Group-Blind Multiuser Detectors for CDMA: Synchronous Systems

Anders Høst-Madsen

TRLabs and the University of Calgary
280 Discovery Place One, 3553 – 31 Street NW
Calgary, Alberta, Canada, T2L 2K7

ABSTRACT
We consider the problem of multiuser detection for CDMA (code division multiple access) systems where the codes of some users are known while others are unknown, group-blind multiuser detection. An example is at the base station of a cellular communication system with interference from both in-cell users, with known codes, and out-of-cell users, with unknown codes. In this paper we develop a number of group-blind multiuser detectors using a subspace approach. One class of detectors developed has a computationally efficient structure of a subspace estimation common to all known users, followed by a simple processing specific to each user. The performance of the detectors is compared to that of the purely blind detector and the non-blind detector, and the group-blind detectors are seen to have the best performance.

KEYWORDS
Multiuser detection, CDMA, spread spectrum, multiuser communications, subspace methods, intercell interference.

I. INTRODUCTION
Multiuser detection is a method to improve detection performance and capacity of multiple access spread spectrum (or CDMA) systems. Multiuser detection was introduced by Verdú in [1], where the optimal multiuser detector was derived. The optimal detector has an exponential complexity in the number of users, and less complex (linear) multiuser detectors were therefore derived in an number of papers, in particular the decorrelating detector [2] and the minimum mean square error (MMSE) detector [3]. For a review of multiuser detection and further references, see, e.g., [4].

The early works on multiuser detection assumed that the codes of all users were known at the receiver, and made a simultaneous detection of all users (therefrom the name multiuser detection or joint detection). If the detection for example is at the base station of a mobile communication system, this seems realistic, as the base station needs to perform detection for all users. On the other hand, it is unrealistic that a mobile station should know the codes of all other users in a cell, and therefore it is desirable to consider multiuser detectors that need
to know only the code of the desired user, blind multiuser detection. The blind MMSE detector was introduced in a number of papers [6,7], and recently it was also shown that the decorrelating detector can be implemented blindly [13,14].

Although a base station knows all codes of the users within a cell, it typically will not know the codes of interfering users from surrounding cells. Even if a base stations could obtain this knowledge from surrounding base stations, it would be a waste of resources if it were to also perform detection for these users just to cancel interference (including synchronization etc.). This is a serious problem to multiuser detection, since typically 1/3 of the interference could be from other cells, intercell interference [8,9]. Thus, also at the base station blind detectors could be relevant. On the other hand, blind detectors do not use neither the fact that the codes of in-cell users are known at the base station, nor that these other users also have to be detected.

This has led us to consider multiuser detectors that can cancel interference from both known and unknown users, while utilizing the information about known users and the fact that detection has to be done for all known users. A blind multiuser detector basically in some way has to estimate the codes of interfering users, and by using the known codes the estimation accuracy can be improved. On the other hand, since several users have to be detected jointly, it is also advantageous, considering computational complexity, if some of the processing can be common to all users. In this paper we will develop detectors that satisfy these two criteria. We call this class of detectors \textit{group-blind multiuser detectors}.

We will only consider linear detectors, i.e., the decorrelating and the MMSE detector. For non-linear approaches see [11] and [12]. We will use the subspace approach of Wang and Poor [14]. We will only consider synchronous systems because of their conceptual simplicity. However, although the theoretical development and the simulations assume that all users are synchronous, the formulas developed can be applied without change to systems where the \textit{unknown} users are not synchronous. The results can therefore be directly applied to systems where the known users are synchronous, e.g., the downlink in a cellular system, and the up-link in some systems. With some additional development, the results can also be applied to a quasi-synchronous CDMA system [16], where the users within a cell are (almost) synchronous, but the interference from other cells is asynchronous.

The truly asynchronous case will be developed in another paper.

\section{SYSTEM MODEL}

Consider a synchronous spread spectrum communications system with the users transmitting through an additive white Gaussian noise channel. The user population consists of $K$ users with known codes, and $\tilde{K}$ users
with unknown codes. As we consider synchronous systems, it is sufficient to consider a single symbol interval
\[ [0, T], \]
where the received signal can be written as
\[
 r(t) = \sum_{k=1}^{K} b_k A_k s_k(t) + \sum_{k=1}^{\tilde{K}} \tilde{b}_k \tilde{A}_k \tilde{s}_k(t) + n(t), t \in [0, T]
\]
where \( s_k \) is the normalized code waveform of the \( k' \)th known user with support in \([0, T]\), \( \tilde{s}_k \) the waveform of the
\( k' \)th unknown user, \( b_k, \tilde{b}_k \) are the transmitted bits (±1), \( A_k, \tilde{A}_k \) the amplitudes and \( n \) white Gaussian noise with
power \( \sigma^2 \). We assume that the codes are given by chip codes, \( s_k(t) = \sum_{i=0}^{M-1} c_k^i \psi(t - iT / M) \), where
\( c_k^i \in \{-1, +1\} \) and \( \psi \) is the chip waveform, and similarly for the unknown users. A sufficient statistic for the
received signal is therefore the output of a chip rate sampled chip matched filter, and we can write this statistic
on vector form as
\[
 r = \sum_{k=1}^{K} b_k A_k s_k + \sum_{k=1}^{\tilde{K}} \tilde{b}_k \tilde{A}_k \tilde{s}_k + n = Sb + \tilde{S}\tilde{b} + n
\]
where \( S = [s_1, s_2, \ldots, s_K] \), \( A = \text{diag}(A_1, \ldots, A_K) \), and \( b = [b_1, b_2, \ldots, b_K]^T \). The \( \tilde{\cdot} \) marked symbols are defined
similarly. The correlation matrix of \( r \) is given by
\[
 R = E[rr^T] = SA^2S^T + \tilde{S}\tilde{A}^2\tilde{S}^T + \sigma^2 I
\]
We will assume that all codes, of both known and unknown users, are linearly independent, so that \([S \quad \tilde{S}]\) has
full rank.

