A BOOTSTRAP MULTI-USER DETECTOR FOR CDMA BASED ON TIKHONOV REGULARIZATION

*Christos Kasparis, *Robert J. Piechocki, **Paul N. Fletcher, *Andrew R. Nix

*University of Bristol, Bristol, U.K.
E-mail : c.kasparis@bris.ac.uk

**QinetiQ Ltd., Malvern, U.K.
E-mail : paul.fletcher@bris.ac.uk

ABSTRACT

A novel Multi-User Detection (MUD) technique for CDMA systems is introduced. The new method is well suited for cases in which the code cross correlation matrix is ill -conditioned. In practice this case coincides with having code lengths approximately equal to the number of users – a desirable condition in terms of bandwidth efficiency. Our approach employs an iterative formulation of a well-known regularization method for linear inverse problems, which is suited to the MUD problem. The technique allows knowledge of the finite set in which the solution belongs to be exploited in a computationally efficient manner in order to iteratively improve the quality of the estimate.

1. INTRODUCTION

Since the introduction of the optimal Maximum-Likelihood (ML) detector (see [3]) a number of lower complexity classes of detectors have been proposed, attempting to approach the ML performance in a computationally and bandwidth efficient manner. Examples of those detector classes are the linear detectors (e.g. MMSE and Decorrelator), the linear and non-linear Interference Cancellation (IC) based detectors (e.g. Serial and Parallel IC, Decision feedback, Turbo MUD etc.) and the Subspace Based Linear detectors. A fundamental difference between the ML detector and the rest is the explicit use of the prior knowledge regarding the solution set, which however results in solving an NP-hard optimization problem.

The proposed detector is based on a well established method for solving ill-posed linear inverse problems, namely Tikhonov Regularization[1] (TR). The fundamental idea in TR is to introduce to the Least Squares (LS) optimization criterion some additional side constraints, which the desired solution needs to fulfill. Those constraints need to be chosen carefully not only to be meaningful but also to be simple enough to allow an analytical solution to the problem. In section 3 we demonstrate that the MMSE detector is a special case of the TR solution to the MUD problem with a particular choice of the regularization parameter.

The clear distinction between the various constraints that a TR solution needs to satisfy gives to the method more design flexibility but also increased difficulties in terms of determining optimal (in some sense) regularization parameters. The new detector is based on a particular formulation of the TR problem in which some prior solution to the problem is assumed to be known. In contrast to other detectors, which make use of prior information, the proposed one does not rely on any external devices. Instead it uses its own initial estimate in conjunction with knowledge of the finite solution set in order to feedback a proposal solution, closer to the true one, and re-solve the TR problem. This procedure is repeated iteratively, giving increased importance to the fed-back solution in each iteration.

2. FORMULATION OF THE BASIC MUD PROBLEM

In this section a simplified multiuser CDMA model is developed and the basic MUD problem is formulated.

A discrete time baseband model of a synchronous DS-CDMA system is considered where U users are active. Each of the $u_i\,(1 \le i \le U)$ users transmits a BPSK modulated bit b_i after this has been spread by some spreading sequnece $\underline{s}_i = \sum_{n=1}^{N} s_n^i \cdot \delta[(k-n)\cdot T_c]$ (of length N) where k is a discrete unit delay variable, T_c is chip period and $\delta[k] = \begin{cases} 1, k = 0 \\ 0, k \neq 0 \end{cases}$. The energy in each of the code sequences is chosen and normalized so that

[1] Also known as Ridge Regression in the statistical literature

$\underline{s}_i \cdot \underline{s}_j^T = \begin{cases} 1 & ,i = j \\ \rho_{ij} < 1 & ,i \neq j \end{cases}$. Assuming an AWGN channel and equal power for each user then the received sequence can be expressed as:

$$\underline{r} = \sum_{i=1}^{U} \sqrt{E_b} \cdot b_i \cdot \underline{s}_i + \underline{n} \qquad (1)$$

where $\sqrt{E_b}$ is the energy in each modulated symbol and $\underline{n}$ is some sampled (in chip rate) realization of a white Gaussian process with zero mean and covariance matrix $\sigma^2 \cdot \underline{\underline{I}}_N$. For simplicity the problem will be formulated involving only with real variables. Extensions for vectors defined in complex spaces is straight forward.

