:

$$\frac{\Delta \hat{x}}{\Delta o} = \left. \frac{d\mathrm{K}(o)}{do} \right|_{o=o_0} = \frac{\left. \frac{d\Phi(\hat{x})}{d\hat{x}} \right|_{\hat{x}=\hat{x}_0}}{\left(\left. \frac{d\Phi(\hat{x})}{d\hat{x}} \right|_{\hat{x}=\hat{x}_0} \right)^2 + \lambda} \tag{14}$$

The regularized inverse characteristics, $\mathrm{K}(o)$, can be calculated by the integration of (14). The integration constant in this case is not important, because the DC component does not affect the sound quality. The resulted characteristics can be used as $\mathrm{K}(o)$, without any further iteration. Thus, the reconstruction itself is a one-step process.

2.3. Determining the optimal regularization parameter

The optimal value of λ has to be determined to design the optimal shape of the regularized inverse characteristics. The optimal value of λ depends on the input signal, on the noise and on the shape of the original nonlinear function.

If the input signal, $x(t)$, is constant, and if the probability density-function of the output noise, $P_n(\nu)$, is known, the norm of the difference between the original signal and the estimate can be written in the following form:

$$\begin{aligned} E\{e(x,\lambda)\} &= E\{\|\hat{x}(\lambda) - x\|\} = \\ &= E\{\|\mathrm{K}(o,\lambda) - x\|\} = \\ &= \int_{-\infty}^{\infty} P_n(\nu) \cdot \|\mathrm{K}((\Phi(x)+\nu),\lambda) - x\| d\nu \end{aligned} , \tag{15}$$

where $E\{\ \}$ denotes the expected value and $e(x,\lambda)$ refers to the difference of the original and estimated signal. Now, the expected error value can be computed for all given $x(t)$ values. If the probability density function, $P_x(\chi)$, of $x(t)$ is known, the expected error after restoration can be written as

$$E\{\varepsilon(\lambda)\} = \int_{-\infty}^{\infty} P_x(\chi) \cdot E\{e(x,\lambda)\} d\chi . \tag{16}$$

The optimal value of λ can be found by minimizing (16). In practice, $P_n(\nu)$ is estimated from that signal parts of the recording, where only noise is present.

$P_x(\chi)$ can be estimated iteratively. As the first step, $P_x(\chi)$ is approximated by the probability density-function, $P_o(o)$, of the observed signal. Now, λ_1 and $\hat{x}_1$ can be estimated. From $\hat{x}_1$, a more exact estimate for $P_x(\chi)$ can be obtained. In our experiments, 3 iterations were enough to estimate a proper λ value.

3. SIMULATION EXAMPLE

To show the capabilities of the proposed algorithms, a multisinusoidal input signal consisting of four sinusoids was generated. This can be seen in Fig. 5a. This signal was distorted by a Gaussian error-function, which is similar to the density function of films:

$$y(t) = 0.5 \cdot erf\left(x(t) + 0.5\right) - 0.5 . \tag{17}$$

A Gaussian white noise was added to the signal to achieve 35 dB signal-to-noise ratio. The distorted, noisy signal can be seen in Fig. 5b.

The offset and amplification parameters of the distortion function were estimated with the proposed method, knowing that the shape of the nonlinear function is the Gaussian error-function. The estimated parameters can be seen in Table I:

Table I: Estimated parameters of the nonlinear function.

Number of sines	Cost	G1 (true=0.5)	O1 (true= - 0.5)	O2 (true=0.5)
1	20.15800	0.3175	-03913	1.6899
2	4.41796	0.4666	-0.5411	5.9754
3	0.49038	0.5012	-0.5220	0.7728
4	0.05798	0.4991	-0.4996	0.4985
5	0.05793	0.4990	-0.4993	0.4964
6	0.05795	0.4989	-0.4990	0.4929
7	0.05793	0.4990	-0.4994	0.4968
8	0.05786	0.4988	-0.4988	0.4921

As it can be seen, the estimates for more than three sinusoid signals are acceptable and the overmodeling did not affect the accuracy of the estimated parameters.

The probability density-characteristics of the original signal is assumed to be not known, and the optimal inverse characteristics was calculated by the method proposed in Section 2. Here, 5 iterations were made and the resulted value was compared with the real optimal value, computed from the density-characteristics of the original signal. The resulted λ values and the squared sum of the difference of the original and the estimated signal can be seen in Table II:

Table II: Regularization and error parameters of different regularized characteristics.

