A Class of Coordinate Descent Methods for Multiuser Detection

J. Luo, G. Levchuk, K. Pattipati, P. Willett

University of Connecticut

Department of Electrical and Systems engineering

Storrs, CT 06269, USA

E-mail: krishna@sol.uconn.edu

ABSTRACT

A class of coordinate descent methods is proposed for the joint detection of binary symbols of K users in a synchronous correlated waveform multiple-access (CWMA) channel with Gaussian noise. We consider the detection problem as one of optimizing a quadratic objective function with binary constraints on decision variables. The proposed coordinate descent methods, while still maintaining a low computational complexity, are shown to provide as much as two orders of magnitude improvement in the probability of error, especially in situations where the existing methods do not perform well. The paper concludes with a discussion of how the proposed methods can be further improved. *

1. INTRODUCTION

Due to the problem of Interuser Interference (IUI) in many multiuser communications, multiuser detection for the symbol-synchronous Gaussian correlated waveform multiple-access (CWMA) channel has received considerable attention over the past ten years. Linear algorithms, such as the Minimum Mean Square Error (MMSE) and the decorrelation methods, are already well known [1-7]. When the user symbols are from a finite alphabet, all linear detectors need to perform a projection to satisfy the integrality constraints. Due to the discrete nature of the objective function, projection can cause significant errors, especially when the signal waveform correlation matrix is ill conditioned or when the signal-to-noise ratio is small. Indeed, since the multiuser detection in a CWMA channel with Gaussian noise, under the Maximum-Likelihood (ML) criterion, is a quadratic optimization problem with binary decision variables, it is NP-hard [1] unless the signal correlations have a special structure [2].

Recently, several sub-optimal and lower complexity alternative algorithms have been proposed for this problem. These include the sequential detection [3], cyclic decision feedback sequential detection [4] and other multistage detectors [5]. Group detection was introduced in [6], where a subset of users is jointly detected. Based on the idea of successive cancellation, a systematic decision feedback approach was also given in [7]. However, when the signature waveforms are significantly correlated, existing methods perform poorly. In this paper, we consider a class of coordinate descent methods to significantly improve the performance of a conventional detectors. Simulation results show that these methods can provide as much as two orders of magnitude improvement in probability of error when compared to the linear and decision feedback methods, while still maintaining a low computational complexity.

The paper is organized as follows. The synchronous multiuser detection problem formulation and existing solution techniques are discussed in section 2. In section 3, the proposed coordinate descent methods are presented. Simulation results and comparative analysis of the various algorithms are provided in section 4. Section 5 concludes with a summary and suggestions for further refinements.

*This work was supported by the Office of Naval Research under contract #N00014-93-1-0793

2. PROBLEM FORMULATION AND EXISTING METHODS

2.1 Problem Formulation

A discrete-time equivalent model for the matched-filter outputs at the receiver of a CWMA channel is given by the K-length vector [7]

$$y = Hb + n \tag{1}$$

where $b \in \{-1,+1\}^K$ denotes the K-length vector of bits transmitted by the K active users. Here $H = W^{\frac{1}{2}} R W^{\frac{1}{2}}$ is a nonnegative signature waveform correlation matrix, R is the symmetric normalized correlation matrix with unit diagonal elements, W is a diagonal matrix whose k-th diagonal element w_k is the received signal energy per bit of the k-th user, and n is a real-valued zero-mean Gaussian random vector with a covariance matrix $\sigma^2 H$. It has been shown that this model holds for both baseband [1] and passband [7] channels with additive Gaussian noise.

When all the user signals are equally probable, the optimal solution of (1) is the output of a ML detector [1] (also a Maximum A Posteriori (MAP) receiver in this case)

$$\phi_{ML}: \hat{b} = \arg\min (y - Hb)^T H^{-1} (y - Hb)$$
$$= \arg\min b^T Hb - 2y^T b$$
$$\text{subject to: } b \in \{-1,+1\}^K \tag{2}$$

The ML detector has the property that it minimizes, among all detectors, the probability that not all users' decisions are correct.

Since ϕ_{ML} is in general NP-hard and exponentially complex to implement, the focus is on developing easily implementable and effective multiuser detectors.

