Hebbian Learning Based Blind Adaptive Multiuser Detection in DS-CDMA Systems

Malay Gupta  
Department of Electrical Engineering  
Southern Methodist University  
Dallas, TX 75275-0338  
Email: malay@engr.smu.edu

Balu Santhanam  
Department of Electrical & Computer Engineering  
University of New Mexico  
Albuquerque, NM 87131-0001  
Email: bsanthan@ece.unm.edu

Abstract- Nonlinear Hebbian learning rules have been instrumental in independent component analysis (ICA) and blind source separation (BSS) problems. We show that in a CDMA system, where the users can be assumed independent, application of these rules give accurate estimates of user signature codes and data from the observations alone. However, the solutions are obtained within a sign and permutation ambiguity. A novel method is then proposed for the resolution of these ambiguities. It is shown via numerical examples that this technique is robust in near-far situations and can be used to estimate various other parameters such as user amplitudes and cross-correlation values. Computer simulations indicate improvements in terms of reduced error probability in detection and increased signal to interference plus noise ratio (SINR). Convergence of the proposed detector is also comparable to that of the blind minimum output energy (MOE) detector.

I. INTRODUCTION

Correlation between the user signature codes introduces multiple access interference (MAI) in a direct sequence-code division multiple access (DS-CDMA) system. Power imbalance between different users worsens the situation and gives rise to the near-far problem [1]. Conventional detector is not near-far resistant and exhibits serious degradation of performance with signature code correlation and absence of power control. Multiuser detection [1] techniques attempt to mitigate these problems and can substantially increase the capacity of a CDMA system.

Optimum multiuser detector was proposed in [2], however the computational complexity of this detector grows exponentially with the number of active users in the channel. Several suboptimum detector have since then proposed to achieve performance close to that of the optimum detector, while keeping the complexity to a manageable level. Some important detectors proposed include, the decorrelating detector [3]; the decision feedback detector [4]; the minimum mean squared error (MMSE) detector [5]; and the multistage detectors [6].

The decorrelating detector [3] is centralized and noniterative and for the detection of a single user it requires the knowledge of signature codes and timings of all the active users in the channel, the decision-feedback detector [4] bears the risk of error propagation, the MMSE detector [5] requires the transmission of training data, and the multistage detector [6] involves long decoding delays.

Blind multiuser detection algorithms are attractive because they need minimal information about the structure of MAI and obviate the dependence on the training symbols which are clearly a waste of precious bandwidth. In [7] a blind multiuser detector based on constrained minimization of output energy (MOE) was proposed. In [8], blind algorithms based on signal subspace are investigated and two algorithms which converge to the decorrelating detector and the MMSE detector are proposed. This, however requires a two stage approach, first finding an estimate of the subspaces, and second the construction of the detector using estimated subspaces.

It is to be noted that above mentioned techniques make use of only the second order statistics (SOS) of the data. SOS based approaches, although computationally attractive, do not exploit the information resident in higher order statistics (HOS) of the data. Hebbian learning based ICA [9] is a recent statistical technique that exploits the nongaussianity of the underlying independent sources to make the observations independent or as independent as possible. In a communication system, it is reasonable to assume the independence of the underlying sources. It can also be shown that most communication sources and their filtered outputs follow subgaussian distributions. Hence, the key assumption of nongaussianity and independence can be exploited to improve upon the existing SOS based multiuser detectors.

ICA solutions are however, obtained up to a permutation and scaling ambiguity [10] that needs to be resolved before ICA technique can be used in multiuser detection. Previous ICA based approaches have relied on a priori knowledge of desired user’s signature code and timing to develop constrained solutions [11], [12]. Another multiuser detection approach was suggested in [13] where ICA was used as an add to the SOS based detector.