	Lambda	Error
Underregularized	$1 \cdot 10^{-10}$	206,087
Optimal	$3.036 \cdot 10^{-4}$	6,388
Estimated	$1.272 \cdot 10^{-4}$	7,499
Overregularized	$1 \cdot 10^{-1}$	191,345

The difference between the optimal and estimated error is relatively small.
The regularized inverses for the underregularized, optimal, estimated and overregularized cases can be seen in Fig. 6. Fig. 7 shows the signal estimates. The underregularized inverse has large errors, due to the noise amplification. The overregularized inverse has strong distortion. In the case of the optimal and estimated characteristics, both the effects of distortion and noise are small. The errors of the noise can be further reduced by additional noise filtering.

4. CONCLUSIONS

In this paper a novel method was proposed for the restoration of nonlinearly distorted movie soundtracks. The shape of the density curve is assumed to be known. The parameters of the density function are estimated from a periodic part of the distorted signal. The distorted signal of the soundtrack is compared to a multisinusoidal signal distorted by the model of the nonlinear function. The parameters are set in order to obtain the best match of these two signals in the least squares sense. In order to avoid noise amplification during restoration, a new method was proposed, which is based on Tikhonov regularization. A novel technique was presented for determining the optimal regularization parameter. The efficiency of the described method was shown on a simulation example.

ACKNOWLEDGEMENT

We would like to thank the help of the Hungarian National Film Archive. The support of the Hungarian Scientific Research Fund (OTKA F-034900 and F-035060) and Research and Developement Fund for Higher Education (FKFP-0098/2001) are also acknowledged.

REFERENCES

[1] Gil M. Raz and Barry D. Van Veen, "Blind Equalization and Identification of Nonlinear and IIR Systems - A Least Squares Approach", IEEE Trans. Signal Processing, vol. 48. pp. 192-200, Jan 2000.

[2] Robert D. Nowak and Barry D. Van Veen, "Volterra Filter Equalization: A Fixed Point Approach", IEEE Trans. Signal Processing, vol. 45. pp. 377-387, Febr 1997.

[3] Wolfgang Klippel, "Compensation for Nonlinear Distortion of Horn Loudspeakers by Digital Signal Processing", AES: Journal of the Audio Engineering Society, vol. 44, No 6, pp. 470-480, June 1996.

[4] Paul T. Troughton and Simon J. Godsill," Restoration of Nonlinearly Distorted Audio using Markov Chain Monte Carlo Methods", Presented at the 104th Convention of the Audio Engineering Society, preprint 4679, Amsterdam, May 1998

[5] S. Kerry Wilson and Patrick Delay, "A Method to Improve Cathode Ray Oscilloscope Accuracy", IEEE Trans. Instrumentation and Measurement, vol. 43, No 3, pp.483-486, June 1994.

[6] Stanley A. White, "Restoration of Nonlinearly Distorted Audio by Histogram Equalization", J. Audio Eng. Soc., vol. 30, No 11, pp. 828-832, November 1982.

[7] D. Preis and H. Polchlopek, "Restoration of Nonlinearly Distorted Magnetic Recordings", J. Audio Eng. Soc., vol. 32, No 1/2, pp. 26-30, January/February 1984.

[8] Tikhonov, A. N. and Arsenin, V. Y. "Solutions of ill-posed problems", New York, John Wiley & Sons, Inc. 1977.

