

Alpha-Stable Distributions in Signal Processing of Audio Signals

Preben Kidmose,
Department of Mathematical Modelling, Section for Digital Signal Processing,
Technical University of Denmark, Building 321, DK-2800 Lyngby, Denmark

Abstract

First, we propose two versions of a sliding window, block based parameter estimator for estimating the parameters in a symmetrical stable distribution. The proposed estimator is suitable for parameter estimation in audio signals.

Second, the suitability of the stable distribution, for modelling audio signals, is discussed. For a broad class of audio signals, the distribution and the stationarity property are examined. It is empirically shown that the class of stable distributions provides a better model for audio signals, than the Gaussian distribution model.

Third, for demonstrating the applicability of stable distributions in audio processing, a classical problem from statistical signal processing, stochastic gradient adaptive filtering, is considered.

1 Introduction

The probability density of many physical phenomena have tails that are heavier than the tails of the Gaussian density. If a physical process has heavier tails than the Gaussian density, and if the process has the probabilistic stability property, the class of stable distributions may provide a useful model.

Stable laws have found applications in diverse fields, including physics, astronomy, biology and electrical engineering. But despite the fact that the stable distribution is a direct generalization of the popular Gaussian distribution, and shares a lot of its useful properties, the stable laws have been given little attention from researchers in signal processing.

A central part of statistical signal processing, is the linear theory of stochastic processes. For second order processes, the theory is established, and numerous algorithms are developed. Applying these algorithms on lower order processes, results in considerably performance degradation or the algorithms may not even be stable. Thus, there is a need for developing algorithms based on linear theory for stable processes, and these algorithms could improve performance and robustness.

2 Modelling Audio Signals

The class of stable distributions is an appealing class for modelling phenomena of impulsive nature; and it is to some extent analytical tractable, because of two important properties: it is a closed class of distributions, and it satisfies the *generalized central limit theorem*.

Audio signals in general are not stationary, the temporal correlation is time varying, and it turns out that the probability density function is more heavy tailed than a Gaussian density. In this work we assume that probability density is symmetric, which is a weak restriction for audio signals. In particular we will consider the Symmetric Alpha Stable, $S\alpha S$, distribution [5]. To show the general applicability of the $S\alpha S$ for modelling audio signals, we examine six audio signals, with very different characteristics.

2.1 The $S\alpha S$ Distribution

A univariate distribution function is $S\alpha S$ if the characteristic function has the form

$$\phi(t) = \exp(-\gamma|t|^\alpha)$$

where the real parameters α and γ satisfies $0 < \alpha \leq 2$ and $\gamma > 0$. The parameter α is called the characteristic exponent. The smaller α -value the more probability mass in the tails of the density function. For $\alpha = 2$ it is the Gaussian distribution, and for $\alpha = 1$ it is the Cauchy distribution. The scale parameter, γ , is denoted the dispersion.

For stable distributions, moments only exist for order less than the characteristic exponent. If x is a S α S random variable, then the fractional lower order moment is

$$E(|x|^p) = C(p, \alpha)\gamma^{\frac{p}{\alpha}} \quad 0 < p < \alpha$$

where $C(p, \alpha) = (2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(\frac{-p}{\alpha})) / (\alpha\sqrt{\pi}\Gamma(\frac{-p}{\alpha}))$ and $\Gamma(\cdot)$ is the usual Γ -function [5, 4].

2.2 Parameter Estimation for S α S Distribution

Several methods, for estimation of the parameters in stable distributions, have been proposed in the literature, see [3, 6] and references herein. In this work, the estimation is performed in the $\ln|S\alpha S|$ -process, as proposed in [3]. Estimation in the $\ln|S\alpha S|$ -process has reasonable estimation characteristics, low computational complexity, and the estimator is on a closed form expression. For a S α S random variable x , the estimation in the $\ln|S\alpha S|$ -process is given as

$$\sigma_z^2 = \frac{\pi^2}{6} \left(\frac{1}{\alpha_x^2} + \frac{1}{2} \right) \quad (1)$$

$$\mu_z = C_e \left(\frac{1}{\alpha_x} - 1 \right) + \frac{1}{\alpha_x} \ln \gamma_x \quad (2)$$

where μ_z and σ_z^2 is the first and second moment of $z = \ln|x|$, and $C_e = 0.57721566\dots$ is the Euler constant. The characteristics of the α -estimator is depicted in Fig.1. The estimator in Eq.1 requires many