In this paper, we develop an ICA based adaptive detector which uses Hebbian learning to obtain the separating matrix for all the active users in the channel. From this separating matrix an estimate of the user signature codes is obtained and based on a priori knowledge of user signature codes the permutation and scaling ambiguities are resolved in ICA.
solution. It is shown in the paper that this method can be used for the extraction of a single user by using the same amount of information as required by the MOE and the conventional detector and can be implemented for the joint multiuser detection at the base station where information about other users in the channel can also be assumed. It will be further shown in the paper that this method can be used to obtain estimates of the amplitudes of active users in the channel as well as the correlation coefficients between them. Simulation results illustrate the superiority of the proposed method over the previously proposed methods.

II. SYSTEM MODEL

While the developments of this work can be extended to the case of an asynchronous CDMA system, we present our results in the theoretical context of synchronous CDMA system because of simplicity and clarity of presentation. First of all, it is well known and understood that every asynchronous system can be modeled as an equivalent synchronous one with higher effective multiuser population [14]. For more comprehensive treatment reader may refer to [5], which shows that in the worst case scenario every asynchronous interferer acts as two synchronous ones.

In this paper, we take the signal model involving $K$ sources, transmitting data sequences $b_k(i)$, where $1 \leq k \leq K$ and $i$ is the discrete time index. We assume that the sources are mutually independent. This is a strong hypothesis but very plausible in practice for physically separated sources. Independence assumption states that the joint pdf of sources is a product of marginal pdfs of individual sources. Let us denote the sources by a column vector, i.e., $b(n) = [b_1, b_2, ..., b_K]^T$. The joint pdf $r(b)$ of $b(n)$ can then be written as

$$r(b) = \prod_{k=1}^{K} r_k(b_k).$$  (1)

Then consider a CDMA system where $K$ users transmit synchronously over an AWGN channel. The continuous time received signal after carrier demodulation and low pass filtering is modeled as follows:

$$y(t) = \sum_{k=1}^{K} \sum_{i=-\infty}^{+\infty} A_k b_k(i) s_k(t - iT) + n(t),$$  (2)

where with respect to the $k^{th}$ user, $A_k$ is the received amplitude, $b_k(i) \in \{-1, 1\}$ is the $i^{th}$ information bit, and $s_k(t)$ is the signature (spreading code) sequence. $T$ is the symbol (bit) period and $n(t)$ is the filtered channel additive white Gaussian noise (AWGN). The signature of every user is actually composed of $N$ spreading chips and it is of the form

$$s_k(t) = \sum_{n=1}^{N} c_k(n) P_{T_c} (t - (n-1)T_c),$$  (3)

where $N$ is the system processing gain, $c_k(n) \in \{-1, 1\}$, $n = 1, \ldots, N$ are the assigned signature bits for the $k^{th}$ user, and $P_{T_c}(t)$ is the spreading pulse with duration $T_c = T/N$.

Without losing the generality the signatures are assumed to be normalized such that

$$\int_0^T s_k^2(t) dt = 1, \quad \forall \quad k = 1, \ldots, K. \quad (4)$$

Since we consider synchronous transmission we can immediately drop the index $i$ from (2) and we can concentrate on a single information bit interval of $T$. After conventional matched filtering and sampling at the chip rate $1/T_c$, we organize the $N$ collected samples in the form of an $N$ dimensional vector as follows:

$$y[n] \triangleq \int_{(n-1)T_c}^{nT_c} y(t) P_{T_c} (t - (n-1)T_c) dt, \quad n = 1, \ldots, N. \quad (5)$$

The $\mathcal{R}^N$ discrete time version of (2) can now be written as

$$y = \sum_{k=1}^{K} A_k b_k s_k + n.$$  (6)

The above equation could be more compactly written as

$$y = \mathbf{S} \mathbf{A} \mathbf{b} + \mathbf{n},$$  (7)

where random vector $\mathbf{n}$ is assumed to be white Gaussian with autocorrelation matrix $E\{\mathbf{n} \mathbf{n}^H\} = \sigma^2 \mathbf{I}_{N \times N}$ and $\mathbf{S} = [s_1, s_2, \ldots, s_K]$ is $N \times K$ matrix of user signature codes. $\mathbf{A} = \text{diag}[A_1, A_2, ..., A_K]$ is a $K \times K$ diagonal matrix of user amplitudes.

