Suboptimal Filtering and Nonlinear Time Scale Transformation for the Analysis of Multiexponential Decays

Jordi Palacín, Santiago Marco, and Josep Samitier

Abstract—**Multiexponential decays may contain time-constants differing in several orders of magnitudes. In such cases, uniform sampling results in very long records featuring a high degree of oversampling at the final part of the transient. Here, we analyze a nonlinear time scale transformation to reduce the total number of samples with minimum signal distortion, achieving an important reduction of the computational cost of subsequent analyses. We propose a time-varying filter whose length is optimized for minimum mean square error.**

Index Terms—**Data reduction, Gardner transform, multiexponential decay, noise, nonuniform sampling, transients.**

I. INTRODUCTION

THE ANALYSIS of multiexponential decays appears in many areas of scientific or technical interest: analysis of deep levels in semiconductors [1], thermal responses [2], fluorescence spectroscopy [3], etc. A recent review confirms the impact of this family of signals and reviews the variety of analysis methods [4]. In many cases, the transient signal under study may contain time-constants differing in several orders of magnitudes. In some applications such as deep level spectroscopy or thermal analysis, as many as eight time decades are recorded. With linear sampling, the transient would contain 10^8 points. Such a huge number of points represent an incredible computational burden, and they are the consequence of a high degree of oversampling in the final time decades. Decaying signals are in fact nonstationary, and their bandwidth decreases with time. In qualitative terms, the "instantaneous bandwidth" of a deterministic signal is related to its derivatives. It is commonly accepted that in the regions of maximum curvature the required bandwidth is maximum. However, for exponential decays all derivatives decay with time, and so also the required bandwidth. In consequence, it is not necessary to keep a very high sampling rate to capture the fastest time-constants occurring at the initial part of the transient to analyze the slowest time-constant that appears at the final part.
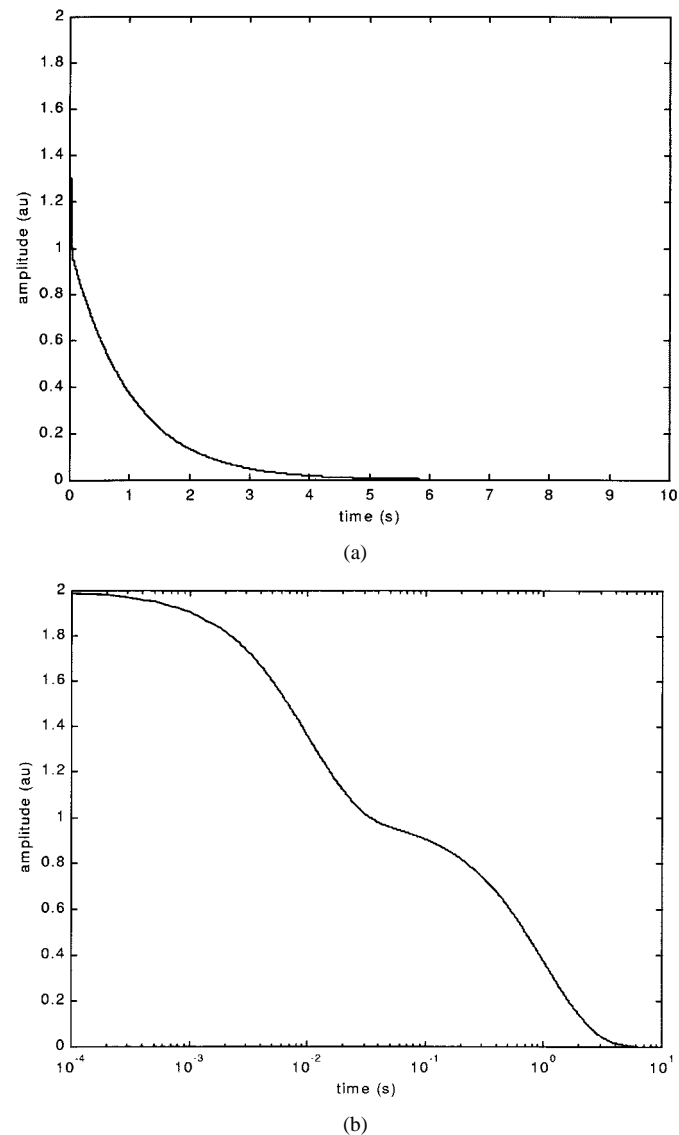

Fig. 1. Multiexponential decay. $x(t) = \exp(-t/0.01) + \exp(-t/1.0)$: (a) linear time scale and (b) logarithmic time scale.

