Optimal Grouping Algorithm for a Group Decision Feedback Detector in Synchronous CDMA Communications

J. Luo, *Member, IEEE*, K. Pattipati, *Fellow, IEEE*, P. Willett, *Fellow, IEEE*, and G. Levchuk

***Abstract*—The group decision feedback (GDF) detector is studied in this letter. Given the maximum group size, a grouping algorithm is proposed. It is shown that the proposed grouping algorithm maximizes the symmetric energy of the multiuser detection system. Furthermore, based on a set of lower bounds on asymptotic group effective energy (AGEE) of the GDF detector, it is shown that the proposed grouping algorithm, in fact, maximizes the AGEE lower bound for every group of users. The theoretical analysis of the grouping algorithm enables the offline estimation of the computational cost and the performance of a GDF detector. The computational complexity of a GDF detector is exponential in the largest size of the groups. Simulation results are presented to verify the theoretical conclusions. The results from this letter can be applied to the decision feedback detector by setting the maximum group size to one.**

***Index Terms*—Code-division multiple access (CDMA), decision feedback (DF), multiuser detection, optimization methods.**

I. Introduction

IN SYNCHRONOUS code-division multiple access (CDMA) communication systems, the near–far problem caused by the multi-access interference (MAI) has been widely studied. With the additive white Gaussian noise (AWGN) assumption and when the source signal is binary valued or integer valued, the conventional detector does not produce reliable decisions for the CDMA channel [1]. The computation of the optimal detector, however, is generally NP-hard, and thus, is exponential in the number of users [2], unless the signature waveform correlation matrix has a special structure [3]. Several new algorithms have been proposed to provide reliable solutions with relatively low computational cost. Among the suboptimal algorithms, the decision-driven detection methods, including decision feedback (DF) [4], [5], group detection [6], and multistage detection [7] are popular. Although the DF method is simple and performs well, there are situations when a marginal increase in computation can provide significant improvement in performance [8].

The main drawback of DF is that detections are made one user at a time; the decision on the strong user is obtained by treating the weak users as noise. However, when user signature sequences are correlated, this noise becomes biased, and thus, is naturally harmful to the userwise detection. The idea of sequential group detection was first introduced by Varanasi in [6] and can be viewed as the group decision feedback (GDF) detector. GDF detector divides users into several groups. The users with relatively high correlations are assigned to the same group, and the correlation between users in different groups are relatively low. Similar to the DF detector, the GDF detector makes decisions sequentially based on successive cancellation. However, instead of making decisions on a single user at a time, in GDF detection, the decisions on users in the same group (the correlated users) are made simultaneously. The computational expense for a GDF detector is approximately exponential in the largest group size, and this is expected to be small if the largest group size is small.

In [6], the sizes of the groups are design parameters. However, in practice, given a user signal set, it is not easy for one to find the correlated users and assign them to groups. Since the largest group size is closely related to the overall computational cost, in this letter, we consider the largest group size as the only design parameter. A grouping and ordering algorithm is proposed to find the optimal size and users for each group. Theoretical results are given to show its optimality in terms of the symmetric energy (SE). Together with a fast computational method modified from [8], the proposed GDF detection method provides an efficient way to improve DF detection with a marginal increase in computational cost. Simulation results on small- and large-size problems are presented to verify the theoretical conclusions.

II. Problem Formulation and Performance Measure of the GDF Detector

A discrete-time equivalent model for the matched-filter outputs at the receiver of a CDMA channel is given by the K-length vector [2]

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{b} \in \{-1, +1\}^K$ denotes the vector of bits transmitted by the K active users. Here, $\boldsymbol{H} = \boldsymbol{W}\boldsymbol{R}\boldsymbol{W}$ is a nonnegative definite signature waveform correlation matrix, $\boldsymbol{R}$ is the symmetric normalized correlation matrix with unit diagonal elements, $\boldsymbol{W}$ is a diagonal matrix whose kth diagonal element, w_k, is the square root of the received signal energy per bit of the kth user, and $\boldsymbol{n}$ is a real-valued zero-mean Gaussian random vector with a covariance matrix $\sigma^2\boldsymbol{H}$. It has been shown that this model holds

Paper approved by G. Cherubini, the Editor for CDMA Systems of the IEEE Communications Society. Manuscript received November 28, 2000; revised June 18, 2001; January 28, 2002; and May 22, 2002. This work was supported by the Office of Naval Research under Contract N00014-98-1-0465 and Contract N00014-00-1-0101, and by the Naval Undersea Warfare Center (NUWC) under Contract N66604-1-99-5021. This paper was presented in part at IEEE GlobeCom 2001, San Antonio, TX, November 2001.

The authors are with the Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT 06269 USA (e-mail: krishna@engr.uconn.edu).

Digital Object Identifier 10.1109/TCOMM.2003.809756

for both baseband [2] and passband [5] channels with additive Gaussian noise.

When all the user signals are equally probable, the optimal solution of (1) is the output of a maximum-likelihood (ML) detector [2]

$$\phi_{\rm ML}: \hat{\boldsymbol{b}} = \arg \min_{\boldsymbol{b}\in\{-1,+1\}^K} (\boldsymbol{b}^T\boldsymbol{H}\boldsymbol{b} - 2\boldsymbol{y}^T\boldsymbol{b}) \tag{2}$$

which is generally NP-hard and exponentially complex to implement.

