USC–SIPI REPORT #253

Maximum A Posteriori Filter Estimation for Cauchy Data Using Lower-Order Statistics

by

Russell H. Lambert and Chrysostomos L. Nikias

March 1994

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

Maximum A Posteriori Filter Estimation for Cauchy Data Using Lower-Order Statistics [1]

Russell H. Lambert and Chrysostomos L. Nikias
Signal and Image Processing Institute
University of Southern California
3740 McClintock Ave. Room 400
Los Angeles, CA 90089

Abstract

A Maximum A Posteriori (MAP) estimation approach is used to determine the multipath channel parameters of a system driven with Cauchy data. This is a blind deconvolution problem with the input driving sequence having a Cauchy probability density function. The fact that the Cauchy distribution is not Bussgang, and has infinite variance, requires the development of a new blind deconvolution formulation which estimates the forward filter, instead of the inverse filter as in traditional deconvolution schemes.

[1] This work was supported by the Office of Naval Research under the AASERT program, contract No. N-00014-93-L-1041.

1 Introduction

The equalization of Cauchy data is addressed in this paper. An intuitive approach to the problem is presented, followed by the mathematical formulation of the Maximum A Posteriori method. The Cauchy distribution is a member of the alpha-stable distribution model which has recently been a topic of research. A tutorial on alpha-stable processes [1], desribes some of the new procedures needed in dealing with infinite variance data. Cauchy, with alpha equal to 1, has finite moments $E|x|^{\alpha}$, only with $\alpha < 1$. Meansquared error methods no longer apply. The Gaussian distribution is also a member of the alpha- stable family with $\alpha = 2$.

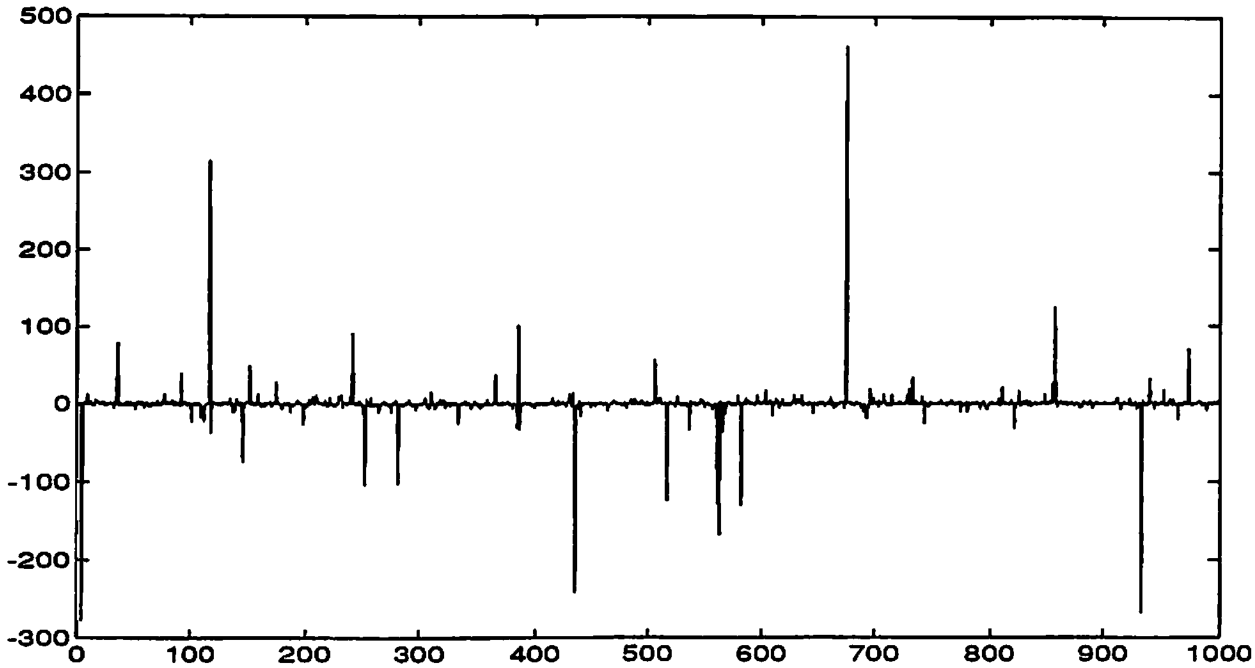
Traditional blind deconvolution schemes usually deal with finite variance data which have a special property which we shall describe. A Bussgang process has the property that the output of a non-linear operation cross-correlated with the input is proportional to the autocorrelation of the input. Bussgang methods for blind deconvolution use the fact that an equalizer can be adjusted using only the output of the equalizer passed through a memoryless nonlinearity. The deconvolution filter taps are the inverse channel which undo the effect of the multipath channel. Alpha-stable processes do not possess a finite autocorrelation and are therefore not amenable to Bussgang equalization. Therefore a completely new approach is required.

2 Intuitive Approach

Researchers in the area of statistics have considered this problem [3] [2]. Hsing in [2], suggested an intuitive approach to the problem. Consider the plot of a 1000 point sample of a Cauchy sequence in Fig. 1. Notice the "spikey" and implusive nature of the data. Actually one large impulse is all that is needed for determination of the impluse response (forward filter) of the system.

Fig. 1

1000 point Sample of Cauchy Data



An example of this type of filter estimation is in image processing where a narrow, intense bright spot and it's immediate surroundings can be taken as the scaled impulse response of the imaging system. The impulse response can then be inverted and used to deblur (or equalize), the image. Hsing, in [2] suggested using a similar idea to this by taking a small section of data located around a local maximum point, scaling it by the value of that local maximum point and calling that section of data an estimate of the filter.

This simple approach works well with Cauchy data. Cauchy tends to be populated with large spikes which are spaced far enough apart that the estimates of such a method can have low variance. As the alpha parameter of a stable process is adjusted lower and made further from Gaussian, the deconvolution process is expected to only improve. As the alpha term is made closer to Gaussian ($\alpha = 2$), the procedure is espected to degrade in performace until no true-phase equalization at all is possible at the Gaussian case.