J. Palacín is with the Laboratory d'Instrumentació i Comunicacions, Departament d'Electrònica, Universitat de Barcelona, 08028 Barcelona, Spain, and also with the Departament d'Informàtica i Enginyeria Industrial, Universitat de Lleida, 25001 Lleida, Spain.

S. Marco and J. Samitier are with the Laboratory d'Instrumentació i Comunicacions, Departament d'Electrònica, Universitat de Barcelona, 08028 Barcelona, Spain (e-mail: santi@el.ub.es).

In different application areas, researchers usually prefer a logarithmic representation (Fig. 1) of the time scale which shows up more clearly the presence of fast time constants which appear hidden in a linear time scale plot. In fact, several algorithms, based on the work of Gardner *et al.* [5], have been proposed which directly operate in this transformed time scale [4], [6]. This is because after the time transformation every exponential

term presents the same signal shape whatever the time constant is. In other words, an increase (decrease) in the time-constant only represents a right (left) shift. In other words, the transformed signal can be expressed as a convolution [5]. The authors have already analyzed in previous works how to take advantage of this invariance [7].

Our purpose in this work is to analyze how to perform this time scale transformation from linear sampling in the time scale to equidistant samples in the logarithm of time for minimum signal distortion in the presence of additive white noise.

II. PROCEDURE

As already mentioned in the previous section, our aim is to carry out a data reduction algorithm that maps the linear time scale to a logarithmic time scale. We will consider first the effect of such transformation in a single exponential

$$x(t) = Ae^{-(t/\tau)}$$
$$z = \ln(t); \qquad \zeta = \ln(\tau)$$
$$x(z) = A\exp\left(-e^{z-\zeta}\right). \tag{1}$$

A point to discuss concerns the adequate sampling rate in the transformed time domain. In the following we will talk about pseudo-frequency and pseudo-spectrum because the independent variable is no longer time, but the logarithm of time. It can be seen that the pseudo-frequency spectrum of an exponential in the transformed domain is clearly low pass. It can be calculated analytically, and it is related to the Gamma function evaluated along the imaginary axis

$$|\Im(g(z)))| = |\Gamma(j\omega)| = \sqrt{\frac{\pi/\omega}{\sinh(\pi\omega)}}. \tag{2}$$

It may be observed that the pseudo-spectrum amplitude does not depend on the time constant value, because a change in the time constant only represents a shift in position in the transformed domain. A few points per decade are necessary to adequately represent the signal, although the algorithm presented later on permits the calculation of the signal in the transformed domain for a large range of sampling rates (in the transformed domain). Our experience indicates that sampling rates exceeding ten to fifteen points per decade produce high-frequency pseudo-spectral bands where the noise content is more intense that the signal power. However, a rigorous analysis of this problem will be published elsewhere [8].

New samples are equidistant in the logarithm of time

$$t_n = t_0 10^{n_N} \tag{3}$$

where t_0 is the initial time, and N is the number of points per decade. Note that to represent eight time decades we will need only 80–120 points as compared with 10^8.

This "decimation" in time must be accompanied by low-pass filtering to avoid aliasing of the superimposed noise that has, in general, a much wider spectrum. Note that this process resembles, but it is not the same as, the usual sampling rate conversion by a rational factor for which a complete theory is available [9]. In addition, the actual time points $\{t_n\}$ do not coincide with the original sampling times, so we must necessarily interpolate. Thus, the procedure will consist of: first, filtering for the minimum square error followed by linear interpolation to calculate the signal value at the points t_n. Note that the new time instants t_n do not depend on the filtering method chosen.