2.2 Existing Sub-optimal Algorithms

A. Conventional Detector

The MMSE-based solution of conventional detector

$$\phi_c : \hat{b} = \arg\min_{b\in\{-1,1\}^K} \left\| b - H^{-1}y \right\|_2^2 \qquad (3)$$

is found in two steps. First, the unconstrained solution $\tilde{b} = H^{-1}y$ is computed. This is then projected onto the constraint set via: $\hat{b}_i = sign(\tilde{b}_i)$.

B. Decorrelation Detector

The decorrelation detector [1] is similar to a conventional detector with a modified objective function:

$$\phi_D : \hat{b} = \arg\min_{b\in\{-1,1\}^K} \left\| b - \left[H + \sigma^2 I\right]^{-1} y \right\|_2^2 \qquad (4)$$

The solution for unconstrained problem (4) can be viewed as a MMSE output of

$$\min (y - Hb)^T H^{-1}(y - Hb) + \sigma^2 (b'b - K) \qquad (5)$$

which is equivalent to (1) under the binary constraints on b_i.

Although there exist other types of linear detectors [1], all suffer from errors caused by the projection operation.

C. Decision Feedback Detector(DFD)

The DFD is described by a set of two matrices (F, B) such that

$$\phi_{C-DFD} : \hat{b} = \operatorname{sgn}\left(Fy - B\tilde{b}\right) \qquad (6)$$

where $\tilde{b} = \arg\min \|b - Fy\|_2^2$ subject to: $b \in \{-1,+1\}^K$.

The general form of DFD based on successive cancellation is described in [7] as

$$\phi_{DFD} : \hat{b}_i = \operatorname{sgn}\left(\sum_{j=1}^{K} F_{ij} y_j - \sum_{j=1}^{i-1} B_{ij} \hat{b}_j \right) \qquad (7)$$

Here, F is an upper triangular matrix and B is a strictly lower triangular part of FH. When the i-th row of F is obtained by the first row of the MMSE linear user-expurgated detector for users i to K, the DFD detector is called "optimal" DFD (O-DFD). It has been shown [7] that the performance of the DFD varies greatly depending on the order of successive cancellation. If we define the most-powerful-user be a user with minimum probability of error among the rest of users corresponding to a MMSE linear detector, then the optimal DFD achieves the best performance for the cancellation order corresponding to the descending order of the most-powerful-users [7]. An alternative order can also be obtained by calculating the descent order of the normalized diagonal elements of Cholesky factorization of R [7]. Due to the discrete nature of the objective function, the linear and decision feedback methods suffer from errors caused by projection operation. Proposed coordinate descent methods are designed to overcome this type of error.

3. COORDINATE DESCENT ALGORITHMS

The problem posed in (2) is equivalent to the following quadratic 0-1 programming problem:

$$\hat{x} = \arg\min_{x\in\{0,1\}^K} \frac{1}{2} x^T Q x + c^T x \qquad (8)$$

where $x = \frac{(b+e)}{2}, c = -(y + He), Q = 2H$ and $e = [1\,1..1]'$.

The problem (8) is equivalent to

$$\hat{x} = \arg\min_{x\in\{0,1\}^K} \sum_{i<j} q_{ij} x_i x_j + \sum_{i=1}^{K} \left(c_i + \frac{1}{2} q_{ii} \right) x_i \quad (9)$$

At each step, the algorithm, called *Descent-1*, finds the largest decrease $w(i)$ to the objective function value for a one-variable change of the form $x_i = 1 - x_i$

$$w(i) = \left(\sum_{j=1, j\neq i}^{K} q_{ij} x_j + c_i + \frac{1}{2} q_{ii} \right) (2x_i - 1) \qquad (10)$$

This process can be viewed as finding a discrete local minimizer in the neighborhood of the point $x = (x_1, \cdots, x_K)$ of the form $\left\{x^k = (x_1, \cdots x_{k-1}, 1 - x_k, x_{k+1}, \cdots, x_K), k = 1, \cdots, K\right\}$. The algorithm stops when no decrease in function value in any neighboring direction can be found.