3 Maximum A Posteriori Formulation

Given a received signal

$$r = a + n,$$

where a is the original data and n is convolutional noise, we can use their probability distributions to find the probability of receiving an r.

In the finite variance world, the convolutional noise would be modeled as Gaussian due to the Central Limit Theorem. In the inifinite varience case we must use the Generalized Central Limit Theorem (see [1]) which states that distributions with power law α, sum together to form another distribution with power law α. In other words, many Cauchy distributions added together will make another Cauchy with a larger scale parameter. Our convolutional noise is then a scaled Cauchy

$$p_n(n) = \frac{\sqrt{D}}{\pi(D + n^2)}.$$

Our data is modelled as Cauchy with unit scale,

$$p_a(a) = \frac{1.}{\pi(1 + a^2)}.$$

We wish to follow the method of Van Trees in [4], and form a likelihood function of a given r:

$$p_{a/r}(a/r) = p_{r/a}(r/a)p_a(a)$$

where

$$p_{r/a}(r/a) = p_n(r - a).$$

So

$$p_{a/r}(a/r) = \frac{\sqrt(D)}{\pi(D + (r - a)^2)} \frac{1.}{\pi(1 + a^2)}.$$

Taking the derivative and setting the result equal to zero yields:

$$0 = dp_{a/r}(a/r)/da = \frac{(r - a)}{(D + (r - a)^2)} - \frac{a}{(1 + a^2)}$$

$$0 = ar^2 - (1 + 3a^2)r + (a + 2a^3 + aD).$$

This equation is cubic in a and quadratic in r. Traditional equalizers at this point solve for a, leading to an inverse filter deconvolution method. In this Non-Bussgang formulation the received vector is the key (not the output of the equalizer). The input to the equalizer $r = a * h$, remains part of the equalizer adaptation. In Bussgang equalizers, the output $\hat{a} = r * \hat{h}^{-1}$ of the equalizer is used and the estimate of the inverse filter can be iteratively updated towards the true inverse. In this new case, the input to the equalizer r, is used, and the forward filter is estimated. Here r must be viewed as a series of scaled forward filters. We shall solve the quadratic equation in r and obtain:

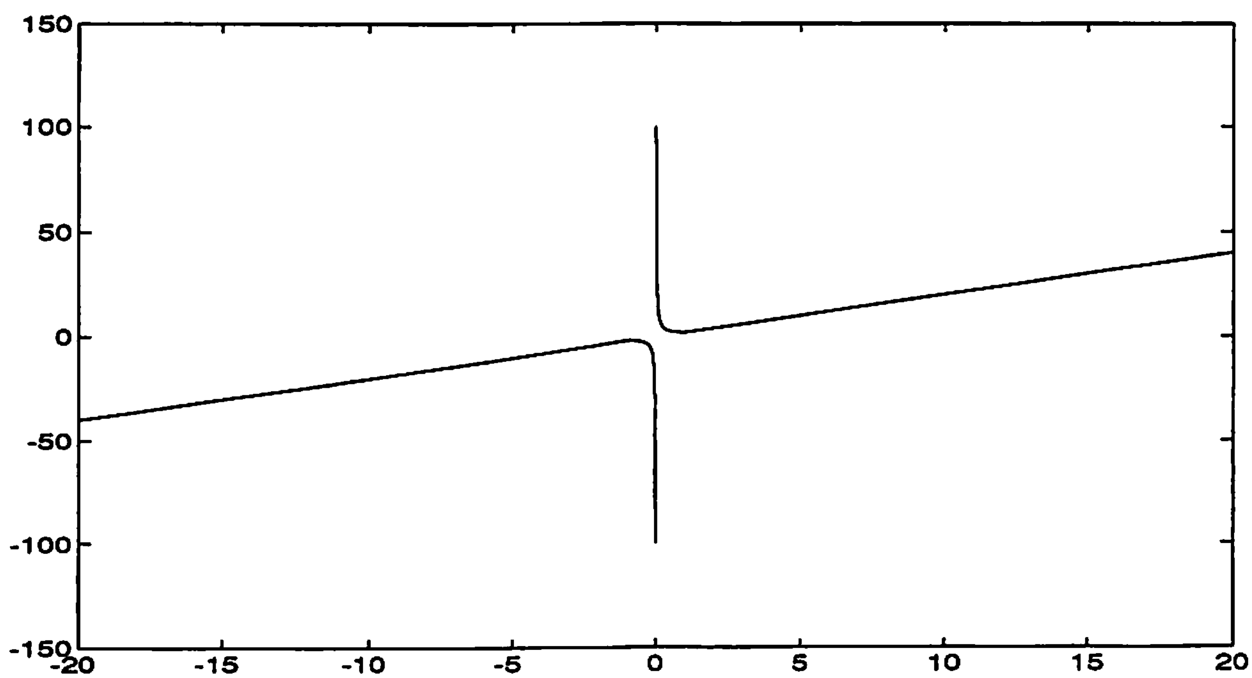
$$r = \frac{(1+3a^2) \pm \sqrt{(1+3a^2)^2 - 4a(a+2a^3+aD)}}{2a}$$

This equation with the posivtive sign is plotted in Fig. 2. Note the linearity above a threshold.

Fig. 2

Plot of $r = \frac{(1+3a^2)+\sqrt{(1+3a^2)^2-4a(a+2a^3+aD)}}{2a}$

X axis is a - Y axis is r



Note that the equation for r, becomes linear above a threshold. This is the desired result. This tells us that the forward filter updates should be weighted linearly with respect to the height of the locally maximum spike.

The term,

$$r = \frac{(1+3a^2) \pm \sqrt{(1+3a^2)^2 - 4a(a+2a^3+aD)}}{2a}$$

can be further analysed by looking only at the higher terms in a. Saying that if $a >> 1$, then $r \simeq \frac{(3\pm\sqrt{5})a}{2}$.

