

ERROR WHITENING NON-PARAMETRIC MAXIMUM LIKELIHOOD CHANNEL ESTIMATOR

Vimal Bhatia and Bernard Mulgrew

Institute for Digital Communications,
The University of Edinburgh
Kings Buildings, Mayfield Road
Edinburgh, U.K. EH9 3JL
Email: [v.bhatia, b.mulgrew]@ee.ed.ac.uk

ABSTRACT

The presence of co-channel interference has been a major hindrance in improving the performance of present day communication systems. In this paper we discuss a iterative block based maximum-likelihood algorithm using kernel density estimates to improve channel estimation in presence of co-channel interference. As it is known that the interference is correlated, we first reduce this correlation by using a whitening filter. After whitening, we estimate this unknown whitened likelihood pdf by using kernel density estimator at the receiver. Thereby combining log-likelihood as cost function with whitening filter and kernel density estimate, a robust channel estimator for correlated noise environments is formed. The simulations for co-channel interference in presence of Gaussian noise, confirms that a better estimate can be obtained by using the proposed technique as compared to the traditional least squares algorithm, which is optimal in the Gaussian noise environments.

1. INTRODUCTION

In communication systems that experience multiple access interference (MAI) or co-channel interference (CCI) the observed noise (noise plus interference) deviates from Gaussianity [1]. Most of the present day systems are optimized under the Gaussian assumption and their performance is degraded by the occurrence of correlated non-Gaussian noise i.e. Gaussian noise in presence of coloured interference. The least squares (LS) criterion is considered optimal and is equivalent to maximum-likelihood (ML) for channel estimation when the transmitted symbols are equi-probable, unknown parameters are deterministic and the noise is additive Gaussian. However, in scenarios where the received data is not a sufficient statistics [1] or is corrupted by non-Gaussian noise, the traditional LS-based methods are inefficient and the LS estimator may not be as efficient (or equivalent) to ML estimator. From here on, in order to avoid confusion,

the traditional (Gaussian assumption) ML is referred to as the LS solution.

In this paper we take a training based channel estimator, where the channel is estimated over a block of data (similar to GSM) [1]. As discussed above due to presence of co-channel interference the observed noise at the receiver does not remain Gaussian. This degrades the performance of traditional LS based channel estimators. We developed a channel estimator in [2] which showed considerable improvement in the estimates in presence of interference. However in [2] we assumed that the input noise plus interference although correlated, could be modelled by independent and identical distribution. In this paper, we extend this earlier work by first whitening the noise plus interference and finding a ML estimate for the channel as well as the whitening filter. Techniques which whiten the noise plus interference before suppressing the interference has been proposed in [3, 4]. This forms a powerful technique to improve the performance, but since, in practice, the tap length of this whitening filter cannot be increased to a large value, the ideal assumption of white Gaussian noise (after the linear prediction error (LPE) filter [1]) does not hold. In this paper we make no such (Gaussian) assumption on the distribution of the whitened noise, which makes this technique robust to the various noise distributions.

It is observed from [1, 5, 6] that various types of noises encountered in communications can be modelled as a Gaussian mixture. In order to estimate this unknown noise pdf at the receiver we use the kernel density estimator. This is a non-parametric method of density estimation that allows the data to define the density directly. We here propose kernel density estimation based technique operating iteratively on a given block of data at each iteration. It is assumed that the corrupting noise pdf can be effectively modelled by a Gaussian mixture. We here make no *a priori* assumption on the number of Gaussian mixtures or their relative probabilities. Instead the received data is exploited using kernel

density estimators to estimate this pdf. An improvement in mean square error (MSE) performance in channel estimation over the traditional LS estimate is observed by using the proposed algorithm.

The paper is organized as follows. First, the problem statement is formulated in section-2 for a general communication system. Followed by short discussion on kernel density estimators in section-3. The non-parametric maximum likelihood algorithm using LPE filter and kernel density estimator is discussed in section-4. In section-5 simulation results are presented. Conclusions based on analysis and simulation are drawn at the end.