III. INDEPENDENT COMPONENT ANALYSIS

ICA is a generalization of SOS based principal component analysis (PCA). It is a blind statistical techniques that tries to make the observed data independent or as independent as possible. Under an appropriate signal model ICA can be used to solve the BSS problem. ICA algorithms that exploit nongaussianity require that there exist at least one Gaussian source among the underlying sources. Most notable ICA/BSS algorithms are given in [15]–[20]. Nongaussianity based ICA algorithms find a transformation of the mixture data which restores the independence. This computation requires the use of HOS, which in general make use of nonlinear statistics that utilize nonlinearities which decay towards infinity. Although, ICA can be realized by making use of explicit HOS, i.e., fourth order cumulants or kurtosis, general nonlinearity based algorithms are computationally attractive and more robust to outliers in the data [21].

ICA algorithms yields solutions up to a scaling and permutation ambiguity [10]. These ambiguities or indeterminacies might be tolerable in the cases where most of the information is carried in the waveform of the signal (rather than its amplitude) and ordering is not important, for example, in speaker separation problem or the ‘cocktail party’ problem. However, in a communication system ordering and scaling ambiguities cannot be tolerated. Scaling ambiguity can be resolved up to a sign ambiguity by constraining the input source variance to unity. To remove the permutation ambiguity some further constraints have to be applied, which typically
in a CDMA system can be formulated in terms of a priori knowledge about the user signature codes [11], [12].

IV. HEBBIAN LEARNING BASED ICA ALGORITHM

Hebbian learning rule is one of the oldest unsupervised learning rules and is based on Hebb’s postulate of learning. The reason of this being so important in the ICA is due to the fact that these learning algorithms exhibit an interesting data dependent orientation selectivity. In these algorithms whitening of the observations before the separation is a preprocessing step for the optimum performance of the algorithm. This can easily be performed by using PCA like algorithm. The whitening transformation of the data in (7) is given as

\[ \tilde{y} = D^{-1/2}E^T y, \]

where \( D \) is a diagonal matrix of eigenvalues and \( E \) is the corresponding eigenvector matrix of the data correlation matrix \( R = E \{ yy^T \} \). Other more efficient and adaptive whitening transforms may be found in [22]. The estimated user data \( \hat{b} = W^T \tilde{y} \), and that \( W \) is expected to converge to an orthonormal matrix due to the whitening step.

For the extraction of multiple independent components we update the following \( N \times K \) separating matrix \( W \) [9]

\[ W(t + 1) = W(t) + \mu \tilde{y}(t)f(\tilde{y}(t)^T W(t))C(t) + \alpha W(t)(I - W(t)^T W(t)), \]

where \( \alpha \) is a constant, in this paper \( \alpha = 0.5 \), and the last term in (9) ensures the convergence to different independent components. \( \mu \) is the learning rate, \( f(\cdot) \) is the nonlinear function, that is applied separately on every component of the row vector \( \tilde{y}(t)^T W(t) \). A particular reason behind selecting (9) besides other ICA algorithms is its relative insensitivity to the choice of the nonlinear function \( f(\cdot) \) [21], which ideally should match the cumulative distribution function (CDF) of the underlying sources. In this paper we have selected \( f(y) = \tanh(ay) \), which is a general purpose nonlinearity and has been shown to work well with almost all types of sources [9], \( a \) is a constant and refer to slope of the nonlinearity, which is taken as \( a = 2 \). \( C(t) \) is a diagonal matrix of elements \( \pm 1 \) which decides if the learning is Hebbian or anti-Hebbian and ensures the convergence of the algorithm to the independent components. \( \tilde{y}(t)f(\tilde{y}(t)^T W(t)) \) is called the Hebbian term in (9) and is responsible for the learning.