4 Simulations for Cauchy Distributed Data

4.1 Simulation 1

In using spikey Cauchy data the forward filter will be nearly perfectly estimated as soon as a large outlier occurs. This is shown is Fig. 3. The channel was modelled as MA(2) and coefficients where choosen as follows:

$$MA(z) = 1. + .8z^{-1} - .75z^{-2}.$$

In Fig. 3 the second coeffient is plotted every time a sufficient local maximum is found. The original Cauchy sequence is plotted in the lower plot. The criterion used here is that a point be above the value of 40. After a maximum is declared, the index must be moved at least a filter length before the next local maximum can be declared. A filter of order 10 was estimated. Although 10 is greatly overestimating the order, it does not matter, since zeros are found for the extra filter terms. The filter was initialized with

$$[1,0,0,0,0,0,0,0,0,0].$$

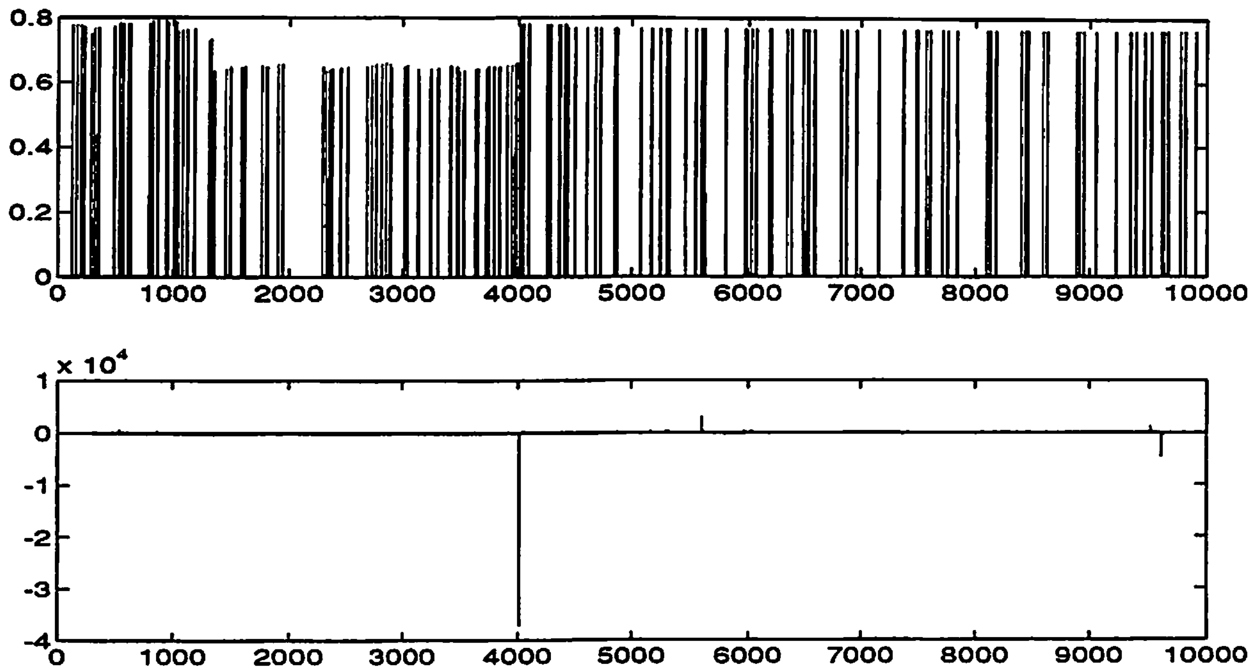
The final estimated forward filter using 10000 points was:

$$[1.0000, 0.7635, -0.7460, -0.0041, 0.0005, 0.0073, 0.0015, -0.0555, -0.0386, 0.0388].$$

Note that the convergence is almost instananeous. Convergence speed is in some sense determined by the average time of arrival of the first large spike.

Fig. 3

Plot of Simulation 1 Results: top plot is second filter coefficient (true value .8), lower plot is the original Cauchy data



4.2 Simulation 2

The channel was modelled as MA(1) and coefficients where choosen as follows:

$$MA(z) = 1. + .5z^{-1}.$$

In this case only 3000 points are used and a filter of order 3. The filter was initialized with

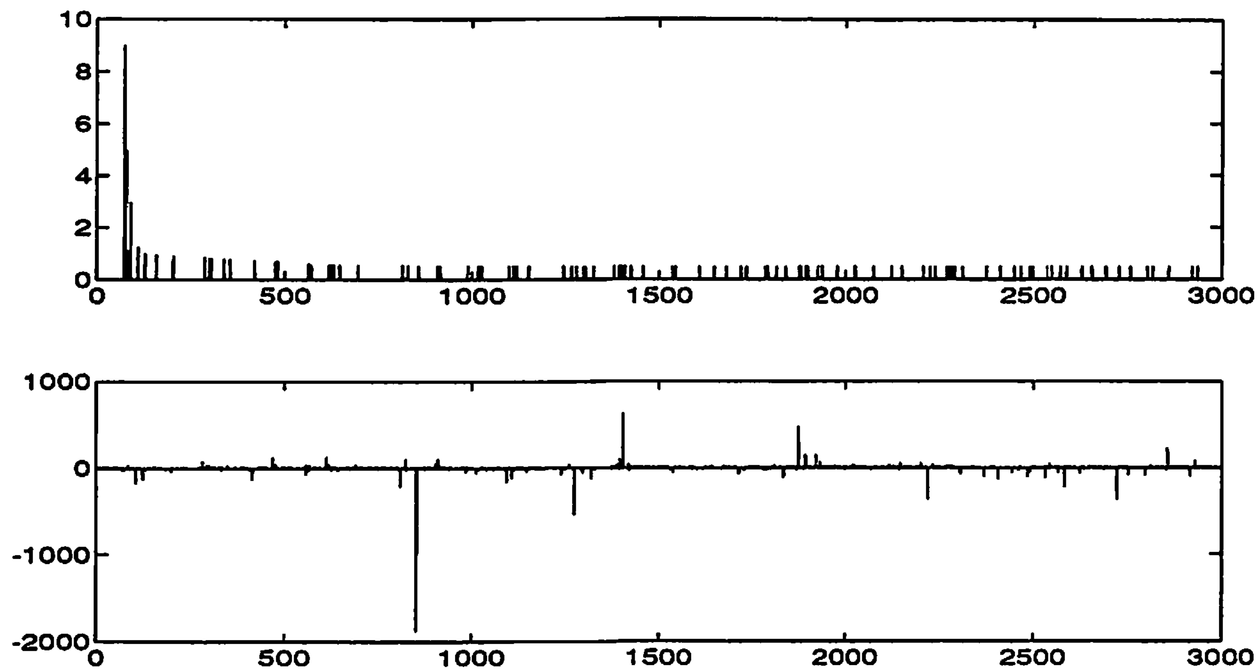
$$[1, 10, 0].$$

The final estimated forward filter using 3000 points was:

$$[1.0000, 0.4915, .03].$$

Fig. 4

Plot of Simulation 2 Results: top plot is second filter coefficient (true value .5), lower plot is the original Cauchy data