2. FORMULATION OF THE PROBLEM

The discrete-time model in the low-pass equivalent form of the communication system channel estimator is shown in Fig. 1. Without loss of generality, the input signal is assumed to be randomly generated binary anti-podal PAM signal, so that the transmitted symbols are $\mathbf{a} \in \{\pm 1\}$. Here 'r' represents the received signal and 'w' is a sequence of additive noise. The model is simplified by assuming that the channel is of order $N-1$ i.e. $\mathbf{h} = [h(0), h(1), \dots, h(N-1)]$.

More precisely, the received signal $r(k)$ sampled once per symbol can be written as

$$r(k) = \sum_{n=0}^{N-1} h(n)a(k-n) + w(k) \quad (1)$$

The problem is to estimate the channel coefficients from the received signal assuming that the input signal (as in supervised training mode) and the channel (tap) length is known at the receiver. Thus the problem reduces to the well known problem of system identification. There are various algorithms based on different criterions to estimate the channel taps. Usually the LS solution is taken as the optimum solution for the Gaussian noise environments where it is equivalent to a ML estimate [7]. However here we assume that the noise in presence of interference is correlated, thus LS does not provide the ML solution. To remove this correlation we use a LPE filter. The problem then reduces to the one shown in Fig. 1.

The problem can now be written as:

$$\sum_{i=0}^P z(i)r(k-i) = \sum_{l=0}^{L+P-1} a(k-l)\zeta(l) + \epsilon(k) \quad (2)$$

where $\mathbf{z} = [z(0) = 1, z(1) = -\alpha(1), \dots, z(p) = -\alpha(p)]$ are the coefficients of the LPE filter and the equivalent channel taps vector $\boldsymbol{\zeta} = [\zeta(0)\zeta(1)\dots\zeta(L+P-1)]$, where $\zeta(l) = \sum_i z(i)h(l-i)$. Ideally the $\epsilon(k)$ is a zero-mean white Gaussian process. Since from eq. (1), the model eq. (2) corre-

sponds to assuming

$$\sum_{i=0}^P z(i)r(k-i) = \sum_{i=0}^P z(i) \sum_{l=0}^{L-1} a(k-i-l)h(l) = \sum_{l=0}^{L+P-1} a(k-l)\zeta(l) \quad (3)$$

and

$$\sum_{i=1}^P z(i)w(k-i) = \epsilon(k) \quad (4)$$

Therefore, the effect of the LPE filter is that of whitening the additive disturbance $w(k)$. The formulation eq. (2) permits the description of the channel plus the whitening filter as a vector inner product, which in turn allows the simultaneous estimation of the LPE coefficients and the equivalent channel taps at the output of the LPE filter [3]. In fact, letting $\boldsymbol{\alpha} = [\alpha(1)\dots\alpha(p)]$, eq. (2) can be rewritten as

$$[1 - \alpha(1)\dots - \alpha(p)] \begin{bmatrix} r(k) \\ r(k-1) \\ \vdots \\ r(k-P) \end{bmatrix} = [\zeta(0)\zeta(1)\dots\zeta(L+P-1)] \begin{bmatrix} a(k) \\ a(k-1) \\ \vdots \\ a(L+P-1) \end{bmatrix} + \epsilon(k) \quad (5)$$

or equivalently

$$r(k) = [\alpha(1)\dots\alpha(p)] \begin{bmatrix} r(k-1) \\ \vdots \\ r(k-P) \end{bmatrix} + [\zeta(0)\zeta(1)\dots\zeta(L+P-1)] \begin{bmatrix} a(k) \\ a(k-1) \\ \vdots \\ a(L+P-1) \end{bmatrix} + \epsilon(k) \quad (6)$$

$$= [\boldsymbol{\alpha} \ \boldsymbol{\zeta}] \mathbf{v}(k) + \epsilon(k) \quad (7)$$