Considering the two-variable change, the algorithm, called *Descent-2*, finds the local minimum in the neighborhood of the point $x = (x_1, \cdots, x_K)$ of the form $\left\{x^{k,m} = (x_1, \cdots x_{k-1}, 1 - x_k, x_{k+1}, \cdots, x_{m-1}, 1 - x_m, x_{m+1}, \cdots, x_K), k, m = 1, \cdots, K\right\}$

For each k,m such that $k \neq m$, the "decrease" coefficients are computed as

$$\begin{aligned} v(k,m) &= \left(\sum_{j=1, j\neq k}^{K} q_{kj} x_j + c_k + \frac{1}{2} q_{kk} \right)(2x_k - 1) \\ &+ \left(\sum_{j=1, j\neq m}^{K} q_{mj} x_j + c_m + \frac{1}{2} q_{mm} \right)(2x_m - 1) \\ &- q_{km}(2x_k - 1)(2x_m - 1) \\ &= w(k) + w(m) - q_{km}(2x_k - 1)(2x_m - 1) \end{aligned}$$

In the worst case, the *Descent-1* algorithm requires exponential number of iterations. As noted in [9], it is an open question whether the problem of finding a discrete local minimizer of this form is NP-complete. However, in practice, the *Decent-1* algorithm performs polynomially for the specific problem in (9).

Decent-1 Algorithm:

Initialization:

x=*solution of O-DFD*;

$$w(i)=\left(\sum_{j=1,j\neq i}^{K} q_{ij}x_j + c_i + \frac{1}{2}q_{ii}\right)(2x_i-1)$$

$$f^*=f(x)$$

Step of the algorithm:

Find $i^*=\arg\max\{w(i), i=1,\cdots,K\}$. Values $w(i)$ are updated:

$$w(j)=\begin{cases} w(j)-q_{ji^*}(2x_{i^*}-1)(2x_j-1), & j\neq i^* \\ -w(i^*), & j=i^* \end{cases} \quad (11)$$

Then $x_{i^*}=1-x_{i^*}$ and $f^*=f^*-w(i^*)$.

Termination: When $w(i)\leq 0$ for $\forall i$, stop. Report $x^*=x$ as the solution.

Descent-2 Algorithm:

Initialization:

x=*solution of Descent-1*;

$$w(i)=\left(\sum_{j=1,j\neq i}^{K} q_{ij}x_j + c_i + \frac{1}{2}q_{ii}\right)(2x_i-1)$$

$$v(k,m)=w(k)+w(m)-q_{km}(2x_k-1)(2x_m-1)$$

$$f^*=f(x)$$

Step of the algorithm:

Find $i^*=\arg\max\{w(i), i=1,\cdots,K\}$,

$[k^*,m^*]=\arg\max\{v(k,m), k\neq m\}$

if $\max\{w(i), i=1,\cdots,K\}\geq \max\{v(k,m), k\neq m\}$

then

$$w(j)=\begin{cases} w(j)-q_{ji^*}(2x_{i^*}-1)(2x_j-1), & j\neq i^* \\ -w(i^*), & j=i^* \end{cases}$$

Then $x_{i^*}=1-x_{i^*}$, $f^*=f^*-w(i^*)$

$$v(k,m)=w(k)+w(m)-q_{km}(2x_k-1)(2x_m-1)$$

else

$$w(j)=\begin{cases} w(j)-q_{jk^*}(2x_{k^*}-1)(2x_j-1) \\ -q_{jm^*}(2x_{m^*}-1)(2x_j-1), & j\neq k^*,m^* \\ -w(k^*)-q_{jm^*}(2x_{m^*}-1)(2x_j-1), & j=k^* \\ -w(m^*)-q_{jk^*}(2x_{k^*}-1)(2x_j-1), & j=m^* \end{cases}$$

Then $x_{k^*}=1-x_{k^*}$, $x_{m^*}=1-x_{m^*}$

$$v(k,m)=w(k)+w(m)-q_{km}(2x_k-1)(2x_m-1)$$

and $f^*=f^*-v(k^*,m^*)$

endif

Termination: When $w(i)\leq 0$ and $v(k,m)\leq 0$ for $\forall i,k\neq m$, stop. Report $x^*=x$ as the solution.

4. SIMULATION RESULTS

In this section, we compare the performance of coordinate descent methods with other existing algorithms. A 10-user CWMA channel is considered. The signature waveform correlation matrix H is generated as follows:

First, the signature correlation matrix R is found by computing the correlation between randomly generated signature sequences of 14-bit length each. The condition number of R is chosen to be greater than 20. Second, the signal energies (i.e., the elements of matrix W) are generated within the range $[1,3.0]$, and H is calculated as $H=W^{\frac{1}{2}}RW^{\frac{1}{2}}$.