5 Can the MAP Method as Developed for Cauchy Data Work With Other Types of Infinite Varience Data?

One of the characteristics of the "forward filter" approach that was derived above, is that the method depends on the spikey nature of the data. The impulsive data reveals the channel as the data itself is a series of scaled impulse responses of the channel. We showed that a linear combination of all "scaled filters" which lie above a given threshold is the optimal estimation procedure for the Cauchy data case. Mixed-Phase channels do not present a problem.

Cauchy is not the only alpha-stable distribution which has a simple closed form expression. The $\alpha = .5$ distribution can be expressed:

$$p(x) = \frac{1}{\sqrt{2\pi}} x^{-3/2} exp(-1/2x).$$

Some preliminary calulations using this density show that the same technique, as developed for the Cauchy data case, will work with the $\alpha = .5$ case as well. The threshold above which the funciton for r is linear, is lower in the $\alpha = .5$ case. The $\alpha = .5$ case will also be able to handle longer filters because the spikes are spaced further apart. This leads us to suspect that perhaps the method will also perform with other stable distributions, even those with alpha closer to Gaussian (alpha = 2) than Cauchy (alpha = 1).

The simulations in this report will test this hypothesis. We shall test the limitations on the length of the channel which can be acurately estimated using our simple technique with various alpha-stable distributions.

6 Simulations to Test Feasibility of the MAP Method on Alpha-Stable Data with Various Values of Alpha

A filter of different orders is used in the tests. Orders 40, 15, 5 and 2 are tested. The filter coefficient where the true value is .8 is plotted. The filter is

$$[1, 0, 0, 0, ..., 0, 0, 0, .8].$$

Alpha-Stable Data was generated with location parameter $\beta = 0$ and $\alpha = .5, 1.0, 1.5, 1.8$. The threshold used determines how many estimates are averaged. The threshold used is noted and an attempt was made to keep it as low as possible while still getting reasonable results.