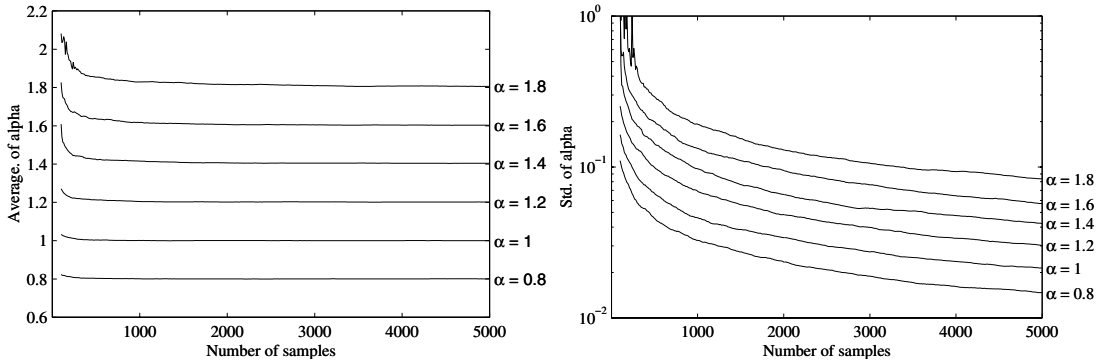


Figure 1 Characteristics of the α -estimator in Eq.1. Average (left) and standard deviation (right) of the α -estimate versus number of samples in estimation.

samples to give low variance estimates, and the variance is dependent of the characteristic exponent.

2.3 Sliding Window, Block Based Parameter Estimation

In this section we propose two versions of a sliding window, block based parameter estimator, suitable for audio signals. The basic idea of the estimator, is based on two observations: Audio signals often has strong short term correlations, due to mechanical resonans. These short term correlations have big influence on the short term distribution, this is particularly the case for mechanical systems with low damping combined with heavy tailed exitation signals. And the stationarity characteristics of the audio signals, necessitate the use of a windowed parameter estimation.

Thus, there is a need for a windowed estimator, that is robust to the influence from short terms correlations. Basically the proposed estimator has two steps, that makes the estimator suitable for handling these characteristics: A short term decorrelation, based on a linear prediction filter. And a sliding window, block based updating of the S α S-parameters of the decorrelated signal sequences. The short term decorrelation is performed over a block of M samples, and the updating of the S α S-parameters is performed over N blocks.

Let k denote the current block, the decorrelation is then performed over $x(n)|_{n=(k-1)M+1, \dots, kM}$, and the total window length, MN samples, applies over $x(n)|_{n=(k-N)M+1, \dots, kM}$.

Mechanical resonans is well modelled by a simple low order AR system, thus the resonans part of the signal can be removed by a linear predictor. The linear predictor coefficients for the k th block, $a_k(l)$, $l = 1, \dots, L$, is determined by

$$\sum_{l=1}^L a_k(l)r_x(m-l) = r_x(m)$$

where $r_x(m)$ is the autocorrelation sequence of $x(n)|_{n=(k-1)M+1, \dots, kM}$. The decorrelated signal, $y(n)$, in the k th block, is determined as the inverse filter

$$y(n)|_{n=(k-1)M+1, \dots, kM} = x(n) - \sum_{l=1}^L a_k(l)x(n-l)$$

In order to apply the estimator in the $\ln|\text{S}\alpha\text{S}|$ -process, we define the signal $z(n) = \ln|y(n)|$. A windowed estimator of $E(z)$ over N blocks of M samples is