As mentioned in the introduction, the system model can still be used if the unknown users are
asynchronous. In that case, different bits within the interval \([0, T]\) from the same user are considered as different
‘virtual’ users, and \( \tilde{K} \) is now the number of ‘virtual’ unknown users, which can be between one and two times
the actual number of unknown users.

### III. DETECTOR STRUCTURES

A general linear detector for user \( i \) is given by
\[
 \hat{b}_i = \text{sgn}(w_i^T r)
\]
For jointly detection of all users, the detector can be written
\[
 \hat{b} = \text{sgn}(W^T r)
\]
where \( W = [w_1 \quad w_2 \quad \cdots \quad w_K] \). As argued in \([10]\), any interesting detector must have \( w_i \in \text{range}[S \quad \tilde{S}] \),
since any component of \( w_i \) outside this subspace will only increase noise without reducing interference.
A. The decorrelating detector

Prior to developing the group-blind detector we will discuss general decorrelating detectors in a geometrical framework useful to understanding the decorrelating detector. A decorrelating detector for user \(i\) satisfies

\[
\begin{align*}
    w_i^T s_i &= 1 \\
    w_i^T s_k &= 0, k \neq i \\
    w_i^T s_k &= 0, k = 1, \ldots, K
\end{align*}
\]

(1)

These conditions imply that the decorrelating detector cancels the interference from other users completely,

\[
w_i^T r = A_i b_i + w_i^T n
\]

The conditions (1) conditions together with \(w_i \in \text{range} [S \, \bar{S}]\) determine \(w_i\) uniquely, since \([S \, \bar{S}]\) has full rank. We can also write \(w_i\) as \(w_i = M_i s_i\) form some \(M_i\). Although \(w_i\) is unique, clearly \(M_i\) need not be unique, and in the following we will discuss some different natural choices of \(M_i\). A practical solution for \(M_i\) is a solution \(M\) that does not depend on \(i\), since then \(w_i^T r = s_i^T M^T r\) and \(M^T r\) can be calculated common to all users. Especially if \(M\) is low-rank, or if it has to be estimated, as is the case when some user codes are unknown, this can yield an efficient solution.

To understand how the decorrelating estimator can be estimated, we will at first analyze the case when all codes are known, i.e., \(\tilde{K} = 0\). Without loss of generality we can consider detection for user 1. There are two natural approaches to the decorrelating detector: the projection approach and the orthogonalization approach. In the projection approach the received vector \(r\) is projected on the subspace orthogonal to the codes of the other users, whereby their interference has been removed, followed by a projection on the subspace spanned by user 1. An explicit solution can be derived as follows. Let

\[
S_1 = [s_2, s_3, \ldots, s_K]
\]

(2)

and let

\[
P_1 = S_1 \left( S_1^T S_1 \right)^{-1} S_1^T
\]

(3)

represent the projection unto the subspace spanned by \(S_1\), with \(P_1^\perp = I - P_1\) the projection unto the orthogonal subspace. Then we can write the decorrelating detector as

\[
w_1 = \frac{1}{s_1^T P_1^\perp s_1} P_1^\perp s_1
\]
The factor $s_1^T P_1^\perp s_1$ is simply a normalizing factor, and can be omitted if we consider BPSK, where we are only interested in the sign of the output. It is easy to see that the conditions [1] are satisfied. The disadvantage of this solution is that $P_1^\perp$ is specific to user 1, and each user will therefore use a different projection.

We will therefore consider the orthogonalization approach. The problem with the traditional matched filter solution is that the vectors $s_1, s_2, ..., s_K$ are not orthogonal. However, since $s_1, s_2, ..., s_K$ are linearly independent, there exists some inner product in $\mathbb{R}^M$, with respect to which $s_1, s_2, ..., s_K$ are orthonormal. Specifically, there exists some (positive semi definite) matrix $W \in \mathbb{R}^{M \times M}$ so that $s_i^T W s_j = \delta_{i,j}$. Since $s_1, s_2, ..., s_K$ are not a basis of $\mathbb{R}^M$, the matrix $W$ is not unique, but two solutions are

$$W = S (S^T S)^{-2} S^T$$

and

$$W_A = S A (A^T S A)^{-2} A^T,$$

which gives

$$w_1 = W s_1$$

or

$$w_1 = \frac{1}{s_1^T W A s_1} W_A s_1.$$

The conditions [1] are satisfied (seen most easily by noticing that $S^T W S = I$, and that scaling of vectors do not influence orthogonality). This solution is in fact equivalent to the “traditional” decorrelating detector [2], since $S^T W r = (S^T S)^{-1} S^T r$. The solution further has the advantage that $W$ is independent of the user considered and that it is a low rank matrix, which makes the transformation efficient to calculate.