Fig. 5

1000 points; $\alpha = .5$ Sample of Data

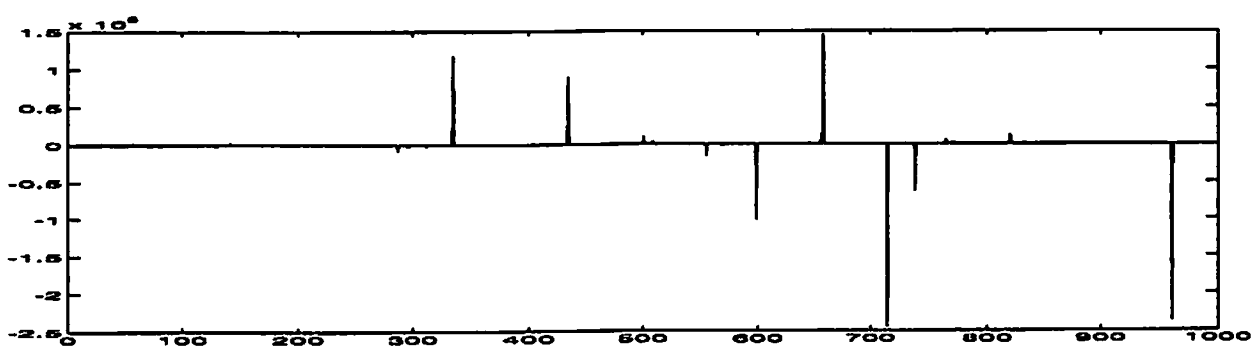


Fig. 6

1000 points; $\alpha = 1$ Sample of Data

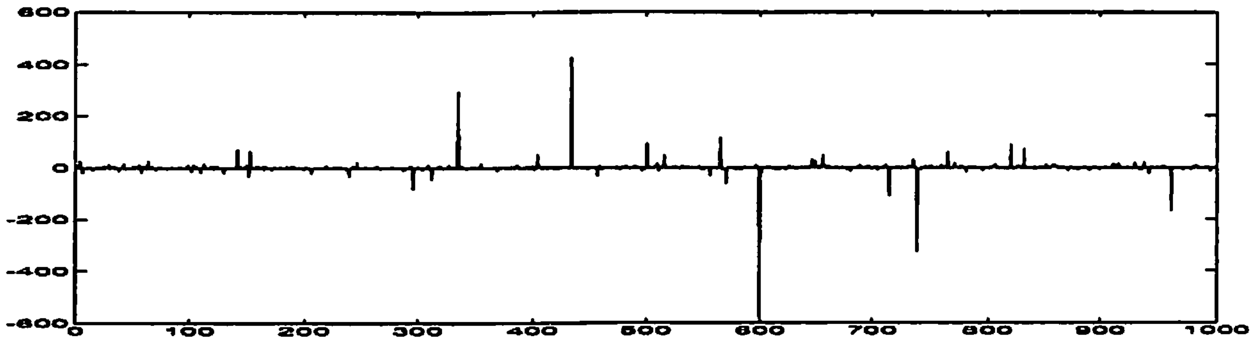


Fig. 7

1000 points; $\alpha = 1.5$ Sample of Data

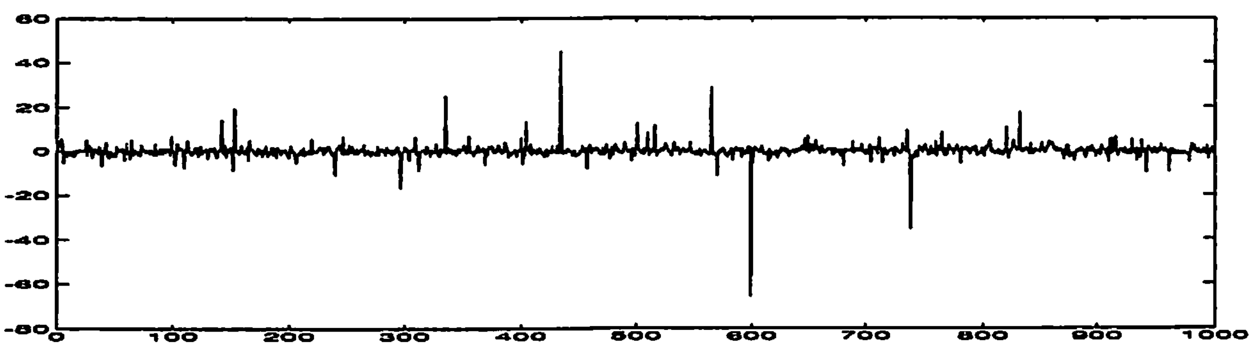
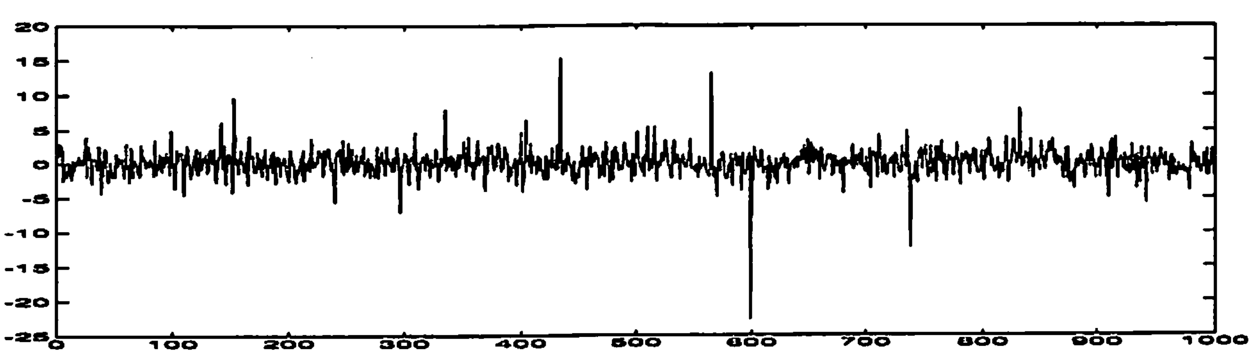


Fig. 8

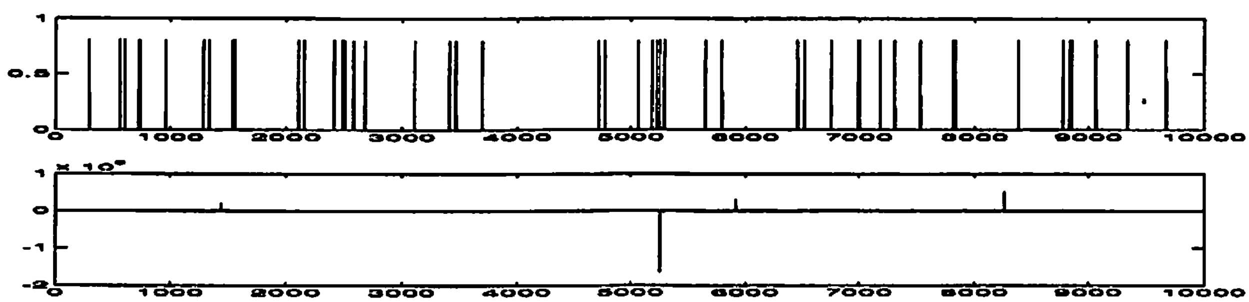
1000 points; $\alpha = 1.8$ Sample of Data



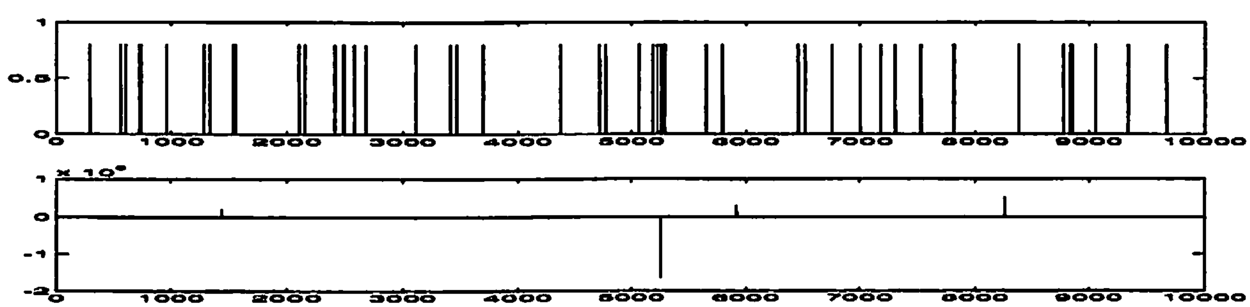
$\alpha = .5$ Data

$\alpha = .5$, Order = 40, Thres=5000, .7926

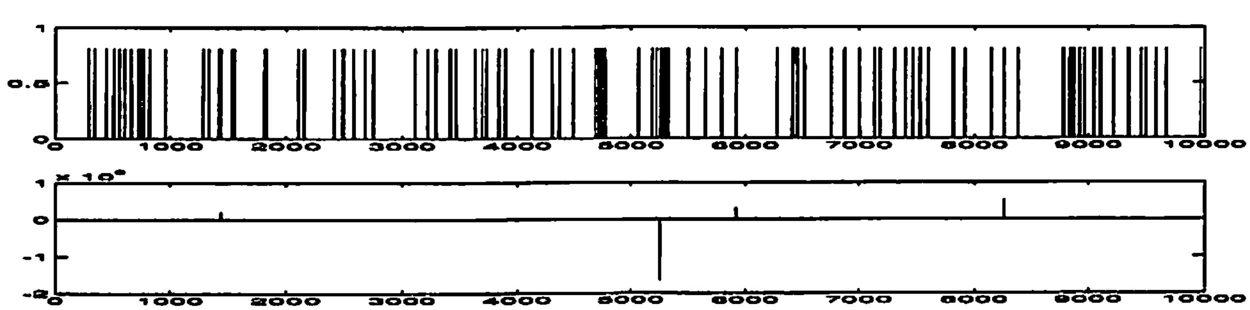
$\alpha = .5$, Order = 15, Thres=5000, .7989