Concerning the filtering step, we will propose three different approaches. The first one is an heuristic method, and it is considered as the baseline algorithm. The second is based in an optimum filter introduced by Papoulis [10]. The third procedure is a time-varying filter, the length of which is optimized for minimum mean square error.

A. Heuristic Method

The heuristic method is as follows: new time points are generated in the logarithmic time scale. Typically the sampling rate in this transformed domain is chosen to be between 10 and 20 points per decade. In the computation of each new sample $x(t_n)$ we use all original samples (P samples) within the interval $(t_n - \Delta/2, t_n + \Delta/2)$ where $\Delta = t_{n+1} - t_n$. Moreover, the filter length N increases linearly with time, and consequently the cutoff frequency decreases linearly

$$N = \frac{2}{5}k \tag{4}$$

where "k" is the sample index in the linear domain. To calculate the new sample we use polynomial linear or quadratic interpolators, obtained after least squares fitting of the points within the interval to a straight line or parabola, respectively. This polynomial filtering can be considered as a generalization of the Savitzky–Golay SG-FIR filters [11], which provide linear phase, and maximum noise rejection ratio subjected to the condition of frequency flatness at DC. The main difference is: since the new sample does not coincide with the original, the filter coefficients may vary slightly depending on the parameter ξ, $0 < \xi < 1$ which represents the fractional position of the new sample within the original time sampling period T. In this way, we can write the filter operation as

$$yf(k+\xi) = \sum_j b_j(\xi, M)x(k-j) \tag{5}$$

with $t_n = (k+\xi)T$ and, $t_n - \Delta/2 < (k-j)T < t_n + \Delta/2$, and b_j the filter coefficients that are linear or parabolic functions of ξ if we are considering first- or second-order Savitzky–Golay filters, respectively. Their closed forms are not written here due to space constraints.

The first consequence of the above procedure is that the noise signal becomes nonstationary and its variance decreases with time. Very important noise reduction is achieved in the final time-decades because of the very low cutoff frequency of the interpolating filter, which is a consequence of the huge number of points within the interval. From this point of view the process is very advantageous because it provides simultaneously significant data reduction and reduces the noise content.

However, at this point we must consider the signal distortion that can arise after this time-varying filtering and "decimation."
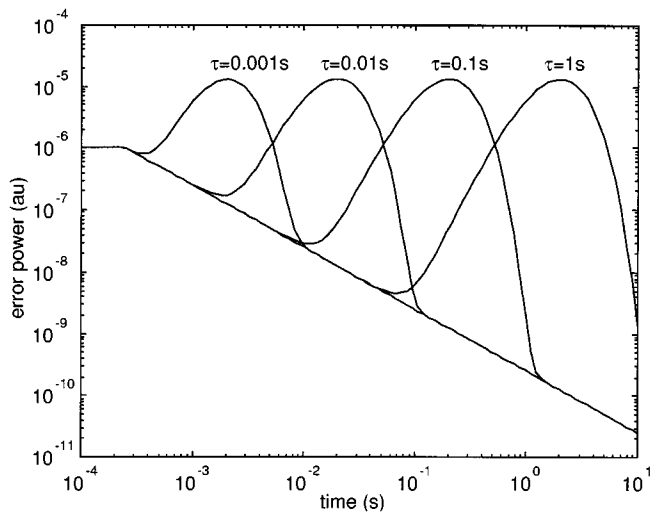
Fig. 2. Estimated influence of the exponential time-constant location ($\tau = 0.001$ s, 0.01 s, 0.1 s, 1 s) in the error power: Noise power $= 10^{-6}$.
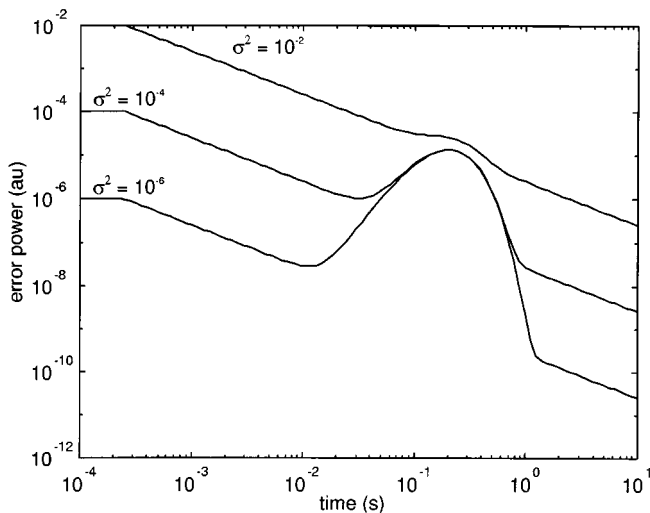