V. REMOVAL OF PERMUTATION AND SIGN AMBIGUITIES

In this section, we provide a method of removing the permutation and sign ambiguities from the ICA solution based on the knowledge of user signature codes. As we noted in the previous section that to obtain optimum performance the data used in (9) should be prewhitened, hence in order to obtain an estimate of the mixing matrix or the user amplitude scaled signature code matrix we need to multiply the steady state matrix \( W \) obtained from (9) by the dewhiteing matrix, hence

\[ \hat{S} = ED^{1/2}W, \]

where \( \hat{S} \) is an estimate of signature code matrix which bears the effect of near-far situation. The matrix \( \hat{S} \), however, has its columns permuted and scaled with respect to the true signature code matrix \( S \). Let us focus on the detection of the kth user, given the information about its signature code and timing. The algorithm can be described in following steps

1) Compute the correlation vector of the desired user’s signature sequence with the estimated signature code matrix ; \( r_k = s_k^T \hat{S} = [\rho_{k1}, \rho_{k2}, \cdots, \rho_{kK}] \)
2) Compute the component wise absolute value of the vector \( \tilde{r}_k = [|\rho_{k1}|, |\rho_{k2}|, \cdots, |\rho_{kK}|] \)
3) Let \( r_{kj} \) denote the maximum of \( r_k \), where \( 1 \leq j \leq K \), is the index of the element where the maximum occurs.
4) Similarly let \( \tilde{r}_{kj} \) denote the maximum of \( \tilde{r}_k \)
5) If \( r_{kj} \neq \tilde{r}_{kj} \) a sign ambiguity is detected. \( j \) gives the index of the desired source among the separated sources.

For the detection of multiple users similar procedure can be extended to form a \( K \times K \) permutation matrix \( P \) with elements \( \pm 1 \), whose \( (kj)^{th} \) element is modified with \( +1 \) if no ambiguity is detected and \( -1 \) if the ambiguity is detected. The matrix \( S \) will then have its columns arranged with permutation and sign ambiguity resolved. In the reverse link where the processing is done at the base station, we are allowed to make the assumption of a priori information about all the active users in the channel. The proposed method then performs joint mutliuser detection in a parallel fashion, with no decoding delays.

VI. ESTIMATION OF USER AMPLITUDES AND CROSS-CORRELATION COEFFICIENTS

After the permutation and sign ambiguities are resolved, the matrix \( SP \) bears the effects of power imbalance between the users. If we denote \( SP \) as \( S_{per} \), then user amplitudes can be estimated as diagonal elements of the matrix \( \hat{A} \), where

\[ \hat{A} = S^T S_{per} \]

It is easy to verify that in this case, the first column of matrix \( \hat{A} \) will give the user cross-correlation coefficients. In the next section, we provide some numerical results and simulation examples to demonstrate the validity of the algorithm presented in this paper.

VII. SIMULATION RESULTS

In this section, we simulate the performance of the proposed algorithm in the presence independent users with correlated signature codes. Correlation between the user signature codes can arise due to the multipath propagation or may be induced intentionally in a bid to increase the capacity of the system. For a given length of spreading code, by allowing correlation between the users, more users can be accommodated in the system as compared to a system employing orthogonal codes.

While simulating the system, we assume that the system is operating at its full capacity, i.e., the number of the users in the channel is equal to the length of the spreading code, i.e., \( N = K \). The system is simulated in a simple case of four users and the length of the spreading code is also taken as

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.
4. This length of code has partly been selected in an effort to illustrate the numerical examples without losing too much space in writing the matrices involved. The data modulation taken in the simulations is BPSK, 30,000 data points are taken while simulating the system. Following is the cross-correlation matrix between the signature codes of the users

\[
R = \begin{bmatrix}
1.000 & 0.200 & -0.120 & 0.107 \\
0.200 & 1.000 & 0.200 & -0.120 \\
-0.120 & 0.200 & 1.000 & 0.200 \\
0.107 & -0.120 & 0.200 & 1.000
\end{bmatrix}.
\]

(12)