$\alpha = .5$, Order = 5, Thres=5000, .7997

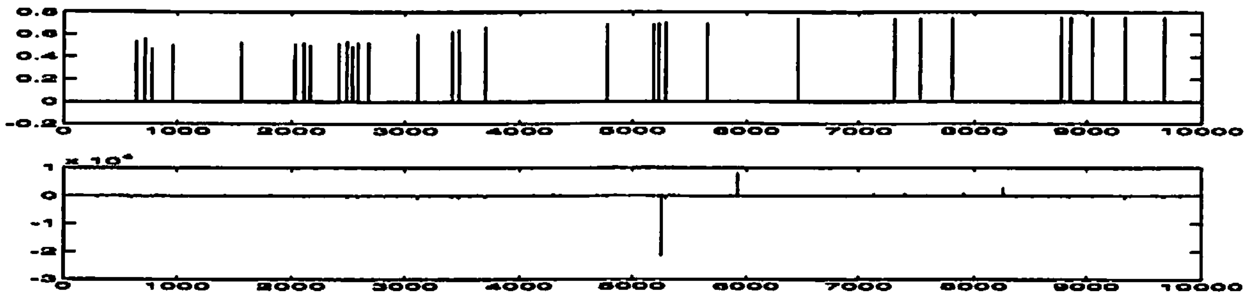


$\alpha = .5$, Order = 2, Thres=5000, .8

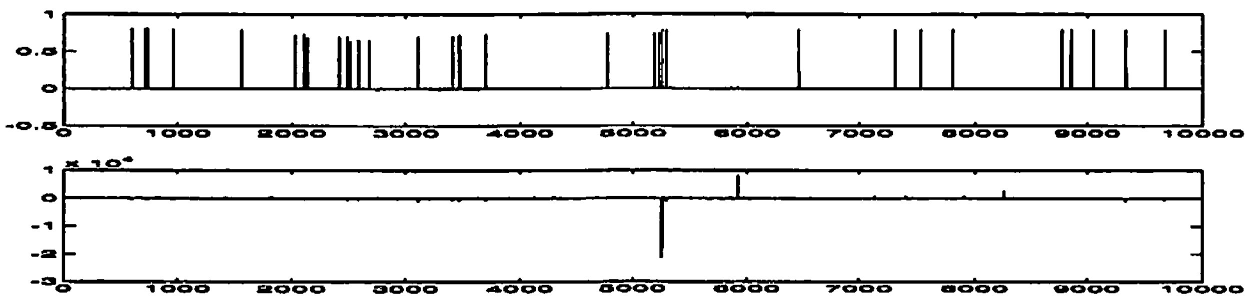


$\alpha = 1.$ (Cauchy) Data

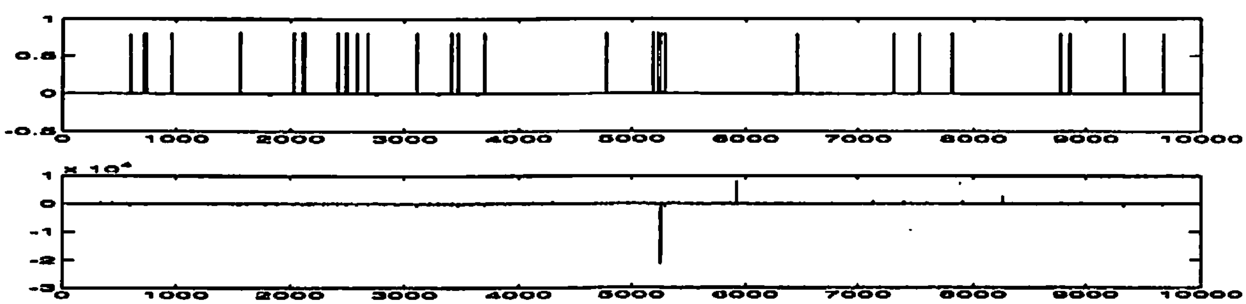
$\alpha = 1.$, Order = 40, Thres=100, .7553



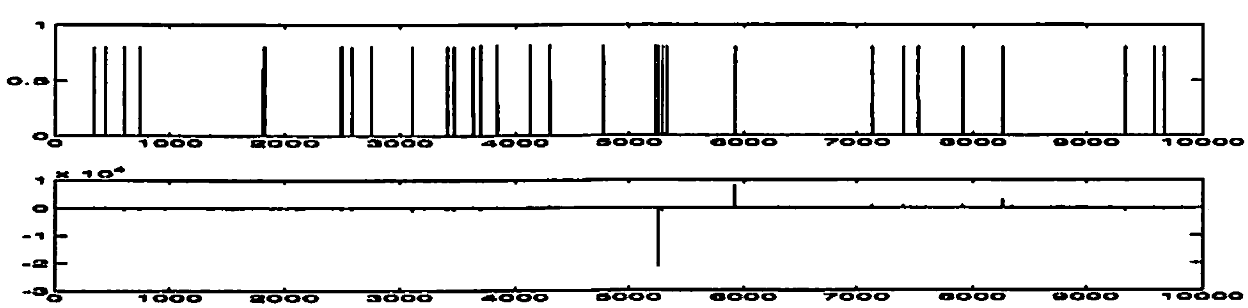
$\alpha = 1.$, Order = 15, Thres=100, .7877