In our example,

$$H=\begin{bmatrix} 1 & -0.41 & -0.10 & -0.42 & -0.50 & 0.66 & 0.26 & -0.09 & 0.42 & -0.30 \\ -0.41 & 1.13 & -0.69 & 0.09 & -0.32 & -0.3 & -0.28 & 0.10 & -0.09 & -0.53 \\ -0.10 & -0.69 & 1.43 & 0.10 & 0.59 & 0.56 & 0.52 & 0.11 & -0.50 & 0.84 \\ -0.42 & 0.09 & 0.10 & 1.20 & 0.55 & -0.31 & 0.10 & -0.72 & -0.46 & -0.11 \\ -0.50 & -0.32 & 0.59 & 0.55 & 1.67 & -0.36 & 0.11 & -0.37 & -0.11 & 0.39 \\ 0.66 & -0.3 & 0.56 & -0.31 & -0.36 & 1.49 & 0.53 & 0.12 & -0.10 & 0.37 \\ 0.26 & -0.28 & 0.52 & 0.10 & 0.11 & 0.53 & 1.27 & -0.32 & -0.09 & -0.11 \\ -0.09 & 0.10 & 0.11 & -0.72 & -0.37 & 0.12 & -0.32 & 1.50 & 0.31 & 0.37 \\ 0.42 & -0.09 & -0.50 & -0.46 & -0.11 & -0.10 & -0.09 & 0.31 & 1.17 & -0.33 \\ -0.30 & -0.53 & 0.84 & -0.11 & 0.39 & 0.37 & -0.11 & 0.37 & -0.33 & 1.7 \end{bmatrix}$$

The SNR is chosen so that the probability of error (calculated based on importance sampling of 10000 Monte-Carlo runs) is similar to that observed in real applications.

The maximum likelihood detector is also presented to show the gap in probability of error between the optimal and sub-optimal solutions. This optimal solution is obtained by the branch-and-bound algorithm [8]. The idea of variable fixing according to the gradient range [10] is utilized. It is based on the idea that variables whose partial derivatives have fixed sign in $[0,1]$ can be fixed to 0 or 1. This remains true for the $\{0,1\}^K$ case as well. In our problem, the maximum range of the gradient is

$$lb_i \leq \frac{\partial f(x)}{\partial x_i} \leq ub_i,$$

where $lb_i=\sum_{j=1,j\neq i}^{K} q_{ij}^{-} + \frac{1}{2}q_{ii} + c_i, ub_i=\sum_{j=1,j\neq i}^{K} q_{ij}^{+} + \frac{1}{2}q_{ii} + c_i$

The complexity of ML detector is exponential.

Figure 1 shows the probability of error for group detection. Detailed data is given in table 1.

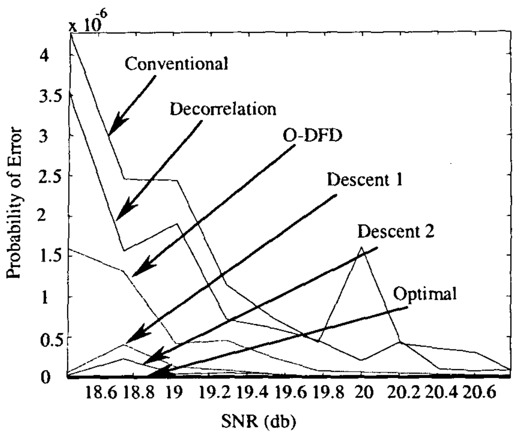

Figure 1. Performance of various methods

SNR	Conventional	Decorrelation	O-DFD	Descent 1	Descent 2	Optimal
18.45	4.28E-06	3.54E-06	1.59E-06	5.74E-08	3.05E-08	1.68E-13
18.75	2.44E-06	1.56E-06	1.31E-06	4.02E-07	2.32E-07	5.41E-13
19.03	2.43E-06	1.89E-06	4.08E-07	1.23E-07	3.81E-08	3.51E-16
19.29	1.15E-06	7.17E-07	4.55E-07	8.45E-08	5.05E-08	7.24E-17
19.54	7.32E-07	5.99E-07	2.41E-07	3.75E-08	3.38E-08	1.45E-15
19.78	4.29E-07	4.45E-07	7.87E-08	5.41E-09	1.73E-09	1.58E-15
20	1.61E-06	2.01E-07	6.65E-08	7.76E-09	2.76E-09	9.41E-20
20.21	4.17E-07	4.39E-07	5.44E-08	3.83E-09	1.06E-09	1.13E-20
20.41	3.47E-07	1.02E-07	4.32E-08	2.24E-08	7.23E-10	3.85E-22
20.61	3.04E-07	7.92E-08	2.83E-08	8.14E-10	7.77E-10	4.54E-24
20.79	8.24E-08	9.20E-08	2.27E-08	9.31E-09	1.13E-10	1.92E-26

Table 1. Detailed data for figure 1

In this simulation, the coordinate descent methods start from the O-DFD detector and provide up to two orders of magnitude improvement in probability of error compared to the O-DFD detector.