Corresponding signature code matrix is given by

\[
S = \begin{bmatrix}
0.5000 & 0.5899 & 0.5374 & 0.5837 \\
0.5000 & -0.3899 & 0.3087 & -0.2222 \\
0.5000 & -0.3899 & -0.4287 & -0.6210 \\
0.5000 & -0.3899 & -0.6574 & 0.4736
\end{bmatrix}.
\]

(13)

Amplitudes of users in the case of no power control are selected such that interfering users are 10dB above the amplitude of the desired user. The amplitude scaling matrix was chosen to be

\[
A = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 3.1623 & 0 & 0 \\
0 & 0 & 3.1623 & 0 \\
0 & 0 & 0 & 3.1623
\end{bmatrix}.
\]

(14)

An estimate of the signature code matrix in perfect power control by the proposed algorithm is obtained as

\[
\hat{S}_1 = \begin{bmatrix}
-0.5170 & -0.6029 & 0.5993 & 0.5525 \\
-0.5272 & 0.4095 & -0.2392 & 0.3290 \\
-0.5150 & -0.6122 & -0.6422 & -0.4438 \\
-0.5106 & 0.3985 & 0.4865 & -0.6790
\end{bmatrix}.
\]

(15)
whereas in the absence of power control, when the desired user is 10 dB below the interfering users, the estimated signature code matrix is obtained as

$$\hat{S}_2 = \begin{bmatrix} 1.8796 & 1.8648 & -0.4979 & 1.7094 \\ -1.2573 & -0.7169 & -0.5382 & 1.0010 \\ 1.8905 & -1.9899 & -0.5196 & -1.3766 \\ -1.2438 & 1.5110 & -0.5020 & -2.0972 \end{bmatrix}. \quad (16)$$

It can be seen from (15) that the desired user’s signature code appears in the first column of the estimated signature matrix within a sign ambiguity. From (16) we see that in this case the desired users’ code appears in the third column again with the sign ambiguity, whereas the signature codes of other users are scaled by their respective amplitudes and signs as applicable. Now, by using the algorithm presented in section-V, we remove the permutation and sign ambiguity to obtain a permuted form of (16) as

$$\hat{S}_{\text{per}} = \begin{bmatrix} 0.4979 & 1.8796 & 1.7094 & 1.8648 \\ 0.5382 & -1.2573 & 1.0010 & -0.7169 \\ 0.5196 & 1.8905 & -1.3766 & -1.9899 \\ 0.5020 & -1.2438 & -2.0972 & 1.5110 \end{bmatrix}. \quad (17)$$

From this ambiguity resolved matrix, an estimate about the amplitudes or powers of different users can be obtained as the diagonal elements of $\hat{A} = S^T \hat{S}_{\text{per}}$. We obtained the matrix $\hat{A}$ as

$$\hat{A} = \begin{bmatrix} 1.0288 & 0.6345 & -0.3817 & 0.3345 \\ 0.1947 & 3.1991 & 0.6237 & -0.3834 \\ -0.1191 & 0.6289 & 3.1964 & 0.6406 \\ 0.0860 & -0.3865 & 0.6371 & 3.1991 \end{bmatrix}. \quad (18)$$

It can be seen from the above matrix that in fact estimates of user amplitudes (14) appear as the diagonal elements and cross-correlations (12) appear as the first column of this matrix. These estimates require a priori information about all the active users in the channel and are particularly useful at the base station where such a priori information can be assumed. The results were obtained from a single run of proposed algorithm (9), (10), (11) in 10 dB SNR, and the accuracy of these results should improve by averaging.