Fig. 4. Time constant spectrum: exact (o) and estimated ($*$) choosing three exponentials per decade. $x(t) = 1 \cdot \exp(-t/0.008) + 0.5 \cdot \exp(-t/0.03)$, $\sigma^2 = 10^{-4}$, $F_s = 10$ kHz.

that it depends on the noise content and on the second derivative of the noise signal

$$yf(n) = \frac{1}{N} \sum_{k=-(N-1)/2}^{k=(N-1)/2} x(n-k)$$

$$N(t) = \frac{1}{t_0} \sqrt[5]{\frac{9\sigma^2}{2f''(t)^2}} \qquad (6)$$

where

$x(n)$	original signal;
$yf(n)$	filtered signal;
N	optimal filter length;
t_0	linear sampling period;
σ^2	noise variance;
$f''(t)$	second derivative of the underlying noise-free transient.

Fig. 3. Estimated influence of the noise power ($\sigma^2 = 10^{-6}$, 10^{-4}, 10^{-2}) in the error power for a fixed exponential time constant: $\tau = 0.1$ s.

As real exponentials are not bandlimited signals, any filtering may induce some degree of signal distortion and, in fact, this is the case. Figs. 2 and 3 show the mean square error which appears when filtering a single exponential decay with the first-order modified S–G filter according to the above procedure. In Fig. 2, we observe the distortion depending on the time constant value: maximum distortion appears in the vicinity of the time constant value. In Fig. 3, we may observe that this distortion is more apparent for low noise levels.