$$\hat{\mu}_z(k) = \frac{1}{NM} \sum_m^{kM} z(m)$$

where the sum is over the N blocks, i.e. $m = (k-N)M + 1, \dots, kM$. The estimator of the expectation value is updated as

$$\hat{\mu}_z(k) = \hat{\mu}_z(k-1) - \frac{1}{N} \sum_m^{(k-N)M} z(m) + \frac{1}{N} \sum_m^k z(m) \quad (3)$$

where the first sum is over $m = (k-N-1)M + 1, \dots, (k-N)M$, and the second sum is over $m = (k-1)M + 1, \dots, kM$. Similar a windowed estimator of $E(z(m) - E(z(m)))^2$, over N blocks, is

$$\begin{aligned} \hat{\sigma}_z(k) &= \frac{1}{NM} \sum_m^k (z(m) - \hat{\mu}_z(k))^2 \\ &= \hat{\sigma}_z(k-1) - \frac{1}{N} \left(\sum_m^{(k-N)M} z^2(m) - \sum_m^k z^2(m) \right) + \hat{\mu}_z^2(k-1) - \hat{\mu}_z^2(k) \end{aligned} \quad (4)$$

Note that if the block sum and the quadratic block sum in Eq.3 and Eq.4 is saved for the last N blocks, then it is only necessary to calculate the k th block sum and the k th quadratic block sum for each block update. The windowed, block based estimators for the first and second moment of $z(n)$, combined with the $\text{S}\alpha\text{S}$ parameter estimators in Eq.1 and 2, yields¹:

$$\hat{\alpha}(k) = \left(\frac{6}{\pi^2} \hat{\sigma}_z^2(k) - \frac{1}{2} \right)^{-1/2} \quad (5)$$

$$\hat{\gamma}(k) = \exp \left(\left(\hat{\mu}_z(k) - \left(\frac{1}{\hat{\alpha}(k)} - 1 \right) C_e \right) \hat{\alpha}(k) \right) \quad (6)$$

The dynamic properties of estimators are of crucial importance in the case of non-stationary signals. The α -estimator in Eq.5 is sensitive to abrupt changes in the distribution, which might be an undesirable properties for non-stationary signals. An estimator, that is more robust against abrupt changes in distribution, is

$$\hat{\alpha}(k) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\frac{6}{\pi^2} \left(\frac{1}{M} \sum_{m=0}^{M-1} \left(z(k-m-nM) - \frac{1}{M} \sum_{l=0}^{M-1} z(k-l-nM) \right)^2 \right) - \frac{1}{2} \right)^{-1/2} \quad (7)$$

which is the empirical mean over N blocks of length M of the α -estimator in Eq. 1. This estimator is biased. In Fig.2 the proposed estimators is applied on a block-wise stationary $\text{S}\alpha\text{S}$ signal, and it is apparent that the estimator in Eq.7 is more robust in the case of abrupt changes in distribution.

¹If the update in Eq.3 and 4 is modified to be an iterative block based update, the update equations in Eq.5 and 6 is equivalent to the iterative update equations proposed in [3]. However, in this context, it is the sliding window property that is the important feature.

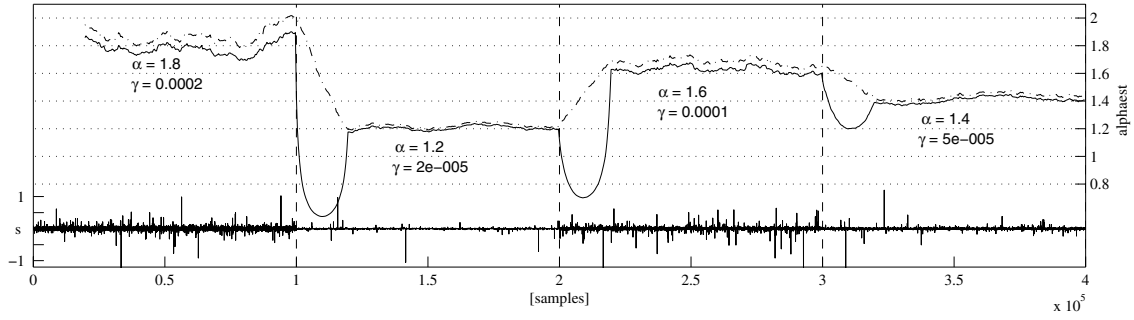


Figure 2 Dynamic properties of the proposed α -estimators. The solid line is the estimator in Eq.5, the dashed-dot line is the estimator in Eq.7. The signal, s , is stationary block-wise over 100000 samples; the estimator block size $M = 400$ and the sliding window is over $N = 50$ blocks.