$\alpha = 1.$, Order = 5, Thres=100, .7985

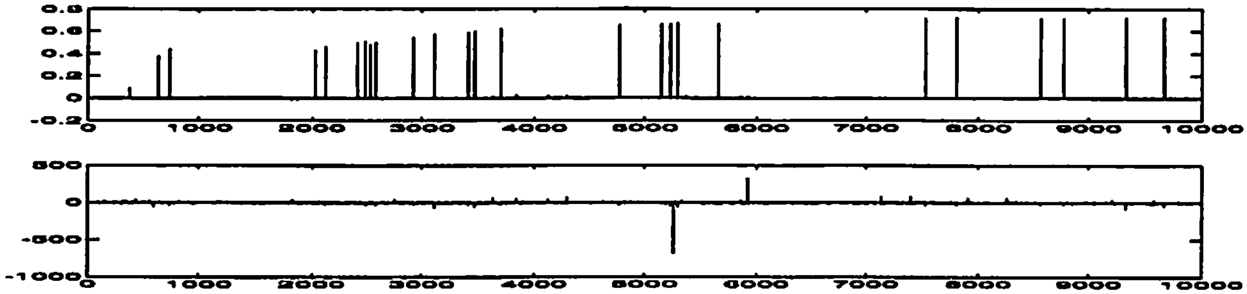


$\alpha = 1.$, Order = 2, Thres=200, .8007

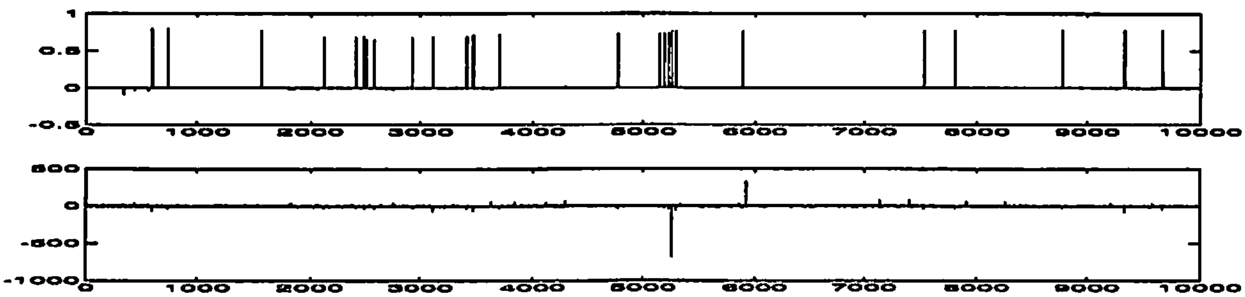


$\alpha = 1.5$ Data

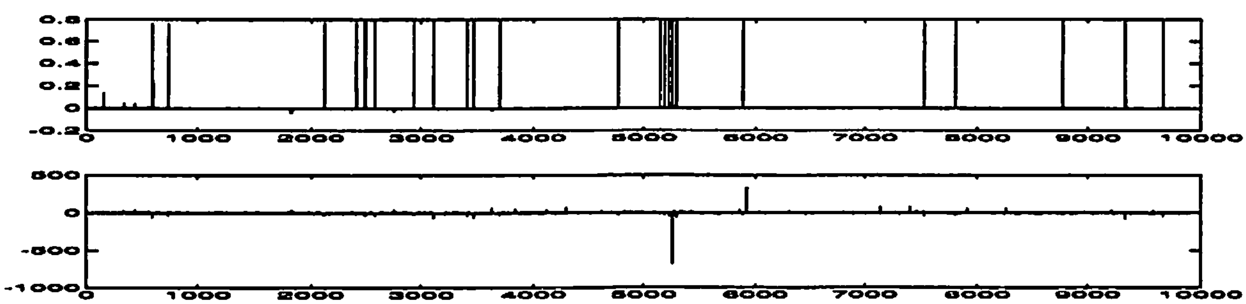
$\alpha = 1.5$, **Order = 40, Thres=20, .7250**



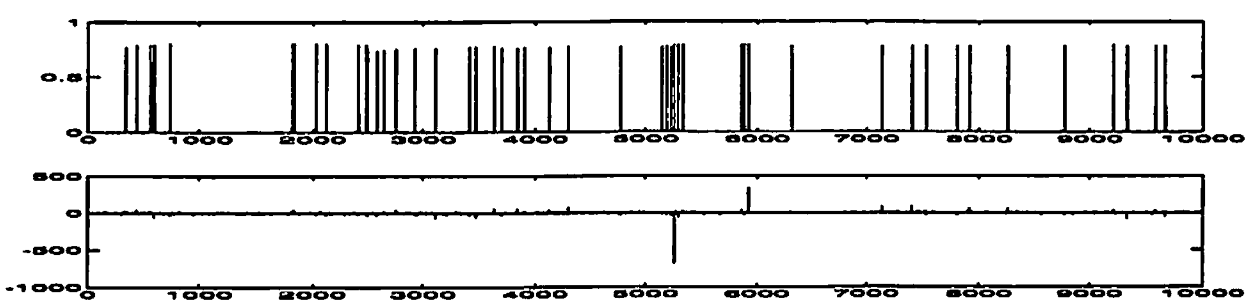
$\alpha = 1.5$, **Order = 15, Thres=20, .7759**