where $\mathbf{v}(k) = [r(k-1)\dots r(k-P)b(k-1)\dots b(k-L-P+1)]^T$ and with $\epsilon(k)$ white. Usually it is assumed that due to LPE filter $\epsilon(k)$ is Gaussian distributed. However in practice the larger the tap-length of the LPE filter, the more the equalizer states, the more computational complexity [1] for maximum likelihood sequence equalizer. Thus in practice the tap-length is usually restricted [1, 4]. This restriction leads $\epsilon(k)$ to non-Gaussianity. The channel estimator proposed in this paper does two tasks: (i) estimating the channel; (ii) estimating the uncorrelated $\epsilon(k)$ pdf at the receiver.

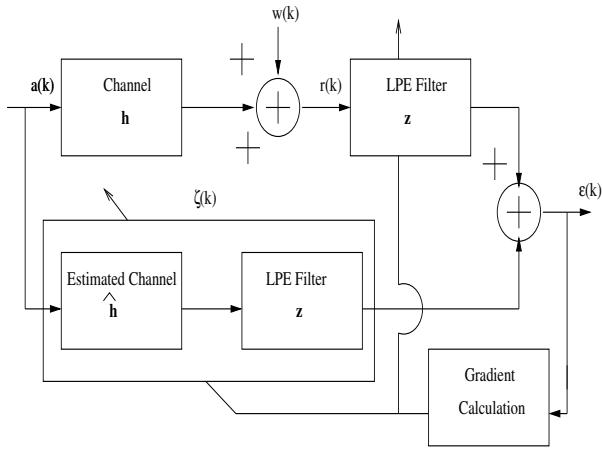


Fig. 1. Communication systems channel estimator with LPE filter

3. KERNEL DENSITY ESTIMATION

To estimate the pdf at the receiver we use the kernel density estimator technique. Parzen window or kernel density estimation assumes that the probability density is a smoothed version of the empirical sample. Its estimate $\hat{f}(y)$ of a random variable Y is simply the average of radial kernel functions centered on M -realizations of Y :

$$\hat{f}(y) = \frac{1}{M} \sum_{j=1}^M \phi(y - y(j)) \quad (8)$$

We here assume ϕ to be Gaussian kernel (Parzen kernel) [1]:

$$\phi(y) = \mathcal{N}(0, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad (9)$$

where σ^2 is defined as the kernel variance (or width) [8]. Other choices of kernel like *Epanechnikov kernel* are also possible. It can be shown that under the right conditions $\hat{f}(y)$ will converge to the true density $f(y)$ as $|M| \rightarrow \infty$ [9].

4. NON-PARAMETRIC MAXIMUM-LIKELIHOOD (NPML) CHANNEL ESTIMATION WITH LPE

For the communication system represented by eq. (2) the ML estimate forms the optimal estimator for the channel. This problem can be viewed as the joint optimization problem [1], where we maximize the likelihood with respect to α and ζ . If the $\epsilon(k)$ was Gaussian then the LS solution as found in [1] could have been applied directly. However, since we assume that $\epsilon(k)$ is non-Gaussian and can be modelled as a Gaussian mixture we use the kernel density estimator to estimate this density. Since the kernel density

estimator is essentially a Gaussian mixture formulation we can't get a closed form estimate of the α and ζ . We then use the iterative scheme as used in [2]:

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \mu(k) \nabla_{\alpha} \mathcal{L}(\alpha | \mathbf{r}, \zeta) \Big|_{\alpha=\hat{\alpha}_{k-1}, \zeta=\hat{\zeta}_{k-1}} \quad (10)$$

$$\hat{\zeta}_k = \hat{\zeta}_{k-1} + \mu(k) \nabla_{\zeta} \mathcal{L}(\zeta | \mathbf{r}, \alpha) \Big|_{\zeta=\hat{\zeta}_{k-1}, \alpha=\hat{\alpha}_{k-1}} \quad (11)$$

where $\mu(k)$ is the adaptation step-size. Since the channel estimator is assumed to have no *a priori* knowledge of the pdf $f_{\epsilon}(\cdot)$, this unknown pdf is then estimated by using the kernel density estimator eq. (8) with Gaussian kernels as shown below. As the kernel estimators are known to be effective in density estimation over short data record, we use this technique over the available data (error) record, of length M , to estimate the unknown density. Using the kernel density estimator [9] we obtain:

$$\hat{f}_{\epsilon}(\epsilon) = \frac{1}{M} \sum_{j=1}^M K(\epsilon - \epsilon(j)) \quad (12)$$

Thus the estimated (joint) log-likelihood function can be written as:

$$\begin{aligned} \hat{\mathcal{L}}(\alpha, \zeta | \mathbf{r}) \Big|_{\alpha=\hat{\alpha}_{k-1}, \zeta=\hat{\zeta}_{k-1}} &= \sum_{i=1}^M \log \left(\frac{1}{M} \sum_{j=1}^M K(\epsilon(i) - \epsilon(j)) \right) \\ &= \sum_{i=1}^M \log \sum_{j=1}^M K(\epsilon(i) - \epsilon(j)) - \log |M| \end{aligned} \quad (13)$$

The gradient $\hat{\alpha}$ of the log-likelihood can be formulated as:

$$\begin{aligned} \nabla_{\alpha} \hat{\mathcal{L}}(\alpha | \mathbf{r}, \zeta) \Big|_{\alpha=\hat{\alpha}_{k-1}, \zeta=\hat{\zeta}_{k-1}} &= \frac{\partial}{\partial \alpha} \hat{\mathcal{L}}(\alpha | \mathbf{r}, \zeta) \Big|_{\alpha=\hat{\alpha}_{k-1}, \zeta=\hat{\zeta}_{k-1}} \\ &= \sum_{i=1}^M \frac{\sum_{j=1}^M \frac{\partial}{\partial \alpha} K(\epsilon(i) - \epsilon(j))}{\sum_{k=1}^M K(\epsilon(i) - \epsilon(k))} \end{aligned} \quad (14)$$

Similarly gradient for $\hat{\zeta}$:

$$\begin{aligned} \nabla_{\zeta} \hat{\mathcal{L}}(\zeta | \mathbf{r}, \alpha) \Big|_{\zeta=\hat{\zeta}_{k-1}, \alpha=\hat{\alpha}_{k-1}} &= \frac{\partial}{\partial \zeta} \hat{\mathcal{L}}(\zeta | \mathbf{r}, \alpha) \Big|_{\zeta=\hat{\zeta}_{k-1}, \alpha=\hat{\alpha}_{k-1}} \\ &= \sum_{i=1}^M \frac{\sum_{j=1}^M \frac{\partial}{\partial \zeta} K(\epsilon(i) - \epsilon(j))}{\sum_{k=1}^M K(\epsilon(i) - \epsilon(k))} \end{aligned} \quad (15)$$

Thereby substituting the estimated gradients in eq. (10) and (11) respectively, and iterating till $\hat{\alpha}_k$ and $\hat{\zeta}_k$ converge we get the ML estimated channel $\hat{\mathbf{h}}$ by deconvolution. The algorithm is initialized by the LS estimate and μ as explained in [1] and [2] respectively. During simulations we

did not converge to a local maxima, however this is not always guaranteed (as with most iterative techniques). However, initializing the channel and whitening filter's taps with the LS estimate reduces the chances of converging to local maxima.

Two possible update methods could be used here to maximize the likelihood. In the first method, we update $\hat{\alpha}$ then update $\hat{\zeta}$ and then estimate the updated likelihood for the next update. The procedure is repeated till both $\hat{\alpha}$ and $\hat{\zeta}$ converge. Second update method is to first have $\hat{\alpha}$ converged and then have $\hat{\zeta}$ converged given the converged $\hat{\alpha}$ on estimated likelihood, and then iterate till no significant change is observed in $\hat{\alpha}$ and $\hat{\zeta}$. In this paper we use the first method of update.