B. Papoulis Filter

Up to now, we have seen the limitations imposed by the first procedure. To introduce our second proposal, we wonder if the filter length evolution along the transient is optimal. Papoulis [10] has obtained the optimal filter length when filtering a deterministic signal embedded in white noise, and he has found
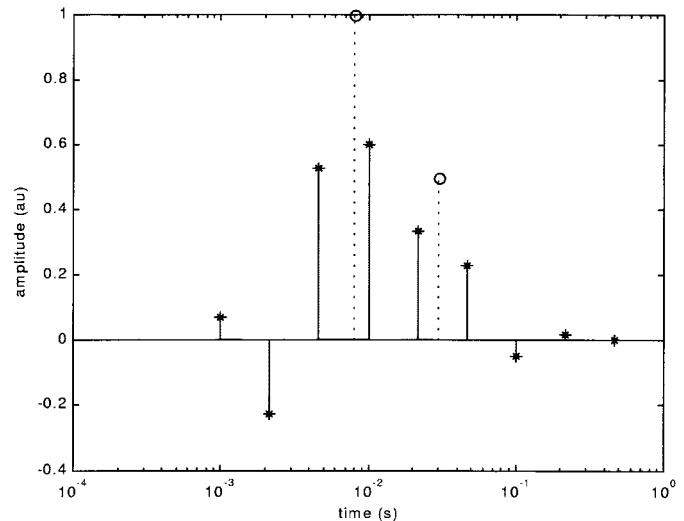
The practical application of this method in our case has a problem: the second derivative has to be estimated. It is well known that differentiation of noisy signals is a risky signal processing operation [12]–[14]. A variety of methods have been proposed in the literature. Among them we have obtained the best results by double differentiation of a signal model obtained by fitting an exponential series [3] to our original data (ESM). Such series contain three exponentials per decade at fixed logarithmically distributed times. Only the amplitudes are estimated by linear least squares. The obtained fitting is very good, and the residuals are white. Fig. 4 shows the goodness of the fitting, and Fig. 5 compares the real and estimated amplitudes for a biexponential decay. Despite the very good time response, we still suffer from unacceptable errors in the estimation of the second derivative at the initial part of the transient (Fig. 6). These errors are undoubtedly related to the inadequacy of the model used. Poor estimation of the second derivative will lead to erroneous filter lengths that will provide either distortion or insufficient noise rejection. Because of that
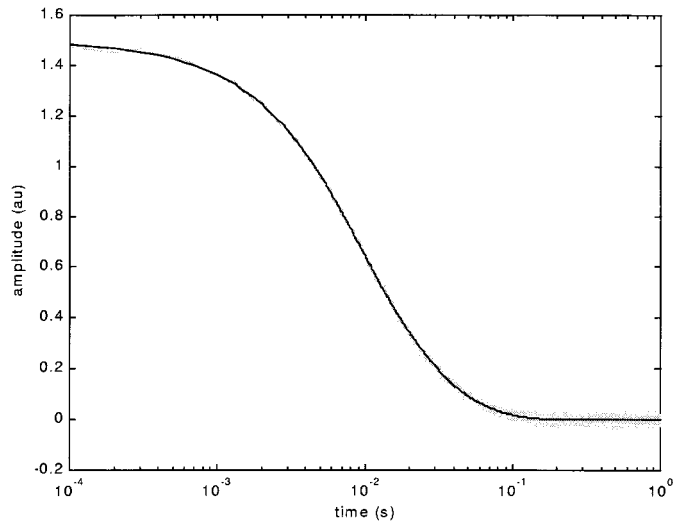
Fig. 5. Signal approximation by the ESM (hard line) and original noise corrupted signal (soft line). $x(t) = 1 \cdot \exp(-t/0.008) + 0.5 \cdot \exp(-t/0.03)$, $\sigma^2 = 10^{-4}$, $F_s = 10$ kHz.
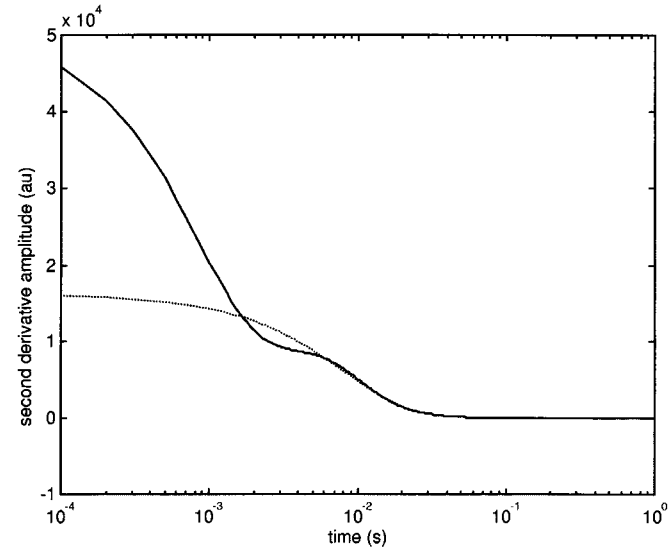

Fig. 6. Exact second derivative of the original signal (dot line), and the second derivative obtained with the signal model (solid line). $x(t) = 1 \cdot \exp(-t/0.008) + 0.5 \cdot \exp(-t/0.03), \sigma^2 = 10^{-4}$, $F_s = 10$ kHz.

we have devised a method which avoids the calculation of the instantaneous second derivative of the transient.

Even when the second derivative is known, (6) is only valid when $N \cdot T_s \leq \tau$ (see the original demonstration for further details [10]), where T_s is the sampling period and τ is one of the time constants of the analyzed transient.

C. Suboptimal Method

In this method, we carry out first a varying time filtering in the linear domain. Transformation to the logarithmic scale is performed after filtering by a simple linear interpolator.