2.4 S α S Modelling of Audio Signal

The signals used for demonstrating the applicability of S α S modelling, are depicted in the bottom part of the plots in Fig. 3, and some additional informations are listed in Tabel 1. In Fig.4 the empirical, the estimated Gaussian, and the estimated S α S density, for the signal sequences, are depicted. These density plots are obtained over the whole signal length. It is important to notice that, due to the non-stationarity

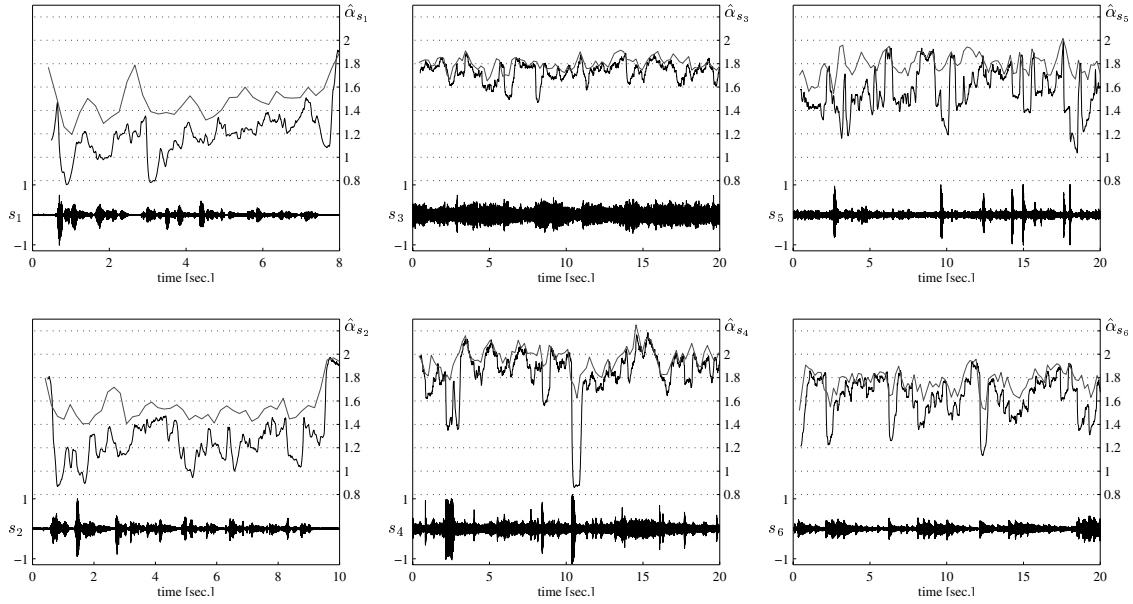


Figure 3 Estimates of the characteristic exponent, α . Solid lines: estimator Eq.5, the linear predictor window is 10 ms and $N = 50$ blocks. Dotted lines: estimator Eq.7, the linear predictor window is over 100 ms, and the mean is over $M = 5$ blocks. For both estimators the linear predictor filter is of order 12.

of the signals, the time window, in which the density is estimated, has deciding influence of the density estimate. Comparing the Gaussian model, which has only one degree of freedom in modelling of the shape of the probability density, with the S α S distribution, which has two degrees of freedom, it is reasonable to conclude that the S α S in general provides a better model for the probability density.