In Fig. 1(a) we compute the probability of error in perfect power control situation, i.e., the user power are all equal. The user correlation values were taken as in (12). The results are averaged over 50 Monte-Carlo experiments. System is simulated for the conventional detector, the MOE detector of [7] and ICA based method proposed in this paper. For comparison purposes theoretical performance of single user system with BPSK modulation and ideal decorrelator are also plotted and have been labeled as “Theory” and “Ideal” respectively. Decorrelator [1] is a non-adaptive approach which gives performance very near to the optimum detector. The MOE detector [7] is a projection based algorithm, which for the detection of $k^{th}$ user is given as $c_k(t+1) = s_k + \tilde{w}_k(t)$, where

$$\tilde{w}_k(t) = \tilde{w}_k(t-1) - \mu Z(t) y(t) - Z_{mf} s_k,$$ \quad (19)

where $Z_{mf} = s_k^T y(t)$ is the output of the conventional detector and $Z(t) = c_k^T y(t)$ is the output of the MOE detector [7]. The step size parameter for the MOE detector was taken as 0.001 while for ICA approach it is 0.01, these parameters were based on the best simulation results obtained. It can be seen from Fig. 1(a) that in perfect power control the performance of ICA based detector is comparable to that of the MOE based detector, however, as revealed from Fig 1(b), the performance of ICA based detector is significantly better than that of the MOE detector. This observation is verified by another set of simulations conducted with the following...
correlation values

\[
R = \begin{bmatrix}
1.0000 & 0.2500 & -0.2200 & 0.2170 \\
0.2500 & 1.0000 & 0.2500 & -0.2200 \\
-0.2200 & 0.2500 & 1.0000 & 0.2500 \\
0.2170 & -0.2200 & 0.2500 & 1.0000
\end{bmatrix}.
\]  

(20)

The correlation values are higher in (20) than in (12). It can be seen from Fig. 2(a),(b) that the performance of ICA and MOE detector is quite similar in the case of perfect power control, whereas ICA detector outperforms the MOE detector in the absence of power control. The above simulations indicate the superiority of the ICA based method over the MOE based method in near-far situation. In other words ICA based detector is more near-far resistant than the MOE based detector. In Fig. 3, we simulate the achievable steady state signal to interference plus noise (SINR) for user 1 in 10dB signal to noise (SNR) ratio. SINR is defined in a CDMA system as

\[
\text{SINR} = \frac{\sum_{k \neq 1} \text{A}^2_k (w^T_k s_k)^2 + \sigma^2 w^T_1 w_1}{\text{A}^2_1 (w^T_1 s_1)^2},
\]

(21)

where \(\sigma^2\) the noise variance and \(w_1\) is the detector for the desired user.

In Fig. 3, the ideal SINR corresponds to the case when no interfering users are present, which is the SNR of the desired user, i.e., 10 dB. The simulations were carried with the correlation values in (12). The SINR of MOE detector, in perfect power control, Fig. 3(a), approaches that of ICA based detector explaining the similar performances of two approaches in terms of bit error rate, see Fig.1 and 2. From Fig. 3(b) we see the effectiveness of ICA based detector in near-far situation, where its performance is rather unaffected, while MOE detectors depicts a degradation in the achievable SINR, these results can also be verified from Fig. 1(a),(b) and Fig. 2(a),(b), where ICA detector exhibits nearly the same performance irrespective of user power imbalance. From looking at Fig. 3 we also notice that the convergence of the proposed algorithm is also very good when compared to the blind MOE detector.

**VIII. CONCLUSIONS**

In this paper, we have presented a Hebbian learning based blind adaptive multiuser detector for DS-CDMA system. We have formulated the problem of blind multiuser detection as the problem of finding independent bases in a CDMA system consisting of independent sources. We provided an algorithm for resolving permutation and scaling ambiguities from the solution of Hebbian algorithm. We demonstrated the efficacy of our scheme in presence and absence of power control. It has also been demonstrated in the paper that user amplitudes and cross-correlation coefficients can also be estimated by the proposed algorithm. Simulation examples with BPSK data illustrate the effectiveness of the proposed algorithm. Probability of error and SINR curves suggest the superiority of the proposed method over MOE detector in similar near-far conditions. The proposed detector was shown to be more near-far resistant than the MOE detector. While single user can be detected with the same information as required by the conventional detector, the structure of the proposed detector is more suitable for joint multiuser detection at the base station, where information about all the active users can be assumed.

**References**