The received signal is despread by cross-correlating $\underline{r}$ with each of the code sequences giving samples $y_i = \underline{r} \cdot \underline{s}_i^T$ each in which the contribution in energy from user u_i is dominant. Nevertheless contributions from the rest of the users are still significant and a detection scheme based directly on these samples is characterised by very poor performance. Instead of treating multiuser interference as some increase in the noise power (compared to the single user case), MUD techniques attempt to reject the interference by processing all the samples $\underline{y} = [y_1, y_2, .., y_U]$ collectively. In particular MUD involves solving an inverse linear problem of the form:

$$\underline{\underline{R}} \cdot \underline{b}^T = \underline{y}^T = \underline{z}^T + \underline{v}^T \qquad (2)$$

where $\underline{\underline{R}}$ is the $U \times U$ code cross-correlation matrix which is symmetric and positive definite, $\underline{b} = [b_1, b_2, ..., b_U]$ which needs to be determined, $\underline{z}$ is the same as $\underline{y}$ for the case where there is no thermal noise present and $\underline{v}^T$ is a sampled (in symbol level) realisation of Gaussian process with zero mean and covariance matrix $\sigma^2 \cdot \underline{\underline{R}}$.

2.1. Brief Review of Linear Multiuser Detectors

The simplest linear detection technique solves the Generalised Least Squares (GLS) problem by assuming that the solution belongs to an infinite space and minimizing the cost function:

$$J_{GLS}(\underline{b}) = \left\| \underline{\underline{C}}^{-1} (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T) \right\|_2^2 \qquad (3)$$

where $\underline{\underline{C}}$ is given by the Cholesky decomposition of $\underline{\underline{R}}$ ($\underline{\underline{R}} = C \cdot C^T$). The solution is the well-known decorrelating detector:

$$\hat{\underline{b}}_{GLS}^T = \underline{\underline{R}}^{-1} \cdot \underline{y}^T \qquad (4)$$

In cases where $\underline{\underline{R}}$ is ill-conditioned, the GLS solution is very poor even in very good SNR conditions, as noise is severely amplified in the directions of the singular vectors (of $\underline{\underline{R}}$), associated to small singular values. In that case, the MMSE detector offers a much more reliable solution, as it effectively dampens components of the solution in these directions. The linear MMSE detector can be derived by finding the matrix $\underline{\underline{M}}$ which minimises:

$$J_{MMSE}(\underline{\underline{M}}) = E\left\{ \left\| (\underline{b}^T - \underline{\underline{M}} \cdot \underline{y}^T) \right\|_2^2 \right\} \qquad (5)$$

where $E\{.\}$ is the statistical expectation operator. The MMSE solution to the problem (derived from (5)) is given as:

$$\hat{\underline{b}}_{MMSE}^T = (\underline{\underline{R}} + \sigma^2 \cdot I)^{-1} \cdot \underline{y}^T \qquad (6)$$

Again the knowledge of the solution set is ignored and $\underline{b}$ is assumed to take values in $\Re^U$.

3. TIKHONOV REGULARIZATION

Tikhonov Regularization is a particular type of Regularization method for linear inverse problems with the attractive feature, from a computational point of view, that no decomposition of any kind (e.g. Singular Value Decompostion (SVD), QR decomposition) needs to be performed on the design matrix ($\underline{\underline{R}}$ in the MUD case). Moreover, in general (but not necessarily) decomposition based methods are better suited for cases in which there is a distinct 'jump' in the magnitude of the singular values, indicating some effective rank. TR on the other hand is better suited for cases in which there is a smooth decay in the spectrum of the design matrix. In MUD the latter case typically holds.

TR, adds to the classical LS constraint an additional regularization constaint :

$$J_{TR}(\underline{b}) = \left\| (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T) \right\|_2^2 + \Omega(\underline{b}) \qquad (7)$$

The side constaint helps to narrow down the set of possible solutions which satisfy the LS constaint, provided the former is consistent with the problem. The side constaint also needs to be a sufficiently simple criterion if an analytical solution to (7) is required. A general choice for $\Omega(\underline{b})$ which proves to be meaningful in many problems and is also simple enough to provide an analytical solution is the one proposed by Tikhonov (see [1], [2]):

$$\Omega(\underline{b}) = \lambda^2 \cdot \left\| \underline{\underline{L}} \cdot \underline{b}^T \right\|_2^2 \tag{8}$$

$\underline{\underline{L}}$ is some linear operator acting on the solution. In many problems $\underline{\underline{L}}$ represents the discrete derivative (of some order) operator in which case the solution is known to fulfill certain smoothness conditions. λ^2 is a smoothing regularization parameter whose value dictates the smoothness on the filtering function which is imposed on the spectrum of the design matrix by the Regularization constraint. Obviously, as $\lambda^2 \to 0$ no weighting is imposed on the singular values of $\underline{\underline{R}}$ and the TR solution coincides with the LS one. On the other hand as $\lambda^2 \to \infty$, an exessively smooth function is applied on the spectrum of $\underline{\underline{R}}$ and information about the solution in the observation is lost in the attempt to over supress noise. Optimal selection of λ^2 is not a trivial task in practice and many methods have been proposed in the literature (e.g. Cross Validation, Generalised Cross Validation, Graphical methods, etc.).