It is instructive to consider the distribution of the signals in shorter time windows. The proposed sliding window estimators in Eq.5 and Eq.7 are applied to the six signals, and the estimate of the characteristic exponent is depicted on the same time axis as the signals in Fig 3. The solid line is the characteristic exponent estimated with the parameter estimator in Eq.5. The linear predictor window is 10 ms, and the estimator window is over 50 blocks. Due to the different sampling frequencies 20 kHz, 32 kHz, and 44.1 kHz, this corresponds to a window length of 200, 320 and 441 samples respectively. The dotted lines is

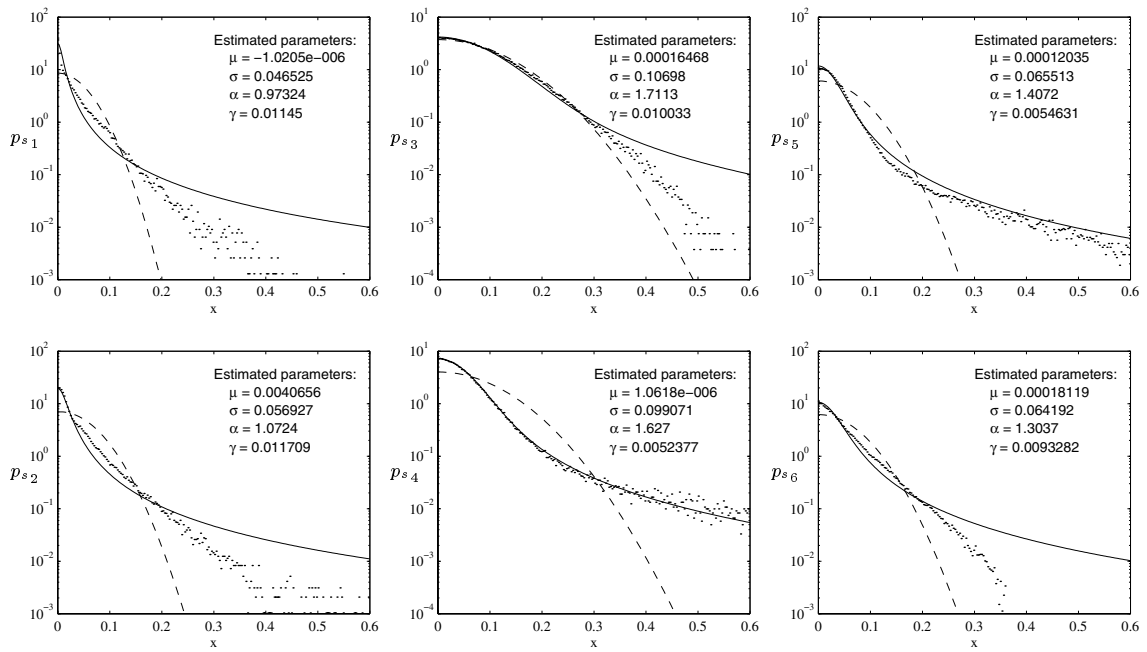


Figure 4 Probability density functions for the signals. The dots indicates the empirical density function. The dashed line is the Gaussian density corresponding to the estimated mean, μ , and standard deviation, σ^2 . The solid line is the S α S density corresponding to the estimated characteristic exponent, α , and dispersion, γ . The estimated parameter values, for each signal, is tabulated in the upper right corner.

characteristic exponent estimated with the parameter estimator in Eq.7. The linear predictor window is over 100 ms, and the mean is over 5 blocks. For both estimators the linear predictor filter is of order 12.

The short term estimates of the characteristic exponent are in general larger than the long term estimates. It is interesting to notice that the speech signals has relative low α -values and in certain intervals approaches the Cauchy distribution. The background noise signals has α -values that in general are considerable larger, but still far below the Gaussian distribution. In Fig.5 the density estimates are depicted for the signal $s_2(n)$ and $s_4(n)$ for two different time windows, and again it is reasonable to conclude that the S α S provide a better fit to the empirical histogram.