Sequential group detection, based on the idea of successive cancellation, was first introduced by Varanasi in [6]. Suppose users are partitioned into an ordered set of P groups, $G_0,\ldots,G_{P-1}$. The number of users in group G_i is denoted by $|G_i|$, and naturally, $\sum_{i=0}^{P-1}|G_i| = K$. The decisions on group $\{G_0\}$ is made by

$$\hat{\boldsymbol{b}}_{G_0} = \arg \min_{\boldsymbol{b}_{G_0}\in\{-1,+1\}^{|G_0|}} \left[\min_{\boldsymbol{b}_{\bar{G}_0}}(\boldsymbol{b}^T\boldsymbol{H}\boldsymbol{b} - 2\boldsymbol{y}^T\boldsymbol{b})\right] \tag{3}$$

where $\boldsymbol{b}_{G_0}$ denotes the part of vector $\boldsymbol{b}$ that corresponds to users in group G_0, and $\bar{G}_0$ denotes the complement of G_0, i.e., the union of $G_1,\ldots,G_{P-1}$. The decisions of (3) are then used to subtract the MAI due to users in G_0 from the remaining decision statistics $\boldsymbol{y}_{\bar{G}_0}$. The detector for the next group G_1 is designed under the assumption that the MAI cancellation is perfect. This process of interference cancellation and group detection is carried out sequentially for users in groups $G_2,\ldots,G_{P-1}$, with the group detector for group G_i taking advantage of the decisions made by group detectors for $G_0,\ldots,G_{i-1}$. Denote the channel model for the user-expurgated channel that only has users in groups $G_i,\ldots,G_{P-1}$ by

$$\boldsymbol{y}^{(i)} = \boldsymbol{H}^{(i)}\boldsymbol{b}^{(i)} + \boldsymbol{n}^{(i)}. \tag{4}$$

The decisions on group G_i can be represented as

$$\begin{aligned}\hat{\boldsymbol{b}}_{G_i} &= \arg \min_{\boldsymbol{b}^{(i)}_{G_i}\in\{-1,+1\}^{|G_i|}} [f(\boldsymbol{b}_{G_i})]\\ f(\boldsymbol{b}_{G_i}) &= \min_{\boldsymbol{b}^{(i)}_{\bar{G}_i}} \left(\boldsymbol{b}^{(i)^T}\boldsymbol{H}^{(i)}\boldsymbol{b}^{(i)} - 2\boldsymbol{y}^{(i)^T}\boldsymbol{b}^{(i)}\right).\end{aligned} \tag{5}$$

In multiuser detection, the asymptotic effective energy (AEE) is an important performance measure. Similar to the definition of AEE [5], for a GDF detector, we can define the asymptotic group effective energy (AGEE) for each user group. Define the probability that not all users in group $\{G_i\}$ are detected correctly in a GDF detector ϕ as $P_{G_i}(\sigma,\phi)$. Then, the AGEE for group G_i is given by

$$E_{G_i}(\phi) = sup\left\{e\geq 0; \lim_{\sigma\to 0}\frac{P_{G_i}(\sigma,\phi)}{Q\left(\frac{\sqrt{e}}{\sigma}\right)} < \infty\right\} \tag{6}$$

where σ^2 is the additive noise variance [see (1)], and $Q(x) = \int_x^\infty (1/\sqrt{2\pi})e^{-z^2/2}dz$. Although an exact performance analysis of the GDF detector is intractable [6], one can obtain upper and lower bounds for the AGEE of all groups. In the above description of the GDF detector, define $\boldsymbol{J}^{(i)} = [\boldsymbol{H}^{(i)}]^{-1}$, and denote $\boldsymbol{J}^{(i)}_{G_iG_i}$ to be the submatrix of $\boldsymbol{J}^{(i)}$ that only contains the columns and rows corresponding to users in G_i. Define d_{G_i} to be the minimum distance of matrix $(\boldsymbol{J}^{(i)}_{G_iG_i})^{-1}$, i.e.

$$d^2_{G_i} = \min_{\boldsymbol{e}\in\{-1,0,1\}^{|G_i|}\backslash\{0\}^{|G_i|}} \boldsymbol{e}^T\left(\boldsymbol{J}^{(i)}_{G_iG_i}\right)^{-1}\boldsymbol{e} \tag{7}$$

where "\" denotes the set subtraction. The AGEE for group G_i can be bounded by

$$\min\left(d^2_{G_0},\ldots,d^2_{G_i}\right) \leq E_{G_i}(\phi) \leq d^2_{G_i}. \tag{8}$$

The upper bound in (8) is reached when all decisions on the users in group G_1 through group G_{i-1} are correct.

III. Optimal Grouping and Detection Order for GDF Detector

It is known that the performance of the decision-driven multiuser detector is significantly affected by the order of the users [1]. Since the overall computation for the GDF detector is exponential in the maximum group size, which is defined by $|G|_{\max} = \max(|G_0|,\ldots,|G_{P-1}|)$, in this section, we develop a grouping and ordering algorithm that maximizes the SE of the GDF detector given $|G|_{\max}$ as a design parameter. In the DF detector case, i.e., when $|G|_{\max} = 1$ for the GDF detector, the optimal user ordering is proposed in [5] and the proof of its asymptotic optimality can be found in [9].