The error can be theoretically evaluated with

$$y[n] = x[n] + w[n]$$
$$yf[n] = h[n] \otimes (x[n] + w[n])$$
$$e[n] = yf[n] - x[n] \tag{7}$$

where $w[n]$ is the noise content, and $x[n]$ is the noise-free multiexponential decay. We suppose now that a multiexponential model of the signal is available, which can be obtained as it has been previously outlined

$$x[n] = \sum_{i=1}^{M} A_i a_i{}^n u[n] \tag{8}$$

where

$h[n]$ filter impulse response;
A_i amplitudes;
a_i $\exp(-T_s/\tau_i)$.

Then, the error power can be estimated for a general case. Note that we consider a centered filter for zero group delay

$$E\left[e^2[n]\right] = \left[\sum_{i=1}^{M} A_i a_i{}^n \left(H_N(a_i) - 1\right)\right]^2 + \sum_{k=-((N-1)/2)}^{(N-1)/2} h^2(k; N)\sigma^2 \tag{9}$$

where

$$H_N(a_i) = \sum_{k=-((N-1)/2)}^{(N-1)/2} h(k; N)a_i{}^{-k}. \tag{10}$$

Expression (9) shows that the error has two contributions: distortion and filtered noise.

Also, when an estimation of the noise and the exponentials contained in the signal is available, (9) can be used to find the optimum length N of the filter in each point to achieve the minimum error power in the filtered signal. We have minimized the above expression in the case of a moving average filter (length N, filter coefficients $1/N$)

$$E\left[e^2[n]\right]$$
$$= \sum_{i=1}^{M} A_i{}^2 a_i{}^{2n} \left(\frac{1}{N} \cdot \frac{a_i{}^{-(1/2)}\left(a_i{}^{N/2} - a_i{}^{-(N/2)}\right)}{1 - a_i{}^{-1}} - 1\right)^2$$
$$+ 2\sum_{i \neq j} A_i A_j a_i{}^n a_j{}^n$$
$$\cdot \left(\frac{1}{N} \cdot \frac{a_i{}^{-(1/2)}\left(a_i{}^{N/2} - a_i{}^{-(N/2)}\right)}{1 - a_i{}^{-1}} - 1\right)$$
$$\cdot \left(\frac{1}{N} \cdot \frac{a_j{}^{-(1/2)}\left(a_j{}^{N/2} - a_j{}^{-(N/2)}\right)}{1 - a_j{}^{-1}} - 1\right) + \frac{\sigma^2}{N}. \tag{11}$$

An algorithm based on golden section search and parabolic interpolation [15] is used to find the local minima corresponding to the optimum filter length N. The computational burden of this iterative minimization is highly reduced if the last value found is provided as the initial value in the next time step.

This kind of filter provides the maximum noise rejection ratio. Moreover for the large filter lengths obtained at the end of the transient, it is not practical to try to optimize the filter coefficients, because huge matrices should be inverted. Note that the optimum length depends on the sample index "n," so we obtain again a time-varying filter as expected due to the signal characteristics. The proposed filtering method is optimal,
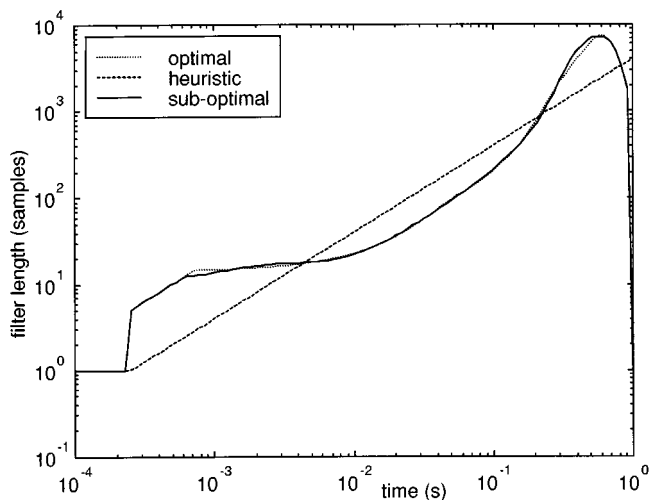
Fig. 7. Comparison between the optimal filter length (dot line), the suboptimal (solid line), and the heuristic method (dashed line). $x(t) = 1 \cdot \exp(-t/0.008) + 0.5 \cdot \exp(-t/0.03)$, $\sigma^2 = 10^{-4}$, $F_s = 10$ kHz.
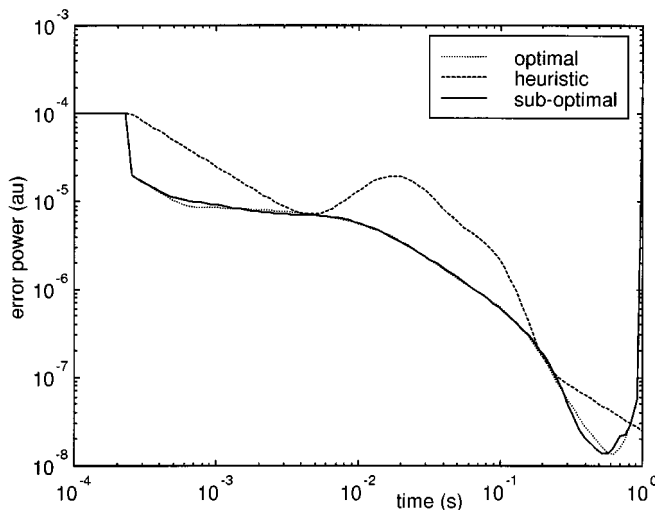

