A SEMI-PARAMETRIC APPROACH FOR ROBUST MULTIUSER DETECTION

Ulrich Hammes and Abdelhak M. Zoubir

Signal Processing Group, Institute of Telecommunications
Technische Universität Darmstadt
Merckstrasse 25, D-64283 Darmstadt, Germany
Email: {u.hammes, zoubir}@ieee.org

ABSTRACT

Robust parameter estimation in impulsive noise has risen a lot of attention in wireless communications. Previously, we proposed a non-parametric estimator for multiuser detection based on non-parametric density estimation. Here, we present a semi-parametric estimator that outperforms its non-parametric counterpart by combating multiple access interference and impulsive noise altogether. The approach is termed semi-parametric since a nonlinear parametric function is used to transform the noise data while non-parametric estimation of the score function is performed using the transformed sample. This estimate is then used to determine the parameters of interest, i.e., the transmitted symbols. We also propose a parametric function and an estimator for its parameter.

Index Terms— Robust Estimation, Multiuser Detection, Impulsive noise

1. INTRODUCTION

Generally, impulsive noise is due to different electromagnetic interference sources and mainly occurs in urban areas and indoor communication channels [1]. A specific problem that we consider here is Multiuser Detection (MUD) in i.i.d. impulsive noise. Conventional estimators for MUD fail in such noise environments since they are optimal only under a specific assumption (e.g. Gaussianity for the least-squares estimator) [2].

Approaches based on M-estimation that cope with deviations from model assumptions have been suggested in the literature. In particular, in [3] a minimax solution over a class of different distributions was suggested. This solution may be suboptimal for a particular model. Even though this approach improves performance over conventional MUD techniques in impulsive noise [2], adapting the estimator to the underlying noise pdf is desirable. In this view, in [4] the focus was set on a parametric approach where a model selection scheme for modeling the noise pdf was suggested. Another approach [5] is based on a non-parametric estimate of the pdf, using adaptive kernel density estimation (AKDE). One limitation of this approach is the necessary selection of local bandwidths to fit the tails of the distribution. An inaccurate choice of local bandwidths leads to poor estimates of the noise pdf and consequently to poor estimates of the parameters of interest.

Here, a new estimator is presented that combines the advantages of parametric and non-parametric approaches to improve small sample performance. This approach uses a parametric, nonlinear function to transform the noise data in a convenient way, so that the problem of local bandwidth selection is overcome. Kernel density estimation (KDE) of the transformed data can then be performed using a global bandwidth only. An estimate of the noise density is provided via back-transformation and used to estimate the parameters of interest.

In Section 2, the signal and noise model are introduced and the concepts of robust estimation are presented. Section 3 describes the novel method whereas Section 4 addresses asymptotic behaviour of the adaptive approaches. Simulation results of different detectors in different noise scenarios are given in Section 5. Section 6 concludes the paper.

2. PROBLEM STATEMENT

2.1. Signal Model

We consider the uplink channel of a CDMA system where K users transmit at the same time. The received signal is

$$y_i = \sum_{k=1}^{K} \mathbf{S}_{ik} A_k b_k + n_i \qquad i = 1, \cdots, N, \tag{1}$$

where $\mathbf{S}_{ik}$ denotes the normalised chip i of user k, A_k is the amplitude and b_k is the transmitted *Binary Phase Shift Keying* (BPSK) symbol of user k and $\theta = \mathbf{Ab}$ is to be estimated. Note that the length of the spreading code equals N and n_i is i.i.d. zero-mean noise with pdf $f(x)$. If $f(x)$ is known, the maximum likelihood estimate (MLE) is defined by

$$\hat{\theta}_{ML} = \arg\min_{\theta} \sum_{i=1}^{N} -\log f\left(y_i - \sum_{k=1}^{K} \mathbf{S}_{ik}\theta_k\right). \tag{2}$$

This work was supported by the Deutsche Forschungsgemeinschaft (DFG)

A unique solution of (2) can be obtained by solving its first-order derivative for zero. This corresponds to solving the following equation system:

$$\sum_{i=1}^{N} \mathbf{S}_{(ik)} \varphi \left(y_i - \sum_{k'=1}^{K} \mathbf{S}_{(ik')} \theta_{k'} \right) = 0, \quad k = 1, \cdots, K, \tag{3}$$

where $\varphi(x) = -f'(x)/f(x)$ is the location score function. The amplitude and symbol of each user are recovered as $\hat{A}_k = |\hat{\theta}_k|$ and $\hat{b}_k = sign(\hat{\theta}_k)$. If $f(x)$ is Gaussian, $\varphi(x)$ is linear and (3) corresponds to least-squares estimation.