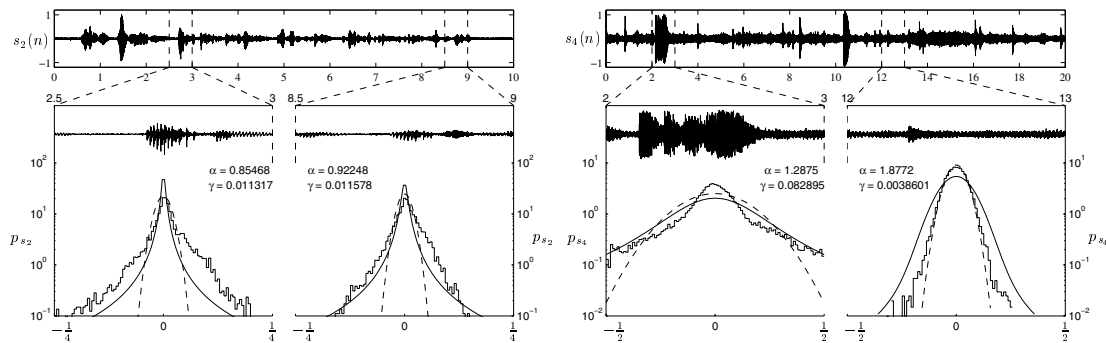


Figure 5 Examples of short term density estimates for the signals $s_2(n)$ and $s_4(n)$ for different time intervals. The stair plot indicates the empirical density function. The dashed line is the Gaussian density corresponding to the estimated mean and standard deviation. The solid line is the S α S density corresponding to the estimated characteristic exponent, α , and dispersion, γ . The block size is 320 and 200 samples respectively which corresponds to 10ms. The estimator in Eq.5 is applied over 50 and 100 blocks respectively, and no linear prediction filter has been used.

It is well-known from the theory of stable distributions, that moments only exists for moments of order less than α . The preceding examinations, that indicates that the S α S is suitable for modelling audio signals, and that the characteristic exponent varies between a Cauchy and a Gaussian distribution, exposes the idea to use the estimates of the characteristic exponent in variable fractional lower order moment adaptive

Signal	Description
$s_1(n)$	Speech signal, male, low background noise. Sampling freq. 32 kHz., 8 sec.
$s_2(n)$	Speech signal, male, low background noise. Sampling freq. 32 kHz., 10 sec.
$s_3(n)$	Cocktail party background noise. Sampling freq. 44.1 kHz., 20 sec.
$s_4(n)$	Background noise recorded in a kitchen. Sampling freq. 20 kHz., 20 sec.
$s_5(n)$	Background noise recorded in an office. Sampling freq. 44.1 kHz., 20 sec.
$s_6(n)$	Music, classical guitar. Sampling freq. 44.1 kHz., 20 sec.

Table 1 Additional information for the audio signals.

algorithms. This idea is the issue of the following section.

3 Adaptive Filtering

An illustrative application for adaptive filtering, is the acoustical echo canceller. The objective is to cancel out the loudspeaker signal from the microphone signal, see Fig.6. An adaptive filter is applied to estimate the acoustical channel from the loudspeaker to the microphone. The echo cancelled signal is obtained by subtracting the remote signal, filtered by the estimated acoustical channel, from the microphone signal. From an algorithmic point of view the local speaker is a noise signal, and the applied adaptiv algorithm must exhibit adequate robustnes against this noise signal.