Denote the Cholesky decomposition of $\boldsymbol{H}$ by $\boldsymbol{L}^T\boldsymbol{L} = \boldsymbol{H}$, where $\boldsymbol{L}$ is a lower triangular matrix. Multiply both sides of (1) by $(\boldsymbol{L}^{-1})^T$ to obtain the white noise model [4]

$$(\boldsymbol{L}^{-1})^T\boldsymbol{y} = \boldsymbol{L}\boldsymbol{b} + (\boldsymbol{L}^{-1})^T\boldsymbol{n}. \tag{9}$$

Define $\tilde{\boldsymbol{y}} = (\boldsymbol{L}^{-1})^T\boldsymbol{y}$, $\tilde{\boldsymbol{n}} = (\boldsymbol{L}^{-1})^T\boldsymbol{n}$, and partition the matrices and vectors according to G_0 and $\bar{G}_0$ to obtain

$$\begin{bmatrix}\tilde{\boldsymbol{y}}_{G_0}\\ \tilde{\boldsymbol{y}}_{\bar{G}_0}\end{bmatrix} = \begin{bmatrix}\boldsymbol{L}_{G_0G_0} & 0\\ \boldsymbol{L}_{\bar{G}_0G_0} & \boldsymbol{L}_{\bar{G}_0\bar{G}_0}\end{bmatrix}\begin{bmatrix}\boldsymbol{b}_{G_0}\\ \boldsymbol{b}_{\bar{G}_0}\end{bmatrix} + \begin{bmatrix}\tilde{\boldsymbol{n}}_{G_0}\\ \tilde{\boldsymbol{n}}_{\bar{G}_0}\end{bmatrix}. \tag{10}$$

Since $\boldsymbol{L}_{\bar{G}_0\bar{G}_0}$ is a full-rank matrix by assumption, the decisions on group G_0 in (3) can be written as

$$\hat{\boldsymbol{b}}_{G_0} = \arg \min_{\boldsymbol{b}_{G_0}\in\{-1,+1\}^{|G_0|}} \left\|\boldsymbol{L}_{G_0G_0}\boldsymbol{b}_{G_0} - \tilde{\boldsymbol{y}}_{G_0}\right\|_2^2. \tag{11}$$

Therefore, the AGEE of group G_0 is determined by the minimum distance of matrix $\boldsymbol{L}^T_{G_0G_0}\boldsymbol{L}_{G_0G_0}$. Since $\boldsymbol{H} = \boldsymbol{L}^T\boldsymbol{L}$, we have

$$\begin{aligned}\boldsymbol{L}^T_{G_0G_0}\boldsymbol{L}_{G_0G_0} &= \left[(\boldsymbol{H}^{-1})_{G_0G_0}\right]^{-1} = \left[\boldsymbol{J}^{(0)}_{G_0G_0}\right]^{-1}\\ E_{G_0}(\phi_{\rm GDFD}) &= d^2_{G_0}.\end{aligned} \tag{12}$$

A similar result can be obtained for group G_i. In the description of GDF detector in Section II, if we let $\boldsymbol{H}^{(i)} = \boldsymbol{L}^{(i)^T}\boldsymbol{L}^{(i)}$, then ${\boldsymbol{L}^{(i)^T}}_{G_i}{\boldsymbol{L}^{(i)}}_{G_i} = (\boldsymbol{J}^{(i)}_{G_iG_i})^{-1}$. Since $\boldsymbol{H}^{(i)}$ is the southeast subdiagonal matrix of $\boldsymbol{H}$, it is easy to see that $\boldsymbol{L}^{(i)}$ is the southeast subdiagonal matrix of $\boldsymbol{L}$ and ${\boldsymbol{L}^{(i)}}_{G_i} = \boldsymbol{L}_{G_i}$. Hence

$$\boldsymbol{L}^T_{G_iG_i}\boldsymbol{L}_{G_iG_i} = \left(\boldsymbol{J}^{(i)}_{G_iG_i}\right)^{-1}. \tag{13}$$

The above result shows that d_{G_i} is determined by the diagonal block matrix $\boldsymbol{L}_{G_i}$ of $\boldsymbol{L}$. Now, given all the decisions on group G_0 to group G_{i-1} are correct, denote the probability that not all the users in group G_i are detected correctly by

TABLE I
CHOICES OF GROUP G_0 AND THE CORRESPONDING $d^2_{G_0}$

User(s)	0	1	2	3	0,1
$d^2_{G_0}$	3.96	1.62	1.14	1.67	1.69
User(s)	0,2	0,3	1,2	1,3	2,3
$d^2_{G_0}$	1.14	1.68	1.74	1.62	1.24

$P_e(G_i|G_0,\ldots,G_{i-1}) \approx Q(d_{G_i}/\sigma)$. The probability that not all the K users are detected correctly can be represented as

$$P(\sigma,\phi) \approx 1 - \prod_{i=0}^{P-1}\left[1 - Q\left(\frac{d_{G_i}}{\sigma}\right)\right]. \tag{14}$$

Therefore, the SE of the GDF detector is given by

$$E(\phi_{\mathrm{GDFD}}) = \min\left(d^2_{G_0},\ldots,d^2_{G_{P-1}}\right). \tag{15}$$

Since $|G|_{\max}$ is given as a design parameter, the problem is then to find an optimal partition and detection order that maximizes $\min(d^2_{G_0},\ldots,d^2_{G_{P-1}})$. Notice that different GDF detectors may have the same $|G|_{\max}$, but different numbers of groups, since P is not a design parameter.