To estimate $W$ (or rather $W_A$) we can make use of the following. The eigenvalue decomposition of $R$ can be written as [14]

$$R = \begin{bmatrix} U_s & U_n \end{bmatrix} \begin{bmatrix} \Lambda_s & 0 \\ 0 & \sigma^2 I \end{bmatrix} \begin{bmatrix} U_s^T \\ U_n^T \end{bmatrix}$$

where $U_s$ is an orthonormal basis of the signal subspace spanned by $S$ and $U_n$ is a basis for the noise subspace orthogonal to $U_s$. The diagonal matrix $\Lambda_s$ contains the $K$ largest eigenvalues of $R$, while $\sigma^2$ is the minimum eigenvalue. We then have, by examining the proofs in [14]
Proposition 1

\[ W_A = U_s \left( \Lambda_s - \sigma^2 I \right)^{-1} U_s^T \]

Thus, as \( R \) can be estimated from the received data only, \( W_A \) can be estimated by EVD, and the orthogonalizing decorrelating detector can be implemented blindly \([14]\).

B. The MMSE detector

The minimum mean square error (MMSE) detector is defined as the detector that minimizes the mean square error (MSE):

\[ w_i = \arg \min E \left[ \left( b_i - w^T r \right)^2 \right] \quad (7) \]

Equivalently, the solution is given as the vector that satisfies

\[ E \left[ \mathbf{r} (b_i - w^T r) \right] = 0 \]

which results in the equation

\[ \mathbf{R} w_i = \mathbf{A}_i s_i \]

Thus, we can write

\[ \hat{b}_i = \text{sgn} \left( s_i^T \mathbf{R}^{-1} r \right) \]
\[ \hat{b} = \text{sgn} \left( S^T \mathbf{R}^{-1} r \right) \quad (8) \]

Notice that this detector can be implemented blindly by replacing \( \mathbf{R} \) by a time-averaged estimate \([17]\). Using the eigenvalue decomposition \([6]\), this can also be written as

\[ \hat{b}_i = \text{sgn} \left( s_i^T U_s \Lambda_s^{-1} U_s^T r \right) \]
\[ \hat{b} = \text{sgn} \left( S^T U_s \Lambda_s^{-1} U_s^T r \right) \quad (9) \]

which is the idea of the subspace method in \([14]\).

Now suppose that \( w_i \) is constrained to lie in some subspace \( U \) of \( \mathbb{R}^M \), let \( \mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m \) be an orthonormal basis of \( U \), and let \( U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m] \). We can then write \( w_i = U \mathbf{c}_i \) for some vector \( \mathbf{c}_i \). Inserting this into \([7]\) gives

\[ \mathbf{c}_i = A_i \left( U^T RU \right)^{-1} U^T s_i \quad (10) \]

if \( P \) is the projection onto \( U \) we can also find \( w_i \) as

\[ \mathbf{c}_i = A_i \left( PRP \right)^\dagger P s_i \]

where \( ^\dagger \) denotes pseudo-inverse (which is needed here, since\( PRP \) does not have full rank).
IV. GROUP-BLIND DETECTORS

A. Group-blind decorrelating detectors

In our first approach to group-blind decorrelating detectors, we use a mixed projection/orthogonalization approach, as described in section III.A. The idea is to first project the received data \( r \) on the subspace orthogonal to \( \text{span}(S_1) \), with \( S_1 \) given by (2). Thereby all interference from the known users has been removed. The orthogonalization approach is then used in \( \text{span}(S_1)^\perp \). The calculations can be done as follows. Let \( P_1 \) be defined by (3), and define \( \tilde{S}_1 = P_1^\dagger \begin{bmatrix} \tilde{S} & s_1 \end{bmatrix} \), \( \tilde{A}_1 = \begin{bmatrix} \tilde{A} & A_1 \end{bmatrix} \) and

\[
\tilde{W}_1 = \tilde{S}_1 \tilde{A}_1 \left( \tilde{A}_1 \tilde{S}_1^T \tilde{S}_1 \tilde{A}_1 \right)^{-2} \tilde{A}_1 \tilde{S}_1^T
\]

We can write the eigenvalue decomposition of \( \tilde{R}_1 = P_1^\dagger R P_1^\perp \) as

\[
\tilde{R}_1 = \begin{bmatrix} U_s & U_n & U_o \end{bmatrix} \begin{bmatrix} \Lambda_s & 0 & 0 \\ 0 & \sigma^2 I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_s^T \\ U_n^T \\ U_o^T \end{bmatrix}
\]

Here \( U_s \) is an orthonormal basis of the signal subspace, which in this case is spanned by \( \begin{bmatrix} \tilde{S} & s_1 \end{bmatrix} \), while \( U_n \) is the noise subspace. The matrix \( U_o \) is an orthonormal basis of the space spanned by \( S_1 \). Furthermore, \( \Lambda_s = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_{K+1}) \) with \( \lambda_i > \sigma^2 \). Similarly to the blind decorrelating detector we now have that

\[
\tilde{W}_1 = U_s \left( \Lambda_s - \sigma^2 I_{K+1} \right)^{-1} U_s^T
\]  

(11)

We can then state the group-blind decorrelating detector as

**Theorem 1:** The decorrelating detector is given by

\[
w_1 = \frac{1}{s_1^T \tilde{W}_1 s_1} \tilde{W}_1 s_1
\]  

(12)

In our setting, \( \tilde{S}_1 \) is unknown. Equation (11), however, outlines a method for estimating \( \tilde{W}_1 \), and the method can therefore be implemented without knowledge of the unknown users’ codes.