2.2. Robust Estimation and Impulsive Noise Model

When the data deviates from the Gaussian assumption, the least-squares estimator is suboptimal. Huber proposed a minimax solution where $\varphi(x)$ in (3) is replaced by an influence function $\psi(x)$ that limits the effect of outlying values [3] and consequently confers robustness on the parameter estimate. In his work the following ε-contaminated Gaussian mixture model

$$f(x) = (1 - \varepsilon) f_G(x, 0, \nu^2) + \varepsilon \mathcal{H}, \tag{4}$$

was considered where $f_G(x, 0, \nu^2)$ is a zero-mean Gaussian pdf with variance ν^2 and $\mathcal{H}$ is an unknown symmetric pdf. The minimax estimator, that minimises the maximum asymptotic variance for the least favourable distribution $\mathcal{H}$, was applied to MUD in [2]. It was shown that significant improvements over conventional estimators can be achieved when the data deviates from the Gaussian assumption. However, this M-estimator, like any minimax estimator, may be far from optimal away from the least favourable distribution. For this reason, adaptive approaches are required which use the residuals $\hat{\mathbf{n}} = \mathbf{y} - \mathbf{S}\hat{\theta}$ to estimate the score function of the underlying noise pdf. A new method based on transformation density estimation is presented hereafter.

In [1] an impulsive noise model for wireless communication channels was suggested. Since this model is mathematically intractable, the Gaussian mixture model has been used as a common approximation [2] and is considered here. In this model, $\mathcal{H}$ of (4) is a zero-mean Gaussian pdf with variance $\kappa\nu^2$ which results in a symmetric, heavy-tailed and unimodal pdf. Typical values for ε and κ are between $0.01 \leq \varepsilon \leq 0.1$ and $10 \leq \kappa \leq 100$.

3. SEMI-PARAMETRIC ESTIMATOR

One approach to obtain adaptivity for unknown noise distributions is to estimate the pdf and its derivative non-parametrically. The estimated score function is then

$$\hat{\varphi}(x) = -\hat{f}'(x)/\hat{f}(x). \tag{5}$$

However, methods such as conventional KDE fail for heavy-tailed data since they produce spurious peaks in the tails. This problem can be partly overcome by selecting local bandwidths for different samples [6], as in [5]. However, inaccuracies in estimating the local bandwidths may still degrade performance and result in non-monotonicities in the tails of the pdf. This would lead to undesired oscillations in the estimated score function and consequently to convergence problems in the parameter estimation algorithm. In particular local bandwidth selection often fails for small sample sizes and heavy tailed data. To circumvent this problem we consider a semi-parametric approach based on transformation density estimation [7].

3.1. Algorithm

First an estimate of θ is obtained by a consistent estimator, e.g., least-squares, and the residuals are calculated. Consider the parametric transformation function $w = g(x, \gamma)$ given by

$$g(x, \gamma) = \begin{cases} x & : \quad |x| \leq \gamma \\ sign(x) \cdot [\ln(|x| - (\gamma - 1)) + \gamma] & : \quad |x| > \gamma. \end{cases} \tag{6}$$

When applying this transformation to the residuals, data points in the tails of the distribution or outliers, which are separated widely in the X-domain, come closer together in the W-domain. This simplifies the estimation in the W-domain for non-parametric KDE without local bandwidth selection. Since $g(x, \gamma)$ is linear in a certain region around zero, the core data remains untransformed. The estimated pdf $\hat{f}(x)$ and its derivative $\hat{f}'(x)$ in the X-domain are obtained using the back transformation $x = g^{-1}(w, \gamma)$ and an estimate of the score function is calculated. Then an iteration step of a Newton-Raphson algorithm suggested in [3] is performed to obtain the next estimate $\hat{\theta}^{i+1}$. This is repeated until convergence is reached (see Table for details). Note that $g(x, \gamma)$ is chosen heuristically and other nonlinear functions may be used for this algorithm. Notice that this algorithm does not require training sequences.