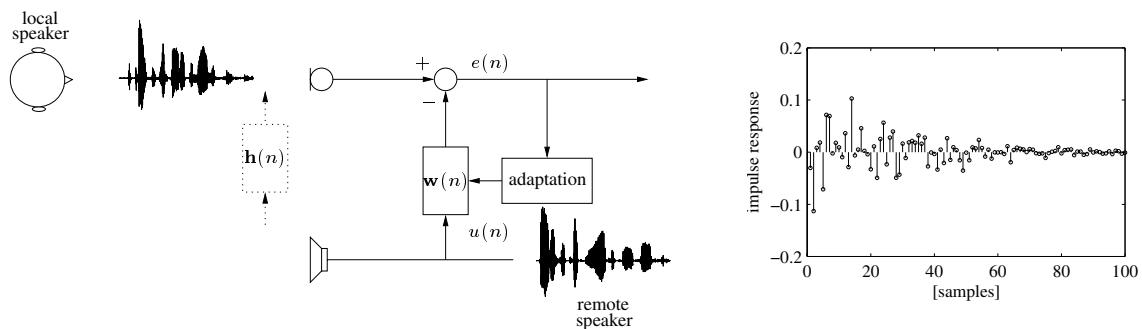


Figure 6 Left: Acoustical echo canceller setup. Right: Impulse response for the acoustical channel, h , from loudspeaker to microphone.

The standard algorithm for adaptive filters, is the Normalized Least Mean Square (NLMS) algorithm, with the update

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{\mathbf{u}(n)e(n)}{a + \|\mathbf{u}(n)\|_2^2}$$

The NLMS algorithm has severe convergence problems for signals with more probability mass in the tails, than the Gaussian distribution. Recently filter theory for S α S signals has been developed [1], and the Least Mean P-norm (LMP) algorithm is proposed. The LMP algorithm is significant more robust to signal with heavy tails. In the following simulation study a normalized LMP update is applied:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{\mathbf{u}(n)e(n)^{p-1}}{a + \|\mathbf{u}(n)\|_p^p}$$

For S α S signals the p -norm must be less than the characteristic exponent, $p < \alpha$. The development of robust adaptive filters is subject to active research, see [2] and references herein.

Consider the adaptiv echo canceller setup, with the signal scenario as depicted in the upper part of Fig.7. The local speaker is the speech signal s_1 , the speaker is deactivated in the time interval 1-2 sec, and added some cocktail party background noise. The remote speaker is the signal s_2 , the speaker is deactivated in the time interval 4-6 sec. and in that time interval the loadspeaker signal is a Gaussian noise signal (comfort noise, $\sigma^2 = .005$). The α -estimator in Eq.5 is applied to the error signal, $e(n)$, the block length is 640, the estimator runs over 10 blocks, and the linear predictor filter is of order 12. Three different adaptive