The problem of the above method is that \( \tilde{W}_1 \) is specific to user 1. It can be used with advantage in the case where some users’ codes are known, but still only one user’s data is of interest, as for example in a mobile station. However, if several users are to be detected simultaneously, an SVD/subspace tracking has to be done for each user, and the method is inefficient.
We will therefore develop a method that requires only one SVD common to all users. This method will be based on the orthogonalization approach. First, define the projection on \( \text{span}(S) \),

\[
P = S(S^T S)^{-1} S^T
\]

with \( P^\perp = I - P \). The eigenvalue decomposition of \( P^\perp R P^\perp \), is then given by

\[
P^\perp R P^\perp = \begin{bmatrix}
\tilde{U}_s & 
\tilde{U}_n & 
\tilde{U}_o
\end{bmatrix}
\begin{bmatrix}
\tilde{\Lambda}_s & 0 & 0
0 & \sigma^2 I & 0
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{U}^T_s
\tilde{U}^T_n
\tilde{U}^T_o
\end{bmatrix}
\]

where \( \tilde{\Lambda}_s = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, ..., \tilde{\lambda}_K) \) with \( \tilde{\lambda}_i > \sigma^2 \), and \( \tilde{U}_o \) has \( K \) rows.

The idea in the group-blind detector is first to apply the projection \( P^\perp \) to the received data. Thereby the signal from known all users have been eliminated. The subspace of the unknown users, projected to \( \text{span}(S)^\perp \), is then found. This is \( \text{span}(\tilde{U}_s) \). The last step is to combine this subspace with the code of each user through some algebraic manipulations.

We can now state the group-blind decorrelating detector as follows

**Theorem 2**: The decorrelating detector is given by

\[
\hat{b} = \text{sgn}\left((S^T S)^{-1} S^T \left(I - R \tilde{U}_s (\tilde{\Lambda}_s - \sigma^2 I)^{-1} \tilde{U}^T_s\right) r\right)
\]  

(13)

*Proof*: see Appendix A.

Notice that all quantities in the theorem can be estimated with knowledge of only the known users’ codes. Notice also that although \( R \) appears explicitly, \( R \) itself does not have to be estimated. Define

\[
T(n) = S^T \hat{R}(n) \tilde{U}_s(n)
\]

where \( \hat{R}(n) \) etc. denote the estimated quantities over \( n \) samples. Then we can rewrite (13) as

\[
\hat{b}(n) = \text{sgn}\left((S^T S)^{-1} \left(S^T - T(n) \left(\tilde{\Lambda}_s(n) - \sigma^2 I\right) \tilde{U}^T_s(n)\right) r_n\right)
\]  

(14)

The quantity \( T(n) \) can be calculated as follows,
The implementation using the formula on the second line only requires \( K K \) multiplications per bit (since \( S^T r \) has to be calculated in any case), instead of the \( M^2 \) multiplications to calculate \( \mathbf{R} \). On the other hand this formula is not directly suitable for recursive implementation, since \( \hat{\mathbf{U}}_s^T(n) \) has to be applied to all previous samples. The formula on the third line requires \( K M \) multiplications, and can be implemented recursively as long as \( \mathbf{S} \) is time-invariant.

**B. Hybrid Group-blind Detectors**

We will in this section consider detectors that are combinations of the decorrelating detector and the MMSE detector. Specifically, we will study detectors that are decorrelating among the known users and MMSE with respect to the unknown users. We will call this class of detectors *hybrid group-blind detectors*. The idea is similar to ideas used in array processing: to direct nulls in the direction of known interferers, and find the MMSE solution for the remaining interference, and has in this context also been studied for multiuser detection [18,19].

As for the decorrelating detector, we will give two solutions for the hybrid detector: one that makes a subspace calculation for each user, and one that makes a common subspace calculation. For the former case we get the following Theorem 3. The notation is the same as for Theorem 1.

**Theorem 3**: The hybrid group-blind detector is given by

\[
\mathbf{w}_1 = \frac{1}{\mathbf{s}_1^T \mathbf{U}_s \Lambda_s^{-1} \mathbf{U}_s^T \mathbf{s}_1} \mathbf{U}_s \Lambda_s^{-1} \mathbf{U}_s^T \mathbf{s}_1
\]

where \( \mathbf{U}_s \) and \( \Lambda_s \) is given from the eigenvalue decomposition of \( \mathbf{R}_1 = \mathbf{P}_1^\perp \mathbf{R} \mathbf{P}_1^\perp \):
where $\Lambda_s = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{K+1})$ with $\lambda_i > \sigma^2$.

This detector is similar to the detector of [18], but is here stated in a subspace context.

For the case of common subspace tracking we get, with the same notation as for Theorem 2

**Theorem 4:** The hybrid group-blind detector is given by

$$\hat{b} = \text{sgn}\left( (s^T s)^{-1} s^T \left( I - R \tilde{U}_s \tilde{A}_s^{-1} \tilde{U}_s^T \right) r \right)$$

(17)

**Proof:** see Appendix B.

We notice that the hybrid group-blind detector (17) is very similar to the group-blind decorrelating detector (13). The only difference is that $\tilde{A}_s - \sigma^2 I$ is replaced by $\tilde{A}_s$. As for (17), $R$ itself does not have to be estimated, as the quantities needed can be calculated using (15).

In general, the MMSE detector has superior performance to the decorrelating detector [10], and it would therefore be natural to study pure MMSE group-blind detectors. However, the hybrid approach has a number of advantages. For the known users, the decorrelating detector does not need any estimated information, whereas the MMSE solution requires estimates and tracking of user powers. On the other hand, the decorrelating detector for the unknown users requires a noise estimate and is therefore rather dependent on an accurate estimate of the number of users, while the MMSE detector does not have these shortcomings. Finally, the group-blind hybrid detector has a compact, simple expression (17), while it seems impossible to derive the true MMSE detector by the methods used here.