1.	*Initialisation* Set i=0. Obtain an initial estimate $\hat{\theta}^0$
2.	*Determine the residuals* $\hat{\mathbf{n}} = \mathbf{y} - \mathbf{S}\hat{\theta}^i$
3.	*Determine γ, transform data via $g(x, \gamma)$, estimate $f(x)$ and $f'(x)$ and estimate the score function $\hat{\varphi}(x) = -\hat{f}'(x)/\hat{f}(x)$*
4.	*Update the parameter estimates* $\hat{\theta}^{i+1} = \hat{\theta}^i + \mu(\mathbf{S}^\mathsf{T}\mathbf{S})^{-1}\mathbf{S}^\mathsf{T}\hat{\varphi}(\hat{\mathbf{n}})$ where $\mu = 1/(1.25\max(\lvert\hat{\varphi}'(\hat{\mathbf{n}})\rvert))$
5.	*Check for convergence* If $\lvert\frac{\hat{\theta}^{i+1} - \hat{\theta}^i}{\hat{\theta}^{i+1}}\rvert < \epsilon$, stop, where $\epsilon \in \mathbb{R}$ is a small number. otherwise $i \rightarrow i + 1$ and go to step 2

3.2. Transformation Density Estimation

Transformations in density estimation have been suggested in [7] to estimate skewed and heavy-tailed data. Here, the approach has been modified and adapted to symmetric noise data which is a common assumption in communications. The smooth pdf obtained in the W-domain is estimated with less difficulty and its estimate is then back-transformed by a monotonic function. This nonlinear function ensures that the pdf in the X-domain inherits the smoothness property of its image in the W-domain. Kernel estimates of the density $f(x)$ and its derivative are given by

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} H\left(\frac{g(x,\gamma)-w_i}{h}\right)\left(\left|\frac{\partial g^{-1}(w,\gamma)}{\partial w}\right|\right)^{-1} \quad (7)$$

$$\hat{f}'(x) = -\frac{1}{Nh}\sum_{i=1}^{N}\left(\frac{g(x,\gamma)-w_i}{h^2}\right) H\left(\frac{g(x,\gamma)-w_i}{h}\right) \cdot\frac{\partial g(x,\gamma)}{\partial x}\left(\left|\frac{\partial g^{-1}(w,\gamma)}{\partial w}\right|\right)^{-1}, \quad (8)$$

where $w_i = g(\hat{n}_i, \gamma)$ and $H(x)$ is the standard Gaussian pdf. Several estimators for selecting the global bandwidth h are available in the literature [6]. Here, we use $\hat{h} = 1.06\hat{\sigma}N^{-1/5}$ where $\hat{\sigma}$ is a robust estimate of scale. This bandwidth estimate is obtained instantly and provides satisfying results here (note that our goal is to estimate θ and not $f(x)$).
In order to satisfy the symmetry constraint we use a symmetrised version $\hat{f}_s(x) = (\hat{f}(x) + \hat{f}(-x))/2$ instead of $\hat{f}(x)$ that improves small sample performance.
Furthermore, the unimodality property of $f(x)$ is incorporated into the estimator in order to reduce the error. In case the estimate $\hat{f}(x)$ contains multiple modes, the global bandwidth is thus increased by a factor η iteratively until a unimodal density is obtained [6]. For the most impulsive noise scenario, i.e. $\varepsilon = 0.1$ and $\kappa = 100$, five iterations on average over 1000 repetitions for the semi-parametric and six for the non-parametric approach were enough to achieve a unimodal density using $\eta = 1.05$, as suggested in [5]. Hence, computational complexity is slightly decreased in this respect.