Fig. 8. Error power of the optimal (dot line), the suboptimal (solid line), and the heuristic method (dashed line). $x(t) = 1 \cdot \exp(-t/0.008) + 0.5 \cdot \exp(-t/0.03)$, $\sigma^2 = 10^{-4}$, $F_s = 10$ kHz.

within the class of uniform moving average filters of variable length, when the exact signal model of the transient is available (labeled optimal filtering in the figures). It is only suboptimal when an estimated model of the transient is used.

Finally, we would like to remark that due to the final linear interpolation step, it is not necessary to filter the complete transient but only those samples needed for the final interpolation. This reduces drastically the number of floating point operations needed to complete the algorithm.

D. Discussion

Results in Figs. 7 and 8 show the evolution of the filter length and the mean square error with time. The heuristic, the optimal and the suboptimal methods are compared. The estimated signal model is obtained as it has been previously outlined (Exponential Series Method) for the second derivative estimation. It can

be seen that, despite the signal model parameters being clearly far from reality, the predicted filter length with the estimated model (*suboptimal*) is almost identical to the exact result (*optimal*). This very good result is obtained with the same signal model that produced problems in the second derivative estimation. Moreover, the results clearly outperform the mean square errors obtained with the heuristic method: the noise rejection is larger and distortion peaks are avoided. The large increase in the error at the final part of the transient is a border effect.

III. CONCLUSIONS

It is not practical to work with linear sampled multiexponential signals containing time-constants differing in several orders of magnitudes. A procedure has been proposed to achieve a dramatic data reduction and noise reduction based on a moving average filter, the length of which is optimized for the minimum mean square error. However, a previous estimation of a signal model and noise level is needed. This estimation is based on the linear least squares of an exponential series with three terms per decade. Despite the fact that the signal model parameters could be quite far from reality, very good time agreement is obtained and the filter length prediction is almost optimal, providing maximum noise rejection with minimum distortion.