One way to make (17) closer to a true MMSE solution, is to replace it with the following expression

$$\hat{b} = \text{sgn}\left( (s^T s + \sigma^2 \Lambda_s^{-2})^{-1} s^T \left( I - R \tilde{U}_s \tilde{A}_s^{-1} \tilde{U}_s^T \right) r \right)$$

(18)

The argument for (18) is heuristic. We first find the MMSE solution in the subspace spanned by the known users, ignoring all interference from unknown users. For each user, we then find the MMSE solution in the subspace spanned by this solution and $\tilde{U}_s$. This solution can avoid the amplification of noise that could happen in (17), but it is not an MMSE solution and does not satisfy any orthogonality or MMSE conditions.

**V. SUBSPACE TRACKING**

All of the group-blind multiuser detectors derived in section [IV] require some form of EVD or SVD. Direct calculation of EVD or SVD is complex, and we therefore consider less complex subspace tracking methods for implementation.
A large number of different subspace tracking methods exist, with varying convergence speed and complexity, and most of them can be applied to the current problem, but with some modifications. To illustrate the principles we have chosen to use the F2 algorithm [20], summarized in Table 1. We use a standard exponential window with

\[
\hat{R}(n) = \lambda \hat{R}(n-1) + r_n r_n^T
\]

Table 1: The F2 algorithm [20] for tracking a \(d\)-dimensional subspace.

<table>
<thead>
<tr>
<th>Let the data vectors be (v_1, v_2,\ldots).</th>
</tr>
</thead>
<tbody>
<tr>
<td>(U(0)) arbitrary (m \times d) matrix, (U(0)^T U(0) = I)</td>
</tr>
<tr>
<td>(\Sigma(0) = I).</td>
</tr>
<tr>
<td>For (n = 1, 2, 3, \ldots)</td>
</tr>
<tr>
<td>Define (W(n) = \sqrt{\lambda(n-1)} \Sigma(n-1) v_n)</td>
</tr>
<tr>
<td>Compute the (m \times d) and (d \times d) matrices, (U(n), \Sigma(n)) from the SVD:</td>
</tr>
</tbody>
</table>
| \[
\begin{bmatrix} U(n) & \Sigma(n) & 0 \\ \\
0 & \sigma & \end{bmatrix} Y(n)^T = W(n)
\] |

For the group-blind algorithms (16) and (12) the input to the F2 algorithm for user 1 are the vectors

\[
v_n = P_{1}^\perp r_n
\]

The subspace dimension is \(d = \tilde{K} + 1\). The matrix \(U_s(n)\) is directly the output \(U(n)\) of the F2 algorithm, while \(\Lambda_s(n) = \Sigma(n)^2\).

For the group-blind detectors (17) and (13), the input to the F2 algorithm is

\[
v_n = P^\perp r_n
\]

and the subspace dimension is \(d = \tilde{K}\). Again, the subspace information can be obtained directly from the output of the F2 algorithm with the matrix \(\bar{U}_s(n)\) equal to \(U(n)\) and \(\bar{\Lambda}_s(n) = \Sigma(n)^2\). However, also \(T(n)\) (equation (T5)) has to be calculated to implement the algorithms. As stated below equation (T5) this formula on the second line is not suitable for recursive implementation. We will here give some approximate formula for a recursive implementation, that can also be used when \(S\) is time-varying.

With the weighting factor included, \(T(n)\) can be rewritten as
The level of approximation used in evaluating (19) should correspond to the level of approximation used in the specific subspace tracking algorithm used, so that there is an agreement in convergence speed and complexity between the subspace tracking algorithm and the update formula for $T(n)$. Thus, a general update formula cannot be given, and we will look at two examples. The first possibility is to replace $\hat{U}_s(n)$ with $\hat{U}_s(i)$ inside the summation in (19), to get

$$T(n) = \sum_{i=1}^{n} \lambda^{n-i} \left( S^T r_i \left( \hat{U}_s^T(i) r_i \right)^T \right)$$

This approximation is essentially the same as the key approximation used in deriving PASTd [21,22], and gives good performance with this algorithm. However, for a fast converging algorithm such as F2, we have found this does not result in good convergence. Another approximation is

$$T(n) = \lambda T(n-1) \hat{U}_s^T(n-1) U_s(n) + S^T r_n \left( \hat{U}_s^T(n) r_n \right)^T$$

The justification of this approximation is geometrical. The quantity $\hat{U}_s(n-1) T^T(n-1)$ represents the orthogonal projection of $\hat{R}(n-1) S^T$ onto the subspace spanned by $\hat{U}_s(n-1)$. To find the projection of $\hat{R}(n-1) S^T$ onto the subspace spanned by $\hat{U}_s(n)$ we should reproject $\hat{R}(n-1) S^T$. Instead we project the previous projection. If the subspace does not change too much from $n$ to $n+1$, this is only a slight approximation. Notice also that this approximation is similar to the key approximation in deriving F2 (equation (27) in [20]).