3.3. Choice of the Transformation Parameter γ

Values of the residuals smaller than γ remain untransformed and may be easily estimated using a global bandwidth. Hence γ has to be chosen such that outliers are transformed nonlinearly to approach the core of the data in the W-domain. For this reason it is important to detect outliers in the data set for determining γ. Here we use an outlier detection method suggested in [8] that normalises the residuals using robust estimates of location and scale. Any value greater than a certain threshold is considered to be an outlier. The standardised residuals are given by

$$z_i = \frac{\hat{n}_i - \text{median}(\hat{\mathbf{n}})}{\text{mad}(\hat{\mathbf{n}})}, \quad (9)$$

where $\text{mad}(\cdot)$ is the median absolute deviation. Here the threshold is set arbitrarily to 3, meaning that each value $\hat{n}_i$ for $|z_i| > 3$ is defined as an outlier. After the outliers are detected the absolute values of the data are ordered. Parameter γ is set to the absolute value of the largest order statistic $|n[l]|$ not belonging to the set of outliers, i.e., $|n[1]| < .. < |n[l]| = \gamma < |n[Outlier]| < .. < max(|\hat{\mathbf{n}}|)$. This choice ensures that only outliers and a small amount of data in the tails are transformed closer to the core. If no outliers are detected γ is set to $\gamma = 3\text{mad}(\hat{\mathbf{n}})$.

4. ASYMPTOTIC BEHAVIOUR

In order to gain insight into the asymptotic behaviour of the proposed estimator, we study an estimate of the asymptotic variance given by

$$\hat{AV}(f,\hat{\varphi}) = \frac{\int_{-\infty}^{\infty}\hat{\varphi}(x)^2 f(x)dx}{\left(\int_{-\infty}^{\infty}\hat{\varphi}'(x)f(x)dx\right)^2}, \quad (10)$$

where $\hat{\varphi}(x)$ is estimated for a noise sample of $N = 31$ points. The relative efficiency (RE) is used as a metric to compare the estimators to the MLE. It is defined as the ratio of the Cramér-Rao-Bound, i.e., the asymptotic variance of the MLE, where $\hat{\varphi}(x)$ in (10) is replaced by $\varphi(x)$ and the estimated asymptotic variance (10). We consider the Gaussian mixture noise model given in Section 2.2 and compare the non-parametric estimator given by (5) using AKDE and the semi-parametric estimator given in Section 3.1. The REs of both estimators are given in Figures 1 and 2.
A possible explanation for the opposite slopes of the two surfaces is that for the non-parametric estimator, local bandwidth selection fails in highly impulsive noise environments, causing oscillations in the score function. This yields to lower efficiency of the estimator. The transformation function used for the semi-parametric approach stabilises the shape of the estimated score function, which increases efficiency for high ε and κ. For non-kurtotic pdfs, nonlinearity may distort the data, thus leading to a small performance loss with respect to the estimator based on AKDE.

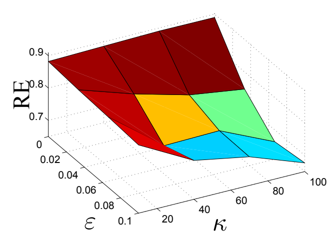

Fig. 1. RE of the non-parametric estimator

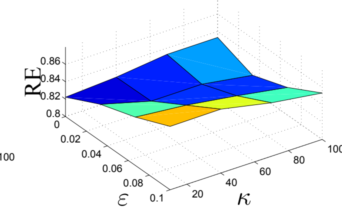

Fig. 2. RE of the semi-parametric estimator

5. SIMULATIONS

The linear decorrelating detector, the soft-limiter, the non-parametric estimator [5] and the semi-parametric estimator are applied to the CDMA scenario given in Section 2. Simulations are performed over 10^5 Monte Carlo runs in different noise environments with a near-to-far ratio (NFR) of 10dB. $K = 6$ users with Gold sequences of length $N = 31$ are used, meaning density estimation is performed with 31 sample points. The scale for bandwidth selection is estimated using $\hat{\sigma} = 1.483\mathrm{mad}(\hat{\mathbf{n}})$ and the clipping point c of the soft-limiter is chosen to be $c = 1.5\hat{\sigma}$, as in [2].
The Bit-Error-Rate (BER) of user 1 versus the Signal-to-Noise-Ratio (SNR) in Gaussian noise is shown in Figure 3. We observe that the linear detector and soft-limiter perform slightly better than the adaptive approaches. At high SNR the performance gap is about 1.5 dB. However, when the impulsiveness of the noise density increases, i.e., either ε or κ increase, the linear detector suffers a significant performance loss whereas the robust detectors are able to cope with these noise environments, as illustrated in Figure 4. Here, one can see that the adaptive approaches perform significantly better in impulsive noise environments than the static M-estimator, whereas the linear detector breaks down. The semi-parametric approach outperforms its non-parametric counterpart by at most 3dB. Similar results were obtained for different ε and κ.