$\alpha = 1.5$, **Order = 5, Thres=20, .7835**

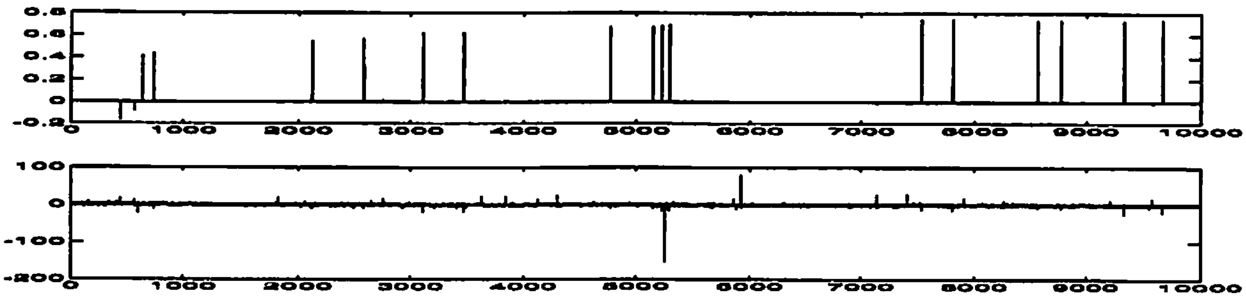


$\alpha = 1.5$, **Order = 2, Thres=20, .7790**

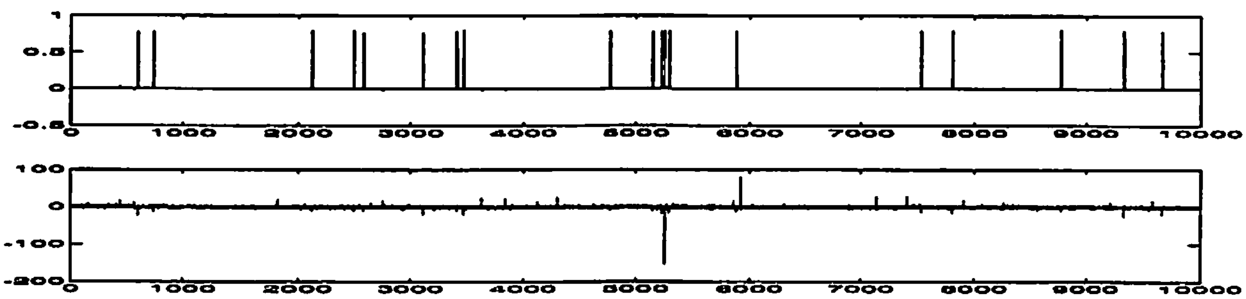


$\alpha = 1.8$ Data

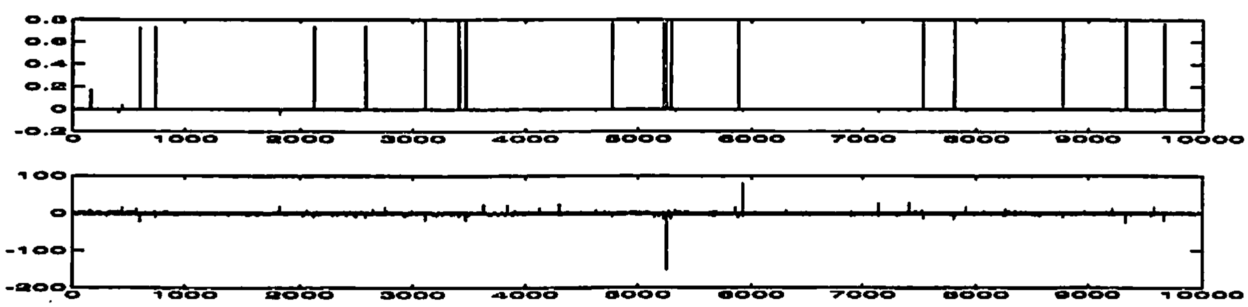
$\alpha = 1.8$, Order = 40, Thres=10, .7377



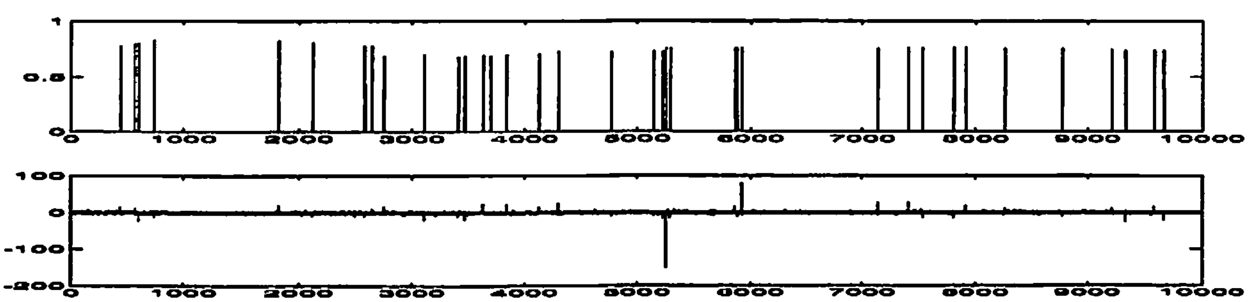
$\alpha = 1.8$, Order = 15, Thres=10, .7909



$\alpha = 1.8$, Order = 5, Thres=10, .7762



$\alpha = 1.8$, Order = 2, Thres=10, .7391



7 Conclusions

A method of deconvolution of Cauchy data is described using a Maximum A Posteriori method. The Maximum A Posteriori formulation supports the simplistic intuitive approach. By modelling the input data in the immediate neighborhood of extreme outliers as scaled versions of the channel, the method simply combines these "scale channels" in a linear way. The only nonlinearity in this technique, is in the specification of a threshold for determination of the presence of an extreme outlier.

We saw the method worked well for Cauchy data and also tried the method using other type of alpha-stable data. The result was that the method resulted in good estimates even when the data became fairly close to Gaussian. The simulations show that by adjusting the threshold we can sacrifice adaptivity for better estimates. An example of this is with the alpha=1.8 data. We found that one can still obtain good estimates of the filter when we are willing to adjust the threshold high and wait longer between updates.

The method appears to behave quite well over a broad range of data types and filter orders. The Maximum A Posteriori method which was developed using the Cauchy data case is shown by simulation to extend to other types of Alpha-Stable data.

8 Further Research

Because of the existence of a simple expression for the characteristic function of alpha-stable data, it may be possible to make a more general derivation. Also the derivative of the alpha-stable pdf is fairly simple, and may yield an even more elegant solution by going directly to a derivative expression and equating it to zero in order to minimize it.

References

[1] Shao, M. and Nikias, C., "Signal processing with fractional lower order moments: Stable processes and their applications," *IEEE Proceedings.*, vol. 7, July 1993, pp. 986-1010.

[2] Hsing, T., "On some estimates based on sample behavior near high level excursions," *Texas A & M Dept of Statistics Research Report,*

[3] Cline, D. and Brockwell, P., "Linear prediction of ARMA processes with infinite varience," *Stochastic Processes and their applications.*, vol. 19, 1985, pp. 281-297.

[4] H. Van Trees, **Detection, Estimation and Modulation Theory Part 1**,J. Wiley and Sons, 1968.