The average CPU times of the algorithms is shown in Figure 2 on a 450 MHz PC based on 1000 Monte-Carlo runs.

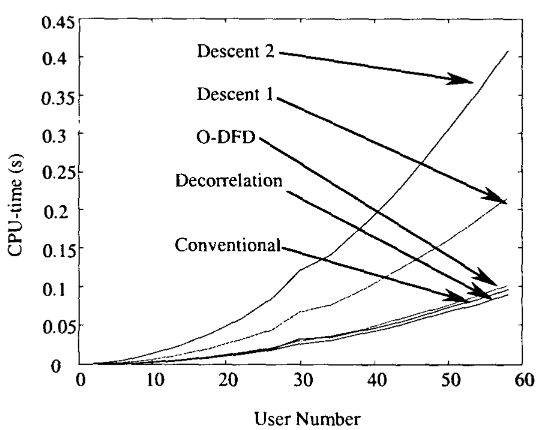

Figure 2. Computational Complexity

The complexity of Descent-1 algorithm (computation of the initial point has been included) is approximately twice that of O-DFD, while the complexity of Descent-2 is almost four times that of linear methods and O-DFD.

5. CONCLUSION

The performance of the coordinate descent algorithms strongly depends on the initialization. Our 1-coordinate descent algorithm (*Descent −1*) starts from the solution obtained by O-DFD, which has similar computational complexity to conventional methods (3) and (4). The 2-coordinate descent algorithm (*Descent-2*) starts from the solution obtained by the 1- coordinate descent algorithm, *Descenet-1*.

Generally, by combining the descent methods with other linear or DFD detectors in an attempt to overcome the projection errors, significant improvements in probability of error are possible, while still maintaining a low computational complexity. However, there is still a big gap in the probability of error between the sub-optimal methods and the ML algorithm.

The ideas presented in this paper can also be used in the asynchronous multiuser detection problems. Further study is needed on how to combine the descent methods efficiently with other search methods (e.g., "tabu" search). Other optimization techniques such as Lagrangian relaxation, quadratic constraint relaxations and semi-definite programming are potential candidates for further improvement of the methods of this paper.

6. REFERENCES

[1] 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.

[2] S. Ulukus and R. Yates, "Optimum detection of code-division multiplexed signals", *IEEE J. Select. Areas Commun.*, vol. 14, pp. 1610-1618, Oct. 1996.

[3] Z. Xie, C. T. Rushforth, and R. T. Short, "Multi-user signal detection using sequential decoding", *IEEE Trans. Commun.*, vol. 38, pp. 578-583, May 1990.

[4] E. Fain and M. K. Varanasi, "Cyclic decision feedback multiuser sequence detection", in *Proc. 32nd Annu. Allerton Conf. Communication, Control, and Computing*, Oct. 1994, pp. 372-381.

[5] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems", *IEEE Trans. Commun.*, vol. 38, pp. 509-519, Apr. 1990.

[6] M. K. Varanasi, "Group detection for synchronous Gaussian code-division multiple-access channels", *IEEE Trans. Inform. Theory*, vol. 41, pp. 1083-1096, July 1995.

[7] M. K. Varanasi, "Decision feedback multiuser detection: a systematic approach", *IEEE Trans. Inform. Theory*, vol. 45, pp. 219-240, Jan. 1999

[8] P.M. Pardalos, G.P. Rodgers, "Computational aspects of a branch and bound algorithm for quadratic zero-one programming", *Computing* 45, 131-144 (1990)

[9] P.M. Pardalos, S. Jha, "Complexity of uniqueness and local search in quadratic 0-1 programming", *Operations research letters*, 11, 119-123 (1992)

[10] G.P. Rodgers, "Algorithms for unconstrained quadratic 0-1 programming and related problems on contemporary computer architectures", *Ph.D. Dissertation*, The Penn. State University (1989)