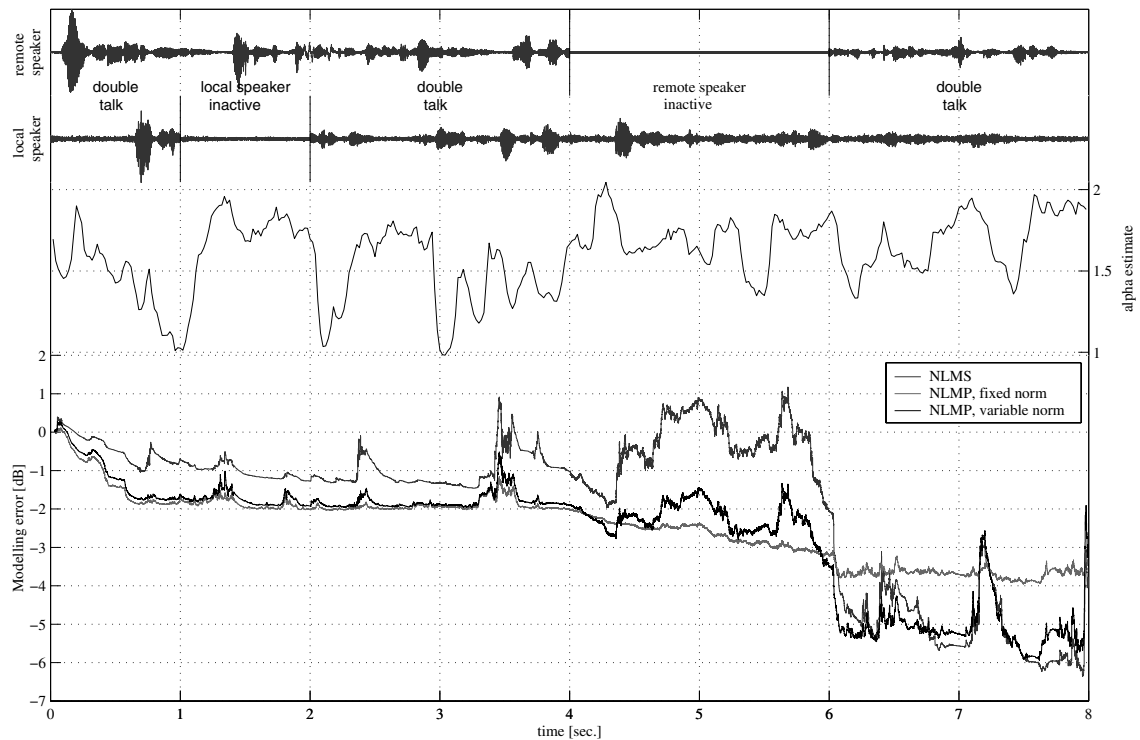


Figure 7 Simulation result for acoustical echo canceller scenario.

filters is applied, the standard NLMS, the NLMP with a fixed norm ($p = 1.1$), and the NLMP with a norm that is adjusted in accordance with the α -estimate ($p = \hat{\alpha} - .2$). The stepsize parameter $\mu = .001$ for NLMS, and $\mu = .002$ for NLMP. The performance of the algorithms is evaluated as the modelling error $10 \cdot \log_{10} E \{ (\mathbf{w} - \mathbf{h})^T (\mathbf{w} - \mathbf{h}) / (\mathbf{h}^T \mathbf{h}) \}$, and is depicted in the bottom part of Fig.7.

The $p = 1.1$ for the fixed norm NLMP was empirical found to be the best norm for the applied signals, and the NLMP algorithm with fixed norm in general performed very good. The modelling error for NLMP algorithm with variable norm, is between the modelling error of the two others algorithms. The variable norm algorithm follows the best of the two other algorithms, and it is thus concluded that the variable norm algorithm has overall better performance. However, the result is far from convincing, and the conclusion might not hold in general, because of the dependence of the α -estimator.

The variable norm NLMP algorithm is computational much more expensive than using a fixed norm algorithm, because of the running α -estimator. The small gain in performance do probably not justify the additional computational expenses. Despite this fact, the simulation study shows, that the choice of norm has deciding influence of the performance of the algorithms.

4 Conclusion

The proposed sliding window, block based parameter estimators has been applied to a broad class of audio signals. Comparing the histogram of the audio signals with the estimated parameters of a $S\alpha S$ distribution, it is concluded that the class of $S\alpha S$ distributions is suitable for modelling audio signals.

The simulation study shows that lower norms algorithms exhibit better robustness characteristic for audio signals, and that the choice of norm has deciding influence of the performance of the algorithm.

Stable distributions provides a framework for synthesis of robust algorithms for a broad class of signals. The linear theory of stable distributions and processes, and the development of robust algorithms for impulsive signals is an open research area.

References

- [1] John S. Bodenschatz and Chrysostomos L. Nikias. Symmetric Alpha-Stable Filter Theory. *IEEE Transactions on Signal Processing*, 45(9):2301–2306, 1997.
- [2] Preben Kidmose. Adaptive Filtering for Non-Gaussian Processes. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pages 424–427, 2000.
- [3] Xinyu Ma and Chrysostomos L. Nikias. Parameter Estimation and Blind Channel Identification in Impulsive Signal Environments. *IEEE Transactions on Signal Processing*, 43(12):2884–2897, December 1995.
- [4] Gennady Samorodnitsky and Murad S. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman&Hall, 1994.
- [5] Min Shao and Chrysostomos L. Nikias. Signal Processing with Fractional Lower Order Moments: Stable Processes and Their Applications. *Proceeding of the IEEE*, 81(7):986–1010, July 1993.
- [6] George A. Tsihrintzis and Chrysostomos L. Nikias. Fast Estimation of the Parameters of Alpha-Stable Impulsive Interference. *IEEE Transactions on Signal Processing*, 44(6):1492–1503, June 1996.