A. Grouping and Ordering Algorithm

Find the optimal grouping and detection order via the following steps.

Step 1) Partition the K users into two groups $\{G_0\}$ and $\{\bar{G}_0\}$ with $|G_0| \le |G|_{\max}$. Among these partitions ($\{G_0\}$ and $|G_0|$ are not fixed), select the one that maximizes d_{G_0} (which is the minimum distance of matrix $[\boldsymbol{J}^{(0)}_{G_0G_0}]^{-1}$).

Step 2) Partition the remaining $K - |G_0|$ users into two groups G_1 and $\bar{G}_1$ with $|G_1| \le |G|_{\max}$. Among these partitions, select the one that maximizes d_{G_1}.

Step 3) Continue this process until all the users are assigned to groups.

Example 1: The algorithm is illustrated by the following four-user example. Suppose the H matrix is given by

$$\boldsymbol{H} = \begin{bmatrix} 4.30 & 1.00 & 0.60 & 0.30 \\ 1.00 & 3.00 & 1.70 & 0.50 \\ 0.60 & 1.70 & 2.20 & 0.70 \\ 0.30 & 0.50 & 0.70 & 1.90 \end{bmatrix}. \tag{16}$$

Assume that the desired maximum group size is $|G|_{\max} = 2$. In Step 1 of the algorithm, the possible choices for group G_0 and the resulting $d^2_{G_0}$ are shown in Table I. The best choice for group G_0 is {user 0}. Then, for the user-expurgated channel, we have

$$\boldsymbol{H}^{(1)} = \begin{bmatrix} 3.00 & 1.70 & 0.50 \\ 1.70 & 2.20 & 0.70 \\ 0.50 & 0.70 & 1.90 \end{bmatrix}. \tag{17}$$

The possible choices for group G_1 and the resulting d_{G_1} are shown in Table II. We can see that the best choice for group G_1 is {user 1, user 2}. Naturally, {user 3} will be the last group. The resulting SE for this partitioning and ordering is $E = 1.78$.

Note that the above example has four users and $|G|_{\max} = 2$. One may think that partitioning users into two groups with two users in each group is a good choice. However, since user 0 is a strong user, this user has to be detected first. And since user 1 and user 2 are strongly correlated, they have to be assigned to the same group. If, for example, we assign two groups as {user 0, user 3} and {user 1, user 2}, then, as a penalty for detecting the weak user (user 3) as part of the first group, we obtain $E = 1.68 < 1.78$.

TABLE II
CHOICES OF GROUP G_1 AND THE CORRESPONDING $d^2_{G_1}$

User(s)	1	2	3	1,2	1,3	2,3
$d^2_{G_1}$	1.69	1.14	1.68	1.78	1.68	1.24

Proposition 1: The proposed grouping and ordering algorithm maximizes the SE in (15).

See the Appendix for the proof.

The grouping and ordering algorithm is also optimal in the following sense.

Proposition 2: The proposed grouping and ordering algorithm maximizes the performance lower bound in (8) for every group. In other words, suppose G is the grouping and ordering result obtained from the proposed algorithm, and G_k is one of the groups in G. Further suppose there is another group and detection sequence $\hat{G}$ with $\hat{G}_l$ being one of the groups in $\hat{G}$, and $\hat{G}_l = G_k$. The following result holds:

$$\min\left(d^2_{G_0},\ldots,d^2_{G_k}\right) \ge \min\left(d^2_{\hat{G}_0},\ldots,d^2_{\hat{G}_l}\right). \tag{18}$$

See the Appendix for the proof.

In addition to the above two propositions, we can derive a fast computational method for the GDF detector, which is a modified version of the method proposed in [8]. We suggest the following steps for the group detection.

B. Computational Method for GDF Detector

Suppose the GDF detector has P groups, $G_0,\ldots,G_{P-1}$.

1) Initialize $\tilde{\boldsymbol{y}}^{(1)} = (\boldsymbol{L}^{-1})^T\boldsymbol{y}$, $\boldsymbol{L}^{(1)} = \boldsymbol{L}$. Let $i = 0$.
2) Form the white-noise system model for the user-expurgated channel, partition the vectors and matrices according to group G_i and its complement $\bar{G}_i$ as

$$\begin{bmatrix} \tilde{\boldsymbol{y}}^{(i)}_{G_i} \\ \tilde{\boldsymbol{y}}^{(i)}_{\bar{G}_i} \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}^{(i)}_{G_iG_i} & 0 \\ \boldsymbol{L}^{(i)}_{\bar{G}_iG_i} & \boldsymbol{L}^{(i)}_{\bar{G}_i\bar{G}_i} \end{bmatrix} \begin{bmatrix} \boldsymbol{b}^{(i)}_{G_i} \\ \boldsymbol{b}^{(i)}_{\bar{G}_i} \end{bmatrix} + \begin{bmatrix} \tilde{\boldsymbol{n}}^{(i)}_{G_i} \\ \tilde{\boldsymbol{n}}^{(i)}_{\bar{G}_i} \end{bmatrix}. \tag{19}$$

Find the decisions on group G_i by

$$\hat{\boldsymbol{b}}_{G_i} = \arg\min_{\boldsymbol{b}_{G_i}\in\{-1,+1\}^{|G_i|}} \left\|\boldsymbol{L}_{G_iG_i}b_{G_i} - \tilde{\boldsymbol{y}}^{(i)}_{G_i}\right\|_2^2. \tag{20}$$

3) Compute $\tilde{\boldsymbol{y}}^{(i+1)}$ by

$$\tilde{\boldsymbol{y}}^{(i+1)} = \tilde{\boldsymbol{y}}^{(i)}_{\bar{G}_i} - \boldsymbol{L}^{(i)}_{\bar{G}_iG_i}\hat{\boldsymbol{b}}_{G_i} \tag{21}$$

Let

$$\boldsymbol{L}^{(i+1)} = \boldsymbol{L}^{(i)}_{\bar{G}_i\bar{G}_i}. \tag{22}$$

4) Let $i = i + 1$. If $i < P$, go to Step 2; otherwise, stop the computation.