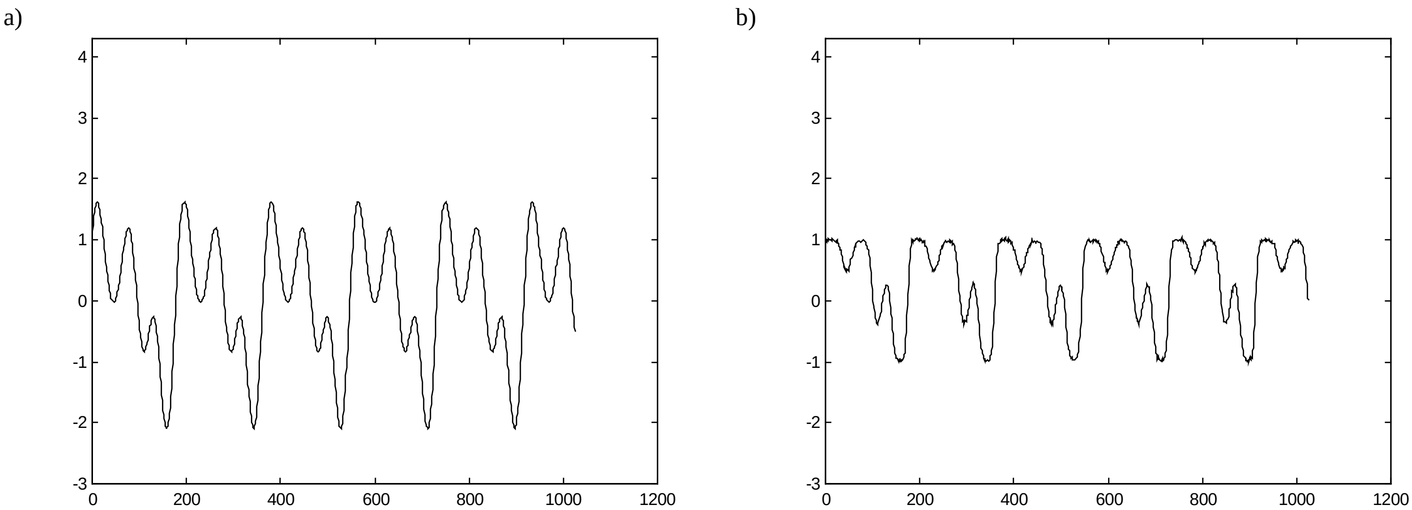

Figure 5: Input signal (a) and the nonlinearly distorted, noisy signal (b).

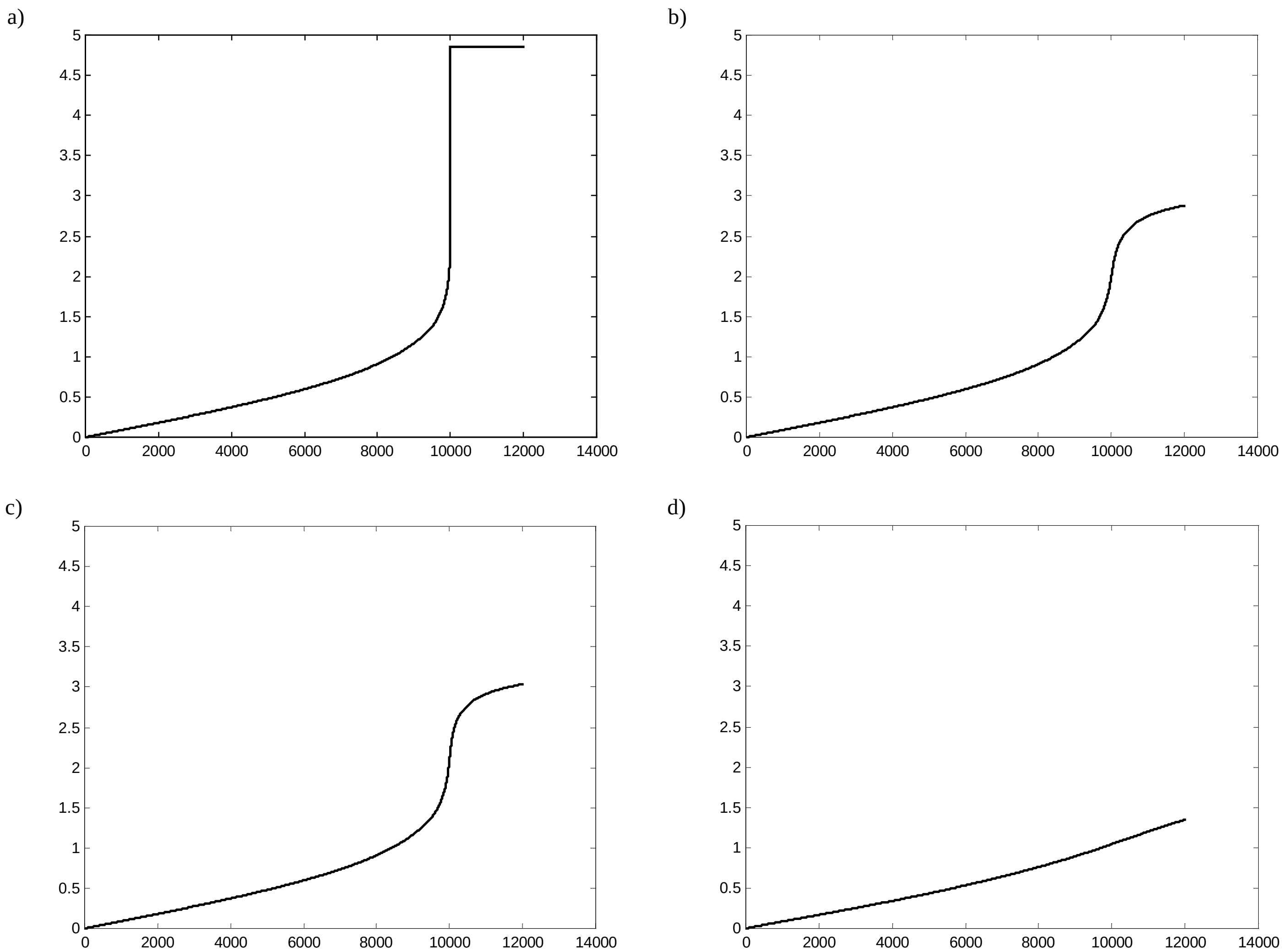

Figure 6: Underregularized (a), optimal regularized (b), estimated regularized (c) and overregularized characteristics (d).