REFERENCES

[1] J. Samitier, S. Marco, A. Pérez-Rodríguez, J. R. Morante, P. Boher, and M. Renaud, "Anomalous optical and electrical recovery process of photoquenched EL2 defect produced by oxygen and boron implantation in GaAs," *J. Appl. Phys.*, vol. 71, pp. 252–259, 1992.
[2] V. Szekely, "Identification of RC networks by deconvolution: Chances and limits," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 244–258, 1998.
[3] A. Siemiarzuk, B. D. Wagner, and W. R. Ware, "Comparison of the methods of maximum entropy and exponential series method for the recovery of distributions of lifetimes from fluorescence lifetime data," *J. Phys. Chem.*, vol. 31, pp. 1661–1666, 1990.
[4] A. A. Istratov and O. F. Vykenko, "Exponential analysis in physical phenomena," *Rev. Sci. Instrum.*, vol. 70, pp. 1233–1257, 1999.
[5] D. G. Gardner, J. C. Gardner, G. Lush, and W. W. Meinke, "Method for the analysis of multicomponent exponential decay curves," *J. Chem. Phys.*, vol. 31, pp. 978–986, 1959.
[6] S. Marco, J. Samitier, and J. R. Morante, "A novel time-domain method to analyze multicomponent exponential transients," *Meas. Sci. Technol.*, vol. 6, pp. 135–142, 1995.
[7] S. Marco, J. Palacín, A. Guibernau, and J. Samitier, "Improved multi-exponential transient spectroscopy by iterative deconvolution," in *Proc. IMTC'98*, St. Paul, MN, 1998, pp. 670–674.
[8] S. Marco, J. Palacín, and A. J. Samitier, "Improved multi-exponential transient spectroscopy by iterative deconvolution," *IEEE Trans. Instrum. Meas.*, submitted for publication.
[9] R. E. Crochiere and L. R. Rabiner, *Multi-Rate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
[10] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1986.
[11] P. Steffen, "On digital smoothing filters: A brief review of closed form solutions and two new filters approaches," *Circuits Syst. Signal Process.*, vol. 5, p. 187, 1986.
[12] A. Miekina and R. Z. Morawski, "Regularized differentiation of measurement data using *a priori* information of the signal and noise spectra," *IEEE Trans. Instrum. Meas.*, vol. 39, pp. 824–826, 1990.
[13] H. J. Woltring, "On optimal smoothing and derivative estimation from noisy displacement data in biomechanics," *Human. Mov. Sci.*, vol. 4, pp. 229–245, 1985.
[14] S. Fioretti and J. Letto, "Accurate derivative estimation from noisy data: A state-space approach," *Int. J. Syst. Sci.*, vol. 20, pp. 33–53, 1989.
[15] G. F. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice-Hall, 1976.

Jordi Palacín received the electronic engineering degree from the Universitat de Barcelona, Barcelona, Spain, in 1997, where he is currently pursuing the Ph.D. degree in the Departament d'Electrònica.

He is Associate Professor at the Departament d'Informàtica i Enginyeria Industrial, Universitat de Lleida, Spain. His research interest includes signal processing, sensors, and robotics.

Santiago Marco received the B.S. degree in physics from the Universitat de Barcelona, Barcelona, Spain, in 1988, and the Ph.D. (honor award) degree from the Departament de Física Aplicada i Electrònica, Universitat de Barcelona, for the development of a novel silicon sensor for *in vivo* measurements of the blood pressure, in 1993.

From 1989 to 1990, he was working in the electro-optical characterization of deep levels in GaAs. From 1990 to 1993, he was a Regular Visitor of the Centro Nacional de Microelectrònica. In 1994, he was Visiting Professor at the Universita di Roma "Tor Vergata" working in electronic noses. Since 1995, he has been Associate Professor at the Departament d'Electronica of Universitat de Barcelona. He has published about 35 papers in scientific journals and books. His current research interests are the modeling, simulation, and test of microsystems and signal processing for intelligent sensor systems.

Josep Samitier received the B.S. degree in physics from the Universitat de Barcelona, Barcelona, Spain, in 1982, and the Ph.D. degree from the Departament de Física Aplicada, Universitat de Barcelona, in 1986.

Since February 1995, he has been Full Professor at the University of Barcelona. From 1983 to 1984, he was a Research Fellow in the Applied Physics Department of the University of Barcelona in the field of GaAs MESFET devices and electro-optical characterization of $III-V$ semiconductors. From February 1984 to June 1985, he was Visiting Research Fellow at the Laboratoire D'Electronique Philips, LEP, Paris, France. In 1988, he was appointed as Associate Professor at the same university working in the development of microsystems and electronic instrumentation. His current research and developed projects concern the design, test, and signal conditioning of microsystems and the design of interface circuits for sensors and microsystems. Currently, he coordinates the projects in Electronic Instrumentation of the Electronic Engineering R&D activities.