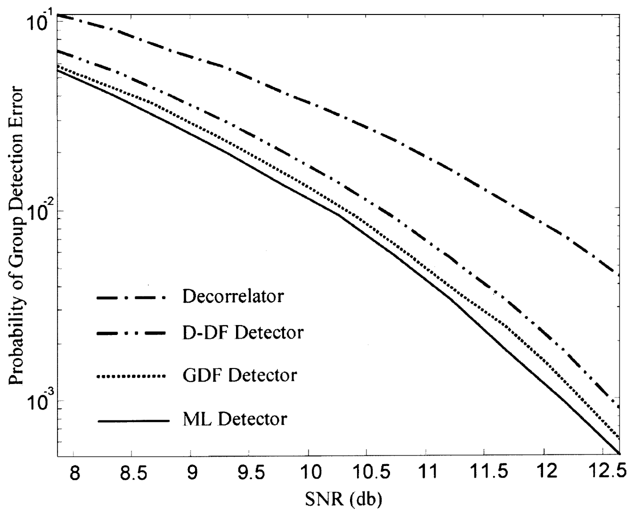

Fig. 1. Performance of various methods (four users, 1 000 000 Monte-Carlo runs).

The computational cost for Step 1 is $K(K+1)/2$ multiplications and $K(K-1)/2$ additions. Assume the computational cost for Step 2 can be bounded by

$$``\times" \leq M\left(|G_i|\right)\ , \quad ``+" \leq S\left(|G_i|\right) \tag{23}$$

where "$\times$" denotes the number of multiplications and "$+$" denotes the number of additions. In Step 3, since $\boldsymbol{b}$ can only take known discrete values, $\boldsymbol{Lb}$ can be precomputed and stored. Thus, only $|G_i|\sum_{k=i+1}^{P-1}|G_k|$ additions are needed. Therefore, the overall computational cost is bounded by

$$\begin{aligned} ``\times" &\leq \frac{K(K+1)}{2} + \sum_{k=0}^{P-1}\left[M\left(|G_k|\right)\right] \\ ``+" &\leq \frac{K(K+1)}{2} + \sum_{k=0}^{P-1}\left[S\left(|G_k|\right) + |G_k|\sum_{j=k+1}^{P-1}|G_j|\right]. \end{aligned} \tag{24}$$

IV. Simulation Results

Example 1-Continued: In the previous four-user example, $E(\phi_{\rm GDFD}) = 1.78$. The SE for the optimal decorrelating DF detector and the ML detector can be obtained from [5] as $E(\phi_{D-{\rm DFD}}) = 1.69$ and $E(\phi_{\rm ML}) = 1.8$. The simulation results are shown in Fig. 1, which are consistent with the theoretical analysis.

Example 2: In this example, we compare the computational loads for different multiuser detectors. We fix the signal-to-noise ratio at 12 dB and fix the maximum group size at 5. The square roots of the powers of the user signals are generated by $w_i \sim N(4.5, 4)$ $(N(.)$ represents the Gaussian distribution) and are limited within the range [2], [7] $2 \leq w_i \leq 7$. The signature sequences are randomly generated and the ratio between the spreading factor and the number of users is fixed at 1.2. Let the number of users vary from 5 to 60. Fig. 2 shows the worst-case computational complexity measured in terms of the number of multiplications plus number of additions of different detectors. Note that for the GDF detector, although the computation for finding the optimal user partitioning and user ordering is $O(K^{|G|_{\max}})$, this needs be done only once, offline. When $|G|_{\max}$ is small, as can be seen from the figure, the increase in the online computational cost of the GDF detector compared to D-DF detector is marginal. The computational methods used for the D-DF detector, the ML detector, as well as the search inside the groups of the GDF detector in this example, may be found in [8].

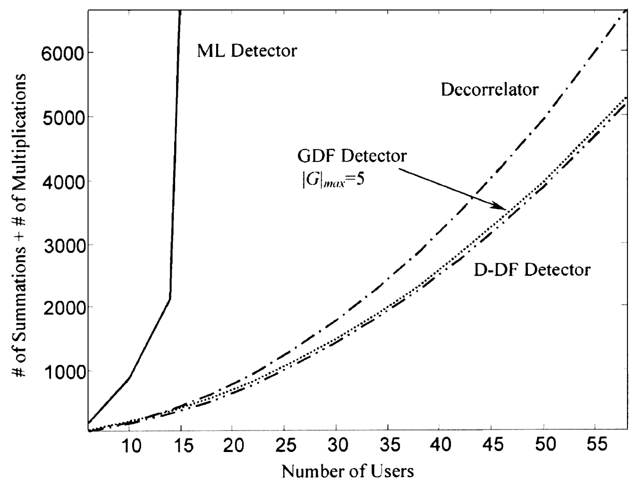

Fig. 2. Comparison of worst-case computational cost (random signature sequences, 10 000 Monte-Carlo runs).