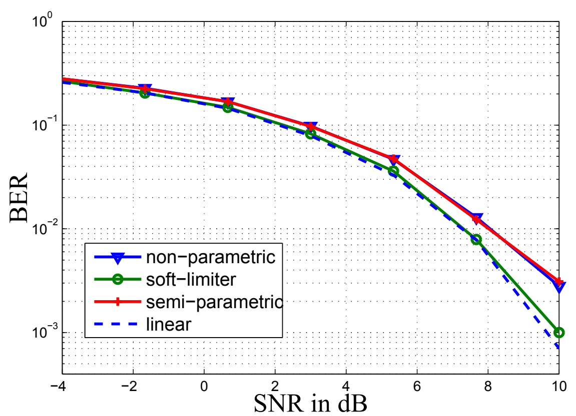

Fig. 3. BER versus SNR for user 1 in a synchronous CDMA channel with Gaussian noise. N=31, K=6, NFR=10dB

6. CONCLUSION

In this paper, a semi-parametric estimator based on transformation density estimation was presented. It was found that in impulsive noise environments the proposed estimator achieves higher relative efficiency than the non-parametric one, which is based on AKDE. For the semi-parametric approach, an estimator for the parameter of the nonlinear transformation function, based on outlier detection, was proposed. Robust MUD was considered and simulation results showed

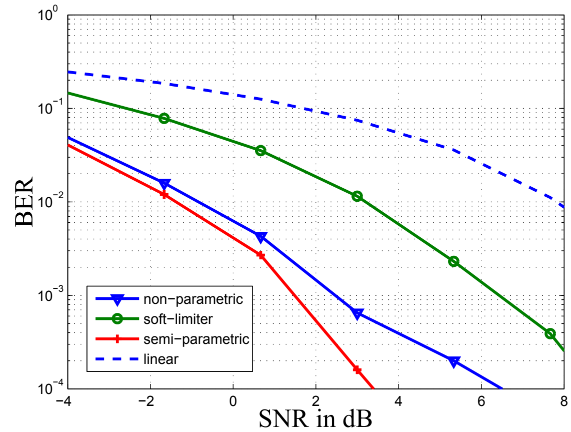

Fig. 4. BER versus SNR for user 1 in a synchronous CDMA channel with Gaussian mixture noise ($\varepsilon = 0.1$, $\kappa = 100$). N=31, K=6, NFR =10dB.

that the new method strongly outperforms conventional detectors in impulsive noise. No difference in performance between the adaptive approaches can be observed in Gaussian noise.
An extension of the semi-parametric approach with different parametric transformation functions, e.g., the Box-Cox transformation, and selection schemes for their parameters are considered for future work.

7. REFERENCES

[1] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and class B noise models," vol. 45, no. 4, pp. 1129–49, May 1999.

[2] X. Wang and H. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 289–304, Feb. 1999.

[3] P. Huber, *Robust Statistics*, Wiley, 1981.

[4] U. Hammes, C. L. Brown, R. F. Brcic, and A. M. Zoubir, "Model selection for adaptive robust parameter estimation and its impact on multiuser detection," in *VIII IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Helsinki, Finland, 2007.

[5] A. Zoubir and R. Brcich, "Multiuser detection in heavy tailed noise," in *Digital Signal Processing: A Review Journal*, April-July 2002, vol. 12(2-3), pp. 262–73.

[6] B. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, 1986.

[7] M. P. Wand, J. S. Marron, and D. Ruppert, "Transformations in density estimation," *Journal of the American Statistical Association*, vol. 86, no. 414, pp. 343–353, June 1991.

[8] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*, Wiley, 2003.