In the case where knowledge about an initial-default solution $\tilde{\underline{b}}$ is known for the problem then (8) can be further generalized as:

$$\Omega(\underline{b}) = \lambda^2 \cdot \left\| \underline{\underline{L}} \cdot (\underline{b}^T - \tilde{\underline{b}}^T) \right\|_2^2 \tag{9}$$

In this case λ^2 controls the bias in the estimator towards the default solution. As $\lambda^2 \to \infty$ the estimator will coincide with the default solution and no useful information will be extracted from the observations. This is not necessarily bad as the default solution might already be close to the true one (in which case the observation will have a negative effect on the quality of the estimator) although it is not easy to verify that in practice.

Starting from the general formulation of TR criterion :

$$J_{TR}(\underline{b}) = \left\| (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T) \right\|_2^2 + \lambda^2 \cdot \left\| \underline{\underline{L}} \cdot (\underline{b}^T - \tilde{\underline{b}}^T) \right\|_2^2 \tag{10}$$

A solution can be found by setting

$$\frac{\partial}{\partial b_i}\left\{\lambda^2 \cdot (\underline{b}^T - \tilde{\underline{b}}^T)^T \underline{\underline{L}}^T \underline{\underline{L}} (\underline{b}^T - \tilde{\underline{b}}^T) + (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T)^T (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T)\right\} = 0 \tag{11}$$

which leads to the following solution:

$$\begin{aligned} &2 \cdot \lambda^2 \cdot \underline{\underline{L}}^T \underline{\underline{L}} (\underline{b}^T - \tilde{\underline{b}}^T) - 2 \cdot \underline{\underline{R}}^T \cdot (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T) = 0 \Rightarrow \\ &\hat{\underline{b}}_{TR} = (\lambda^2 \cdot \underline{\underline{L}}^T \underline{\underline{L}} + \underline{\underline{R}}^T \underline{\underline{R}})^{-1} \cdot (\lambda^2 \cdot \underline{\underline{L}}^T \underline{\underline{L}} \cdot \tilde{\underline{b}}^T + \underline{\underline{R}}^T \cdot \underline{y}^T) \end{aligned} \tag{12}$$

We observe that the Tikhonov estimator resembles strongly the structure of the MMSE one. Indeed by setting $\underline{\underline{L}} = \underline{\underline{I}}$, $\tilde{\underline{b}} = 0$ and $\lambda^2 = \sigma^2$ we get the MMSE solution for the case when the noise is zero mean with covarinance $\sigma^2 \cdot \underline{\underline{I}}$ and uncorrelated with $\underline{b}$. In a similar fashion the MMSE multiuser detector can be derived from the TR criterion by solving a slightly modified problem

$$J_{TR}(\underline{b}) = \left\| \underline{\underline{C}}^{-1} \cdot (\underline{\underline{R}} \cdot \underline{b}^T - \underline{y}^T) \right\|_2^2 + \sigma^2 \cdot \left\| \underline{b}^T \right\|_2^2 \tag{13}$$

in order to take into account the non whiteness of the noise. The solution of (13) is given by:

$$\begin{aligned} \hat{\underline{b}}_{TR} &= (\sigma^2 \cdot \underline{\underline{I}} + \underline{\underline{R}}^T (\underline{\underline{C}} \cdot \underline{\underline{C}}^T)^{-1} \underline{\underline{R}})^{-1} \cdot \underline{\underline{R}}^T \cdot (\underline{\underline{C}} \cdot \underline{\underline{C}}^T)^{-1} \cdot \underline{y}^T = \\ &= (\sigma^2 \cdot \underline{\underline{I}} + \underline{\underline{R}})^{-1} \cdot \underline{y}^T \end{aligned} \tag{14}$$

which is exactly the MMSE detector. This result indicates that choosing $\lambda^2 = \sigma^2$ for the particular problem is a good choice. This is an important observation as far as producing an initial estimate in the proposed detector is concerned, as it bypasses the need for finding a good Regularization (or Ridge) parameter, which usually involves solving a non-trivial optimization problem (e.g. Generelized Cross Validation).

4. BOOTSTRAP MUD BASED ON TIKHONOV REGULARIZATION

The proposed Multiuser detector is based on the formulation of the TR criterion in which $\underline{\underline{L}} = \underline{\underline{I}}$ but some default solution is assumed to be known about the problem. In the initial iteration no such solution is known so the detector reduces to the MMSE one. As soon as some initial estimate is available, we use the fact that the distribution of each estimated symbol is well approximated by a Gaussian distribution [4] in order to make hard decisions only for symbols which satisfy some posterior probability of error criterion. Those, which lie outside the required decision boundaries, are left unchanged. So a proposal solution vector is constructed which consists of both soft and hard estimates, the latter of which we are confident about their correctness. The estimator used in each iteration is given by:

$$\hat{\underline{b}}_{TR} = (\lambda^2 \cdot \underline{\underline{I}} + \underline{\underline{R}})^{-1} \cdot (\lambda^2 \cdot \tilde{\underline{b}}^T + \underline{y}^T) \quad (15)$$

where $\tilde{\underline{b}}$ is the semi-hard proposal solution to the problem. The effect of this procedure is to iteratively limit the variance in the estimation error while at the same time bias the estimator towards the correct solution. Simulation results (see section 5.) have shown that when this method is applied to the MUD problem, it provides a significant performance gain especially if in each iteration we show increased belief in the proposal solution by increasing the value of λ^2. From a computational point of view the latter is not very bad news as $(\lambda^2 \cdot \underline{\underline{I}} + \underline{\underline{R}})^{-1}$ can be efficiently recomputed by decomposing $\lambda^2 \cdot \underline{\underline{I}} = \lambda^2 \sum_{i=1}^{U} \underline{w}_i^T \cdot \underline{w}_i$, (where $\underline{w}_i$ is the all zero vector except for the i^{th} element which is 1) and applying iteratively (U times) the matrix inversion lemma:

$$(\underline{\underline{H}} + \underline{x}^T \cdot \underline{x})^{-1} = \underline{\underline{H}}^{-1} - \frac{\underline{\underline{H}}^{-1} \cdot \underline{x}^T \cdot \underline{x} \cdot \underline{\underline{H}}^{-1}}{1 + \underline{x} \cdot \underline{\underline{H}}^{-1} \cdot \underline{x}^T}.$$

5. SIMULATION RESULTS

The proposed Bootstrap detector has been simulated for the simple MUD problem presented in section 2. Figure 1 illustartes the Bit Error Rate (BER) versus received SNR. Random spreading codes have been used and $U = N = 50$, which results in a significantly ill-conditioned $\underline{\underline{R}}$ indicated by the very poor performance of the decorrelator detector and the moderately bad performance of the MMSE detector. The single user performance which is the target performance is also plotted. We also give the decorrelator and MMSE performance for the case when $N = 2 \cdot U = 100$ which results in a well conditioned $\underline{\underline{R}}$, indicated by the good performance of both the decorrelator and MMSE detectors. In a real system this scenario would translate into doubling the required bandwidth which is not desirable. For the regularization parameter we have chosen $\lambda^2 = \sigma^2$ for the zeroth iteration (MMSE detector) and we have introduced a weighting function for subsequent iterations so that $\lambda_k^2 = k^a \cdot \sigma^2$, for number of iterations $k \geq 1$ with $a = 1$. There is a trade-off involved in the selection of the weighting function between convergence speed and BER performance. Obviously with big weightings the information in the data will be ignored and the estimator will quickly converge to some value.

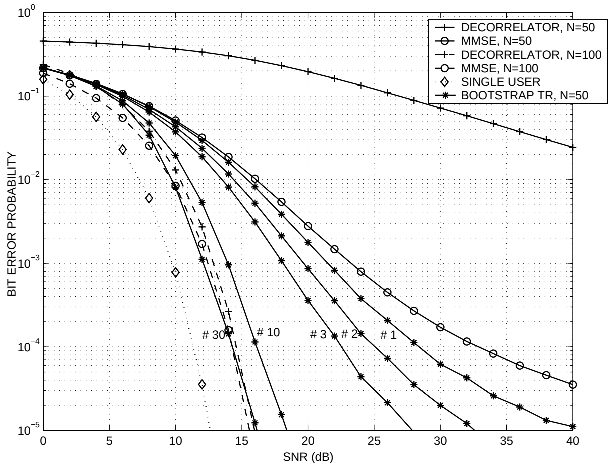

Figure 1 : BER Vs. SNR Performance Evaluation of the Bootstrap Detector

6. CONCLUSIONS

A new type of MUD for CDMA has been presented which is based on a particular formulation of the Tikhonov Regularization criterion where an initial solution to the problem is assumed known. In the proposed method this solution does not originate from some external source but from a sensible use of knowledge about the solution finite space and some initial estimate. The method is in principle applicable to many other communication and engineering linear inverse problems, in which the solution space is finite and known and also some knowledge about the statistics of the solution is available.

REFERENCES

[1] A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston, Washington, D.C., 1977.

[2] P.C. Hansen, *Rank deficient and discrete ill-posed problems,* SIAM monographs on mathematical modeling and computation, Philadelphia, 1998.

[3] S. Verdu, *Multiuser Detection*, Cambridge Press, Cambridge UK, 1998.

[4] H.V. Poor, S. Verdu, "Probability of error in MMSE multiuser detection", IEEE Trans. Information Theory, 45: pp 858-871, May 1997.

ACKNOWLEDGEMENTS

The first author would like to thank QinetiQ Ltd. and EPSRC for funding his research at the University of Bristol.