V. Conclusion

An optimal grouping and ordering algorithm for the GDF detector is proposed. Together with a fast computational method based on the idea of branch and bound, the proposed algorithm provides a systematic way of improving the DF detector, especially when strong correlation exists among the users. In the meantime, when the maximum group size is relatively small, the increase in online computational cost of the proposed method compared to that of the DF detector is marginal. The theoretical results can be easily extended to finite-alphabet signals instead of binary ones.

Appendix

VI. Key Lemmas

Before proving the propositions in this letter, we present the following three lemmas that will be used in the proof.

Lemma 1: Suppose $\boldsymbol{H} = \boldsymbol{L}^T\boldsymbol{L}$ is partitioned on an arbitrary diagonal element as

$$\begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{21}^T \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix}. \tag{25}$$

For any permutation matrix $\boldsymbol{P}$ of the same size as $\boldsymbol{H}_{22}$, if

$$\begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{21}^T\boldsymbol{P} \\ \boldsymbol{P}^T H_{21} & \boldsymbol{P}^T H_{22}\boldsymbol{P} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{L}}_{11} & 0 \\ \tilde{\boldsymbol{L}}_{21} & \tilde{\boldsymbol{L}}_{22} \end{bmatrix}^T \begin{bmatrix} \tilde{\boldsymbol{L}}_{11} & 0 \\ \tilde{\boldsymbol{L}}_{21} & \tilde{\boldsymbol{L}}_{22} \end{bmatrix} \tag{26}$$

the following results hold.

$$\tilde{\boldsymbol{L}}_{11} = \boldsymbol{L}_{11}\ , \quad \tilde{\boldsymbol{L}}_{22}^T\tilde{\boldsymbol{L}}_{22} = \boldsymbol{P}^T L_{22}^T\boldsymbol{L}_{22}\boldsymbol{P}. \tag{27}$$

The proof is quite straightforward and is, therefore, omitted in this letter. □

Lemma 2: Suppose $\boldsymbol{H}$ is an $m \times m$ symmetric and positive definite matrix. Suppose $\boldsymbol{H} = \boldsymbol{L}^T\boldsymbol{L}$ is the Cholesky decomposition. Partition $\boldsymbol{H}$ and $\boldsymbol{L}$ on the last (southeast) diagonal component as

$$\begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{h}_{12} \\ \boldsymbol{h}_{12}^T & h_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{l}_{12}^T & l_{22} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{l}_{12}^T & l_{22} \end{bmatrix}. \tag{28}$$

Now move up the last user to the first, and denote the action and the new Cholesky decomposition matrix by

$$\begin{bmatrix} h_{22} & \boldsymbol{h}_{12}^T \\ \boldsymbol{h}_{12} & \boldsymbol{H}_{11} \end{bmatrix} = \begin{bmatrix} \tilde{l}_{11} & 0 \\ \tilde{\boldsymbol{l}}_{12} & \tilde{\boldsymbol{L}}_{22} \end{bmatrix}^T \begin{bmatrix} \tilde{l}_{11} & 0 \\ \tilde{\boldsymbol{l}}_{12} & \tilde{\boldsymbol{L}}_{22} \end{bmatrix}. \tag{29}$$

Then matrix $\tilde{\boldsymbol{L}}_{22}^T\tilde{\boldsymbol{L}}_{22} - \boldsymbol{L}_{11}^T\boldsymbol{L}_{11}$ is nonnegative definite.

Proof: Substituting (28) into (29) yields

$$\tilde{\boldsymbol{L}}_{22}^T\tilde{\boldsymbol{L}}_{22} - \boldsymbol{L}_{11}^T\boldsymbol{L}_{11} = \boldsymbol{l}_{12}\boldsymbol{l}_{12}^T \geq 0. \tag{30}$$

□

Lemma 3: Suppose $\boldsymbol{L}$ and $\tilde{\boldsymbol{L}}$ are two lower triangular matrices of size $m \times m$, assume that $\boldsymbol{L}^T\boldsymbol{L} - \tilde{\boldsymbol{L}}^T\tilde{\boldsymbol{L}} \geq 0$. Partition $\boldsymbol{L}$ on an arbitrary diagonal component, and partition $\tilde{\boldsymbol{L}}$ accordingly as

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix}, \quad \tilde{\boldsymbol{L}} = \begin{bmatrix} \tilde{\boldsymbol{L}}_{11} & 0 \\ \tilde{\boldsymbol{L}}_{21} & \tilde{\boldsymbol{L}}_{22} \end{bmatrix}. \tag{31}$$

We have

$$\boldsymbol{L}_{11}^T\boldsymbol{L}_{11} - \tilde{\boldsymbol{L}}_{11}^T\tilde{\boldsymbol{L}}_{11} \geq 0\,, \quad \boldsymbol{L}_{22}^T\boldsymbol{L}_{22} - \tilde{\boldsymbol{L}}_{22}^T\tilde{\boldsymbol{L}}_{22} \geq 0. \tag{32}$$