The calculation of $\hat{U}_s^T(n-1) \hat{U}_s(n)$ is complex in itself. Fortunately, this can be obtained as an intermediate result in the F2 algorithm. Notice that since $[U(n) \quad u]^T [U(n) \quad u] = I \quad \text{[Table 1]}$, we have

$$\begin{bmatrix} \Sigma(n) & 0 \\ 0 & \sigma \end{bmatrix} Y^T \begin{bmatrix} \Sigma(n-1) & 0 \\ 0 & \bullet \end{bmatrix}^{-1}$$

$$= [U(n) \quad u]^T W(n) \begin{bmatrix} \Sigma(n-1) & 0 \\ 0 & \bullet \end{bmatrix}^{-1}$$

$$= \sqrt{\lambda} U(n)^T U(n-1) \cdot \cdot \cdot$$
where • indicates elements with values irrelevant to the result. Thus, $\hat{U}_s^T(n-1)\hat{U}_s(n)$ can be found from the upper left $\tilde{K} \times \tilde{K}$ matrix of the above matrix. The total multiplication count for this approach is $K\tilde{K}^2 + 2\tilde{K}$, which compares well with the multiplication count of $3M\tilde{K}^2 + M\tilde{K}$ for the F2 algorithm itself. However, if $\tilde{K}$ is large, the formula in the second line of equation (15) has less complexity.

VI. SIMULATION STUDIES

We consider a system with $K=7$ users with known codes, all with the same power, and a variable number $\tilde{K}$ of users with unknown codes. The users are assigned purely random codes of length $M=31$, and the SNR is 20 dB. An ensemble of 50 different random code assignments is generated, and the median signal to inference and noise ratio (SINR) is calculated over all code choices and users, in total an ensemble of 350. The SINR for a user is calculated using a moving average filter of length 20.

In all cases we consider 8 different detectors:

- **single user**: The conventional single user detector. This is the lower bound on performance of any multiuser detector.

- **Full**: The hypothetical non-blind MMSE multiuser detector that knows all codes. This can be taken as an upper bound of performance for any linear multiuser detector [10].

- **non-blind**: The non-blind MMSE detector that uses only the known codes and ignores the interference from the unknown users.

- **Direct**: The blind MMSE detector calculated using direct inversion of the correlation matrix (equation (8)).

- **Blind**: The blind MMSE detector calculated using (9) and using SVD for subspace calculation.

- **Group-blind SVD**: The group-blind detector calculated using (17) and using SVD for subspace calculation.

- **Group-blind 2**: The group-blind detector calculated using (16) and using SVD for subspace calculation.

- **Group-blind F2**: The group-blind detector calculated using (17) and using the F2 algorithm for subspace calculation with $T(n)$ calculated using (20).

Decorrelating detectors are not considered, since they in general have an inferior performance and are sensitive to estimation errors. Performance of the decorrelating detectors can be seen in [23]. For the F2 algorithm, it was initialized with an ordinary SVD over the first 50 samples, i.e., $U(0)$ and $\Sigma(0)$ in Table 1 was
initialized from the SVD over the first 50 samples. Without this initialization, the F2 algorithm diverged in the beginning, until it would start converging.

![Graph showing convergence with 7 known users and 4 unknown users.](image)

**Figure 1:** Convergence with 7 known users and 4 unknown users. The curve marked 1 is for the algorithms calculated using (17) and the one marked 2 for the algorithm (16).

Figure 1 shows the convergence with \( \tilde{K} = 4 \) unknown users, all of the same power as the known users. It can be seen that the group-blind detector has a performance gain over the blind detector starting at 4 dB and decreasing to 2 dB. As the number of bits increase both detectors should converge towards the full MMSE, and the performance gain go towards zero. It also seen that the performance of the group-blind detector is better than the non-blind MMSE already after 50 bits, whereas the blind detector requires 100 bits.

The performance of the group-blind algorithm using F2 is almost indistinguishable from the performance using SVD, and is therefore difficult to see in the figures. The performance of the group-blind algorithm using (16) is slightly worse than the one using (17). From other simulations we have made, this seems to be a general behavior. Since (17) has much less complexity than (16) for simultaneous detection, (17) must be prefered.

The performance gain of course depends on many factors: SNR, user powers, the codes used, and the number of users. Figure 2 shows the performance when the number of unknown users is increased to \( \tilde{K} = 10 \) users. Of these, 4 have the same power as the known users, while the 6 others have power -6dB. As expected, the performance gain of the group-blind detector over the blind is less than for Figure 1. Basically, when the number of unknown users is high, the information on known users does help relatively less. However, the gain is still around 2dB.
Figure 2: Convergence with 7 known and 10 unknown users.

It can also be seen that the performance of both blind and group-blind detectors is worse than the non-blind detector in the beginning. In this case, the group-blind detector has a further advantage to the blind detector: the non-blind part can be used alone until the blind part has converged. This cannot be done for the blind detector. An even better way is to replace (14) with the following formula

\[
\hat{b}(n) = \text{sgn} \left( \left( S^T S \right)^{-1} \left( S^T - \chi(n) T(n) \Lambda_{\chi}^{-1}(n) \tilde{U}_{\chi}^T(n) \right) \right)_{n}
\]

where \(0 \leq \chi(n) \leq 1\) is a weighting function with \(\chi(0) = 0\) and \(\chi(\infty) = 1\). The optimal (or even a good) way of choosing \(\chi(n)\) is a subject for further study. However, some preliminary tests show that this gives a superior performance in the beginning.

The previous results were under the assumption that \(\tilde{K}\) was known or estimated without errors. In Figure 3 the same data is used as for Figure 2, i.e., \(\tilde{K} = 10\), but the value of \(\tilde{K}\) used for the estimators is set to 4, i.e., equal to the number of high power users. Remarkably, the performance of both the blind and the group-blind algorithms improve, at least for a small number of bits, and in addition, the gain of the group-blind algorithm is increased, now again ranging between 2 and 4 dB. Two conclusions can be drawn

1. The estimators, both blind and group-blind, are relatively insensitive to estimation errors on \(\tilde{K}\).
2. The performance is better if only the largest eigenvalues are used in the subspace algorithms, rather than the full interference subspace.
Figure 3: Convergence with 7 known and 10 unknown users. The subspace algorithm, however, only uses a subspace dimension of 4.