5. SIMULATION RESULTS

For simulation study, we assume a communication channel model, like global system for mobile (GSM), considering CCI with Gaussian noise as a multi-modal, iid, Gaussian mixture interference as discussed in [1]. The performance of channel estimator is calculated by normalized-mean square error (NMSE), as shown in eq.(16).

$$NMSE = \frac{\mathbb{E}\{(h - \hat{h})^2\}}{\mathbb{E}\{h^2\}} \quad (16)$$

where h is the actual channel and \hat{h} is the estimated channel (after deconvolution). For all simulation results, the input symbols of length 100 and ensemble of 1000-runs is considered.

A typical communication system effected by co-channel interference is shown in Fig. 2. The co-channels are each of order $N - 1$ and are represented as h_i and interfering signal as a_i for $i = 2, \dots, I$, where $I - 1$ represents number of interferers. The received signal can be represented as

$$\begin{aligned} r(k) &= \sum_{n=0}^{N-1} h_1(n)a_1(k-n) + \sum_{i=2}^I \sum_{n=0}^{N-1} h_i(n)a_i(k-n) + n(k) \\ &= \sum_{n=0}^{N-1} h_1(n)a_1(k-n) + w(k) \end{aligned} \quad (17)$$

where the middle (double summation) term on the RHS in eq. (17) represents the CCI and $n(k)$ is a zero mean, iid, Gaussian noise process and $k = 1, \dots, M$ represents the number of symbols.

The above presented algorithm is verified for real stationary channel for $N = 5$. The input signal is anti-podal random input sequence. The channels are assumed to be

$$h_1 = [-0.227 \ 0.460 \ 0.688 \ 0.460 \ -0.227] \text{ and}$$

$h_2 = [1.0 \ 0.8 \ 0.6 \ 0.4 \ 0.2]$ where h_1 suffers from amplitude and phase distortion [10], and h_2 is the co-channel considered for the simulation.

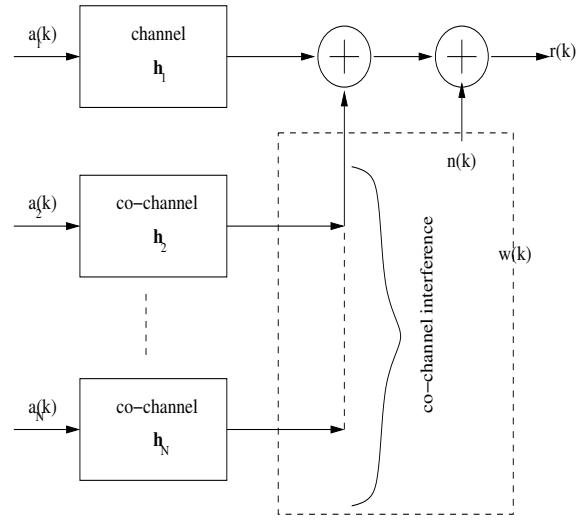


Fig. 2. A typical CCI effected communication system

Fig. 3 depicts the performance plot for the channel estimator presented in this paper. The legends 'LS', 'LS_{white}' and 'NPML_{white}' represent LS without LPE filter, LS with LPE filter and NPML with LPE filter respectively. To observe the performance of the algorithm, the signal to noise ratio (SNR) is kept fixed at 30dBs while signal to interference ratio (SIR) is varied from -13dBs to 9dBs. We can observe that by using the LPE filter with NPML based technique we can gain upto 3.5 dBs at NMSE of 10^{-1} .