Proof: Since $\boldsymbol{L}^T\boldsymbol{L} - \tilde{\boldsymbol{L}}^T\tilde{\boldsymbol{L}} \geq 0$, we can find a lower triangular matrix $\boldsymbol{C}$ which satisfies

$$\boldsymbol{L}^T\boldsymbol{L} = \tilde{\boldsymbol{L}}^T\left(\boldsymbol{I} + \boldsymbol{C}^T\boldsymbol{C}\right)\tilde{\boldsymbol{L}}. \tag{33}$$

According to (31), partition $\boldsymbol{C}$ as

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & 0 \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}. \tag{34}$$

Substitute (31) and (34) into (33) to obtain

$$\begin{aligned} \boldsymbol{L}_{22}^T\boldsymbol{L}_{22} &= \tilde{\boldsymbol{L}}_{22}^T\left(\boldsymbol{I} + \boldsymbol{C}_{22}^T\boldsymbol{C}_{22}\right)\tilde{\boldsymbol{L}}_{22} \\ \boldsymbol{L}_{11}^T\boldsymbol{L}_{11} &= \tilde{\boldsymbol{L}}_{11}^T\left(\boldsymbol{I} + \boldsymbol{C}_{11}^T\boldsymbol{C}_{11}\right)\tilde{\boldsymbol{L}}_{11} + \triangle \end{aligned} \tag{35}$$

where $\triangle$ is a symmetric nonnegative definite matrix. The proof is complete. □

Note that in *Lemma 3*, we can continue partitioning the subdiagonal block matrices, and apply *Lemma 3* iteratively to obtain a result similar to (32) for an arbitrary partition.

VII. Proof of *Proposition 1*

Denote the optimal group and detection sequence determined by the proposed algorithm as G, which has groups $G_0, \ldots, G_{P-1}$. Denote the GDF detector using detection sequence G by $\phi_{G-\mathrm{GDFD}}$. The idea of the proof can be summarized as follows. Suppose there is another group and detection sequence $G^{(i)}$, which has groups $G_0^{(i)}, \ldots, G_{P^{(i)}-1}^{(i)}$. Without loss of generality, assume $\forall j(0 \leq j < i) G_j^{(i)} = G_j$ (the superscript (i) means that the first i groups in $G^{(i)}$ are identical to the first i groups in G).

Now construct a new group and detection sequence $G^{(i+1)}$. The groups of $G^{(i+1)}$ are defined by

$$\begin{cases} G_j^{(i+1)} = G_j^{(i)} = G_j, & 0 \leq j < i \\ G_j^{(i+1)} = G_j, & \text{j=i} \\ G_j^{(i+1)} = G_{j-1}^{(i)} \setminus G_i, & \text{j>i.} \end{cases} \tag{36}$$

To simplify the notation, in the above construction, if $G_j^{(i+1)} = \mathrm{NULL}$, we still keep group $G_j^{(i+1)}$ and define $d^2_{G_j^{(i+1)}} = \infty$. Evidently, $G^{(i+1)}$ has one more group than $G^{(i)}$. The following result holds for $G^{(i+1)}$.

Proposition 3: If $G^{(i+1)}$ is constructed according to the above definition, then:

1) $\forall j(0 \leq j < i)$, $d^2_{G_j^{(i+1)}} = d^2_{G_j^{(i)}}$;
2) $d^2_{G_i^{(i+1)}} \geq d^2_{G_i^{(i)}}$;
3) $\forall j(i < j \leq P^{(i)})$, $d^2_{G_j^{(i+1)}} \geq d^2_{G_{j-1}^{(i)}}$.

Proof:

1) For any $j < i$, the decision for group $G_j^{(i)}$ is made by treating the signal corresponding to $G_{j+1}^{(i)}, \ldots, G_{P^{(i)}-1}^{(i)}$ as noise and minimizing the probability of error in the ML sense. Therefore, any swapping of users within groups of index larger than j will not affect the performance of $G_j^{(i)}$. This result can be formally proved using *Lemma 1*.
2) Since $G_j^{(i)} = G_j^{(i+1)} (\forall j < i)$, this result can be directly obtained from the definition of the optimal grouping and ordering algorithm.
3) The proof for this part is relatively tricky. In fact, the construction of $G^{(i+1)}$ from $G^{(i)}$ can be divided into three stages. Define the users in group G_i as $K_0, \ldots, K_{|G_i|-1}$. For the convenience of discussion, we first consider user K_0.

Stage 1: Suppose, in $G^{(i)}$, user K_0 belongs to group $G_j^{(i)} (j \geq i)$. Define the action "take out user K_0 from group $G_j^{(i)}$," which converts $G^{(i)}$ to $G^{(S1)}$, as

$$\begin{cases} G_k^{(S1)} = G_k^{(i)}, & \text{k<j} \\ G_k^{(S1)} = \{\mathrm{user} K_0\}, & \text{k=j} \\ G_k^{(S1)} = G_j^{(i)} \setminus \{\mathrm{user} K_0\}, & \text{k=j+1} \\ G_k^{(S1)} = G_{k-1}^{(i)}, & \text{k>j+1.} \end{cases} \tag{37}$$

Stage 2: Now in $G^{(S1)}$, we have $G_j^{(S1)} = \{\mathrm{user} K_0\}$. Define the action "move up user K_0 to follow group $G_{i-1}^{(S1)}$," which converts $G^{(S1)}$ to $G^{(S2)}$, as follows:

$$\begin{cases} G_k^{(S2)} = G_k^{(S1)}, & \text{k<i} \\ G_k^{(S2)} = \{\mathrm{user} K_0\}, & \text{k=i} \\ G_k^{(S2)} = G_{k-1}^{(S1)}, & i < k \leq j \\ G_k^{(S2)} = G_k^{(S1)}, & \text{k>j.} \end{cases} \tag{38}$$

Continue performing the above two stages on all users $K_0, \ldots, K_{|G_i|-1}$. Denote the resulting group and detection sequence as $G^{(S3)}$. Denote the number of groups in $G^{(S3)}$ by $P^{(S3)}$.