Item 2 is in particular important. This can be interpreted so that it pays off to track only a few high power unknown users. In a cellular system, this means that it is advantageous to track only the out-of-cell users on the boundary of the cell considered, while the interference from other users inside neighboring cells are considered background noise because of path loss. This means that the number of users (or, rather, eigenmodes) that need to be tracked in the group-blind algorithms can be quite low, and low-complexity subspace tracking algorithms such as PASTd could be employed, resulting in a very low complexity system.

This conclusion is only true for a small number of bits. As the number of bits go towards infinity, the blind and group-blind algorithm with the correct value of $\tilde{K}$ should converge towards the full MMSE detector, which is not true if the estimated value of $\tilde{K}$ is too small. Thus, in a dynamic environment, the best value of $\tilde{K}$ depends not only on the number of users and their power, but also on the time scale for change in the environment.

In order to find the performance in terms of bit error rate (BER) versus SNR a different simulation scenario was used. Three detectors were considered: the non-blind MMSE detector, the blind MMSE detector, and the group-blind hybrid detector implemented using (17), with the latter two implemented using SVD for a fixed block size of 200. The data was generated in the following way: the number of known users was $K=7$, and the number of unknown users $\tilde{K} = 4$. An ensemble of 20 different code matrices was generated. For each code matrix, 100 different ensembles of 10,000 bits were generated. For each ensemble of 10,000 bits, the detectors were estimated over the 200 first bits, and the BER was calculated over all 10,000 bits, and averaged over the 100 different ensembles (i.e., the total number of bits were 1,000,000). Thus, for each SNR value, 20 code
matrices × 7 users = 140 ensembles were available. For these, the median and 90-percentile (i.e., the 14th worst) BER were calculated, shown in Figure 4. This gives a reasonable impression of performance, but does not completely reveal the performance of the detectors. By plotting each of the 140 ensembles and observing them manually, we made the following observations for high SNR (> 10 dB)

Figure 4: BER versus SNR for 7 known and 4 unknown users. For each detector, the lower curve is the median and the upper curve is the 90-percentile.

- The group-blind detector always had a lower BER than the blind detector
- The blind detector was in most cases better than the non-blind detector. In some cases, the non-blind detector was somewhat better (up to 1 dB).
- The group-blind detector was almost always better than the non-blind detector. In a few cases the non-blind detector was better, but then only insignificantly (0.1-0.2 dB).
- In some cases the non-blind detector had a very poor performance (which is reflected in the 90-percentile in Figure 4). The reason must be that there is an unknown user with a code very close to a known user. Thus, even a few high power out-of-cell users can corrupt the performance of a non-blind multiuser detector in a cellular system.

The last observation is reason enough that the non-blind detector should not be used when there is inter-cell interference. Since out-of-cell users cannot be controlled the same way that in-cell users can, this last case could happen, and as the worst-case performance essentially determines the quality of service, this would not be acceptable.
VII. CONCLUSION

In this paper we have developed a class of new multiuser detectors, called *group-blind* linear detectors. They are distinguished by the fact that they can suppress interference from both known and unknown users, like blind detectors, while simultaneously using the knowledge of known users. Simulations have shown that the group-blind detectors have notably better performance than both the pure blind detectors and the non-blind detectors, and furthermore they have a considerably lower computational complexity.

Our simulations have also shown that the interference from even a few out-of-cell users can totally corrupt the performance of a non-blind multiuser detector. This problem can efficiently be handled by the group-blind detector.

REFERENCES


**APPENDIX A: PROOF OF THEOREM 2**

We will approach the proof constructively. Denote in the following by \( \mathbf{w}_1 \) the weight vector for the decorrelating detector for user 1, and let \( \mathbf{v}_1 \) be defined by

\[
\mathbf{v}_1 = \mathbf{w}_1 \mathbf{s}_1, \quad \mathbf{W} = \mathbf{S} \left( \mathbf{S}^T \mathbf{S} \right)^{-2} \mathbf{S}^T
\]

(see also (5), and the discussion), and consider the subspace \( \mathbf{V} = \text{span} \left( \mathbf{U}_s \mathbf{v}_1 \right) \). Notice that \( \mathbf{U}_s^T \mathbf{s} = \mathbf{0} \) and \( \mathbf{U}_s^T \mathbf{v}_1 = \mathbf{0} \) by the definition of \( \mathbf{U}_s \). We must have \( \mathbf{w}_1 \in \mathbf{V} \), since \( \mathbf{V} \) is orthogonal to all interference by the other known users. Thus,
for some vector $c_1$. As in the proof of proposition 1 of [14], we now see that we can write

$$c_1 = \arg \min_c c^T \begin{bmatrix} \tilde{U}_s^T & v_1 \end{bmatrix} \begin{bmatrix} \Sigma A^2 S^T + \tilde{\Sigma A}^2 \tilde{S}^T & v_1 \end{bmatrix} c$$

$$= \arg \min_c c^T \begin{bmatrix} D & \tilde{U}_s^T \tilde{\Sigma A}^2 \tilde{S}^T v_1 \\ v_1^T \tilde{\Sigma A}^2 \tilde{S}^T \tilde{U}_s & v_1^T \begin{bmatrix} \Sigma A^2 S^T + \tilde{\Sigma A}^2 \tilde{S}^T \end{bmatrix} v_1 \end{bmatrix} c$$