6. CONCLUSION

It was shown that after noise whitening better channel estimates can be obtained. It was reconfirmed that the LS estimate with LPE filter produces better channel estimates for interference limited channels than LS estimate without LPE filter. Due to practice constraints, the Gaussian assumption on the whitened noise is not guaranteed, hence a kernel density estimate based ML channel estimator was proposed. From Fig. 3 we observe that better channel estimates can be obtained by jointly estimating the whitening filter and the channel estimates by using kernel density estimator. Thus by combining kernel density estimator with whitening filter forms a robust channel estimator for interference limited communication channels.

Acknowledgment

The authors would like to thank UK-EPSC for supporting this work.

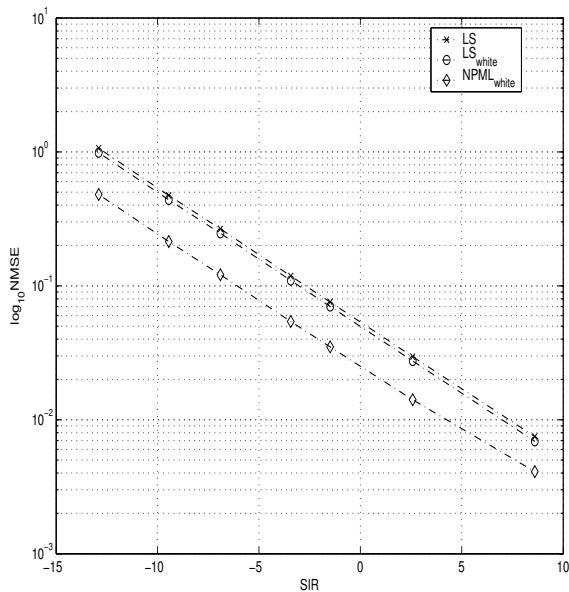


Fig. 3. MSE plot for co-channel effected communication system where $h_1 = [-0.227 \ 0.460 \ 0.688 \ 0.460 \ -0.227]$, SNR=30dBs for 100-symbols over an ensemble of 1000-runs

7. REFERENCES

- [1] C. Luschi and B. Mulgrew, "Nonparametric trellis equalization in the presence of non-Gaussian interference," *IEEE Transactions on Communications*, vol. 51, no. 2, pp. 229–239, Feb 2003.
- [2] V. Bhatia and B. Mulgrew, "A EM-based channel estimator in non-Gaussian noise," in *VTC-2004 (Fall), LA, US*, Sept 2004.
- [3] D. Astely and B. Ottersten, "MLSE and spatio-temporal interference rejection combining with antenna arrays," in *Proc. EUSIPCO*, Sept 1998, vol. III, pp. 1341–1344.
- [4] J-Wei Liang, Jiunn-Tsair Chen, and A. J. Paulraj, "A two-stage hybrid approach for CCI/ISI reduction with space time processing," *IEEE Communications Letters*, vol. 1, no. 6, pp. 163–165, Nov 1997.
- [5] E.E. Kuruoglu, C. Molina, and W.J. Fitzgerald, "Approximation of alpha-stable probability densities using finite gaussian mixtures," in *Proceedings of EUSIPCO 98, Signal Processing IX: Theories and Applications*, Sept 1998, vol. 2, pp. 989–992.
- [6] R.J. Kozick, R.S. Blum, and B.M. Sadler, "Signal processing in non-Gaussian noise using mixture distributions and the EM algorithm," in *Thirty-First Asilomar*

Conference on Signals Systems and Computers, Nov 1997, vol. 1, pp. 438–442.

- [7] J. K. Tugnait, Tong Lang, and Ding Zhi, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 17–28, May 2000.
- [8] B. W. Silverman, Ed., *Density Estimation for Statistics and Data Analysis*, Chapman Hall, London, 1986.
- [9] N. N. Schraudolph, "Gradient-Based Manipulation of Non-parametric Entropy Estimates," *IEEE Transaction on Neural Networks*, vol. 15, no. 4, Jul 2004.
- [10] J. G. Proakis, Ed., *Digital Communications*, McGraw-Hill, 1989.