Stage 3: In $G^{(S3)}$, combine groups $\{K_{|G_i|-1}\}, \ldots, \{K_0\}$, which converts $G^{(S3)}$ to $G^{(i+1)}$, as

$$\begin{cases} G_k^{(i+1)} = G_k^{(S3)}, & \text{k<i} \\ G_k^{(i+1)} = \{\text{user} K_0, \ldots, K_{|G_i|-1}\}, & \text{k=i} \\ G_k^{(i+1)} = G_{k-|G_i|+1}^{(S3)}, & \text{k>i.} \end{cases} \quad (39)$$

In the first stage, without loss of generality, suppose user K_0 is the first user in group $G_j^{(i)}$. The "take out" action does not change the order of the users, thus, the Cholesky decomposition matrix $\boldsymbol{L}$ remains unchanged. This shows that $\boldsymbol{L}_{G_{j+1}^{(S1)} G_{j+1}^{(S1)}}$ is the southeast diagonal subblock of $\boldsymbol{L}_{G_j^{(i)} G_j^{(i)}}$. Therefore

$$d^2_{G_{j+1}^{(S1)}} \geq d^2_{G_j^{(i)}}. \quad (40)$$

In the second stage, since the "minimum distance" of a subblock is the performance measure for the corresponding user group, given all the user groups with smaller indexes are correctly detected, putting more users into the detected user list will result in a better performance and a larger "minimum distance." In fact, from *Lemma 2* and *Lemma 3*, for any groups $G_k^{(S2)} = G_{k-1}^{(S1)}$, $i < k \leq j$, we have

$$\boldsymbol{L}^T_{G_k^{(S2)} G_k^{(S2)}} \boldsymbol{L}_{G_k^{(S2)} G_k^{(S2)}} - \boldsymbol{L}^T_{G_{k-1}^{(S1)} G_{k-1}^{(S1)}} \boldsymbol{L}_{G_{k-1}^{(S1)} G_{k-1}^{(S1)}} \geq 0. \quad (41)$$

Hence, in $G^{(i+1)}$, for any $j > i$, $d^2_{G_j^{(i+1)}} \geq d^2_{G_{j-1}^{(i)}}$, which proves part (3) of *Proposition 3*. □.

Proposition 3 shows that

$$E(\phi_{G^{(i+1)}-\text{GDFD}}) \geq E(\phi_{G^{(i)}-\text{GDFD}}). \quad (42)$$

By iteratively using the above construction procedure in the proof of *Proposition 1*, we will finally get $G^{(P)} = G$ and

$$E(\phi_{G^{(P)}-\text{GDFD}}) \geq E(\phi_{G^{(i)}-\text{GDFD})} \quad (43)$$

which completes the proof. □

VIII. Proof of *Proposition 2*

In the above proof for *Proposition 1*, let $G^{(i)} = \hat{G}$. Construct $G^{(i+1)}$ using the same procedure. Note that $G_l^{(i)} = \hat{G}_l = G_k$, and $G_k \cap G_i = \text{NULL}$. Therefore, in $G^{(i+1)}$, we have $G_{l+1}^{(i+1)} = G_k$. And

$$\min\left(d^2_{G_0^{(i+1)}}, \ldots, d^2_{G_{l+1}^{(i+1)}}\right) \geq \min\left(d^2_{\hat{G}_0}, \ldots, d^2_{\hat{G}_l}\right). \quad (44)$$

By iteratively using the construction procedure, we will finally get $G^{(P)} = G$, which satisfies

$$\min\left(d^2_{G_0}, \ldots, d^2_{G_k}\right) \geq \min\left(d^2_{\hat{G}_0}, \ldots, d^2_{\hat{G}_l}\right). \quad (45)$$

Hence, the proof is complete. □

Acknowledgment

The authors would like to thank the anonymous reviewers for their careful reviews and valuable comments.

References

[1] S. Verdu, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.
[2] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-devision multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.
[3] C. SanKaran and A. Ephremides, "Solving a class of optimum multiuser detection problems with polynomial complexity," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1958–1961, Sept. 1998.
[4] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channel," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.
[5] M. K. Varanasi, "Decision feedback multiuser detection: a systematic approach," *IEEE Trans. Inform. Theory*, vol. 45, pp. 219–240, Jan. 1999.
[6] ——, "Group detection for synchronous Gaussian code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1083–1096, July 1995.
[7] ——, "Multistage detection for synchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, Apr. 1990.
[8] J. Luo, K. Pattipati, P. Willett, and L. Brunel, "Branch-and-bound-based fast optimal algorithm for multiuser detection in synchronous CDMA," in *Proc. IEEE ICC'03*, Anchorage, AK, May 2003.
[9] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSE*, Pisa, Italy, Sept. 1998.