subject to $w_1^T s_1 = 1$, with

$$D = \tilde{\Lambda}_s - \sigma^2 I_k$$

Now notice that

$$p_1 \overset{\text{def}}{=} \tilde{U}_s^T R v_1 = \tilde{U}_s^T \begin{bmatrix} \Sigma A^2 S^T + \tilde{\Sigma A}^2 \tilde{S}^T + \sigma^2 I \end{bmatrix} v_1 = \tilde{U}_s^T \Sigma A^2 \tilde{S}^T v_1$$

and

$$v_1^T R v_1 = v_1^T \begin{bmatrix} \Sigma A^2 S^T + \tilde{\Sigma A}^2 \tilde{S}^T + \sigma^2 I \end{bmatrix} v_1 = v_1^T \begin{bmatrix} \Sigma A^2 S^T + \tilde{\Sigma A}^2 \tilde{S}^T \end{bmatrix} v_1 + \sigma^2 v_1^2$$

Thus,

$$c_1 = \arg \min_c c^T \begin{bmatrix} D & p_1 \\ p_1^T & v_1^T R v_1 - \sigma^2 v_1^2 \end{bmatrix} c$$

Following the proof of [14] proposition 1, we then find that

$$w_1 = \frac{1}{v_1^T M v_1} M v_1$$

with

$$M = \begin{bmatrix} \tilde{U}_s & v_1 \end{bmatrix} \begin{bmatrix} D & p_1 \\ p_1^T & v_1^T R v_1 - \sigma^2 v_1^2 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{U}_s^T \\ v_1^T \end{bmatrix}$$

Here we apply a matrix inversion theorem (see, e.g., [15], section 2.9) to get

$$M = \begin{bmatrix} \tilde{U}_s & v_1 \end{bmatrix} \begin{bmatrix} D^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ -v_1^T R v_1 - p_1^T D^{-1} p_1 - \sigma^2 v_1^2 \end{bmatrix} \begin{bmatrix} D^{-1} p_1 p_1^T D^{-1} - D^{-1} p_1 \\ -D^{-1} p_1 p_1^T D^{-1} \end{bmatrix} \begin{bmatrix} \tilde{U}_s^T \\ v_1^T \end{bmatrix}$$
To calculate $\mathbf{Mv}_1$ notice that since $\tilde{\mathbf{U}}_s^T \mathbf{v}_1 = \mathbf{0}$ in fact only the last row of the matrix inside {} is needed, which yields

$$
\mathbf{Mv}_1 = \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_1^T \mathbf{R} \mathbf{v}_1 - \mathbf{p}_1^T \mathbf{D}^{-1} \mathbf{p}_1 - \sigma^2 \mathbf{v}_1^T \\ -\mathbf{D}^{-1} \mathbf{p}_1 \end{bmatrix} \begin{bmatrix} -\mathbf{D}^{-1} \mathbf{p}_1 \\ \mathbf{1} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_1^T \mathbf{R} \mathbf{v}_1 - \mathbf{p}_1^T \mathbf{D}^{-1} \mathbf{p}_1 - \sigma^2 \mathbf{v}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 - \tilde{\mathbf{U}}_s \mathbf{D}^{-1} \mathbf{p}_1 \end{bmatrix}
$$

Thus,

$$
\mathbf{w}_1^T \mathbf{r} = \kappa \mathbf{v}_1^T \left( \mathbf{I} - \mathbf{R} \tilde{\mathbf{U}}_s \mathbf{D}^{-1} \tilde{\mathbf{U}}_s^T \right) \mathbf{r}
$$

for some positive constant $\kappa$. Combining the solutions for the different users, we obtain \[13\].

**APPENDIX B: PROOF OF THEOREM 4**

The proof follows the proof of theorem 2 closely. Thus, at first we see that we can write

$$
\mathbf{w}_1 = \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} \mathbf{p}_1
$$

For some vector $\mathbf{c}_1$. By \([10]\) we now have, except for a scaling constant

$$
\mathbf{c}_1 = \left( \begin{bmatrix} \tilde{\mathbf{U}}_s^T \\ \mathbf{v}_1^T \end{bmatrix} \mathbf{R} \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} \right)^{-1} \begin{bmatrix} \tilde{\mathbf{U}}_s^T \\ \mathbf{v}_1^T \end{bmatrix} \mathbf{s}_1
$$

$$
\mathbf{w}_1 = \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} \left( \begin{bmatrix} \tilde{\mathbf{U}}_s^T \\ \mathbf{v}_1^T \end{bmatrix} \mathbf{R} \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} \right)^{-1} \begin{bmatrix} \tilde{\mathbf{U}}_s^T \\ \mathbf{v}_1^T \end{bmatrix} \mathbf{s}_1
$$

Similarly to the proof of theorem 2, we see that

$$
\begin{bmatrix} \tilde{\mathbf{U}}_s^T \\ \mathbf{v}_1^T \end{bmatrix} \mathbf{R} \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{U}}_s & \mathbf{p}_1 \\ \mathbf{p}_1^T \mathbf{R} \mathbf{v}_1 \end{bmatrix}
$$

We can now follow the same steps as the proof of theorem 2 with $\mathbf{D}$ replaced by $\tilde{\mathbf{U}}_s$ and $\mathbf{v}_1^T \mathbf{R} \mathbf{v}_1 - \sigma^2 \mathbf{v}_1^T$ replaced by $\mathbf{v}_1^T \mathbf{R} \mathbf{v}_1$, arriving at \[17\].