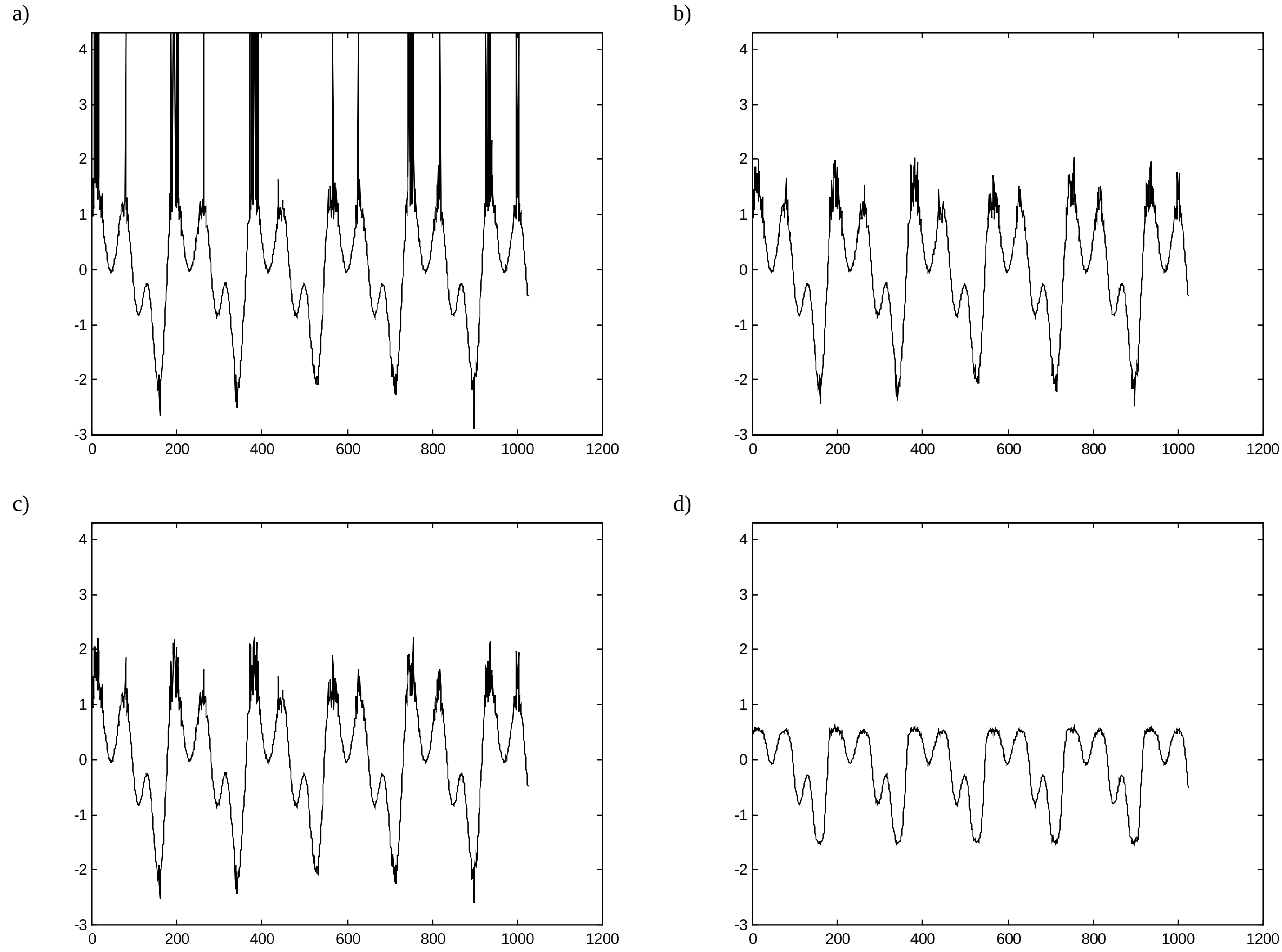

Fig. 7 Reconstructed signals: with underregularized (a), with optimal regularized (b), estimated regularized (c) and overregularuzed characteristcs (d).