DETERMINISTIC WAVEFORM-PRESERVING BLIND SEPARATION OF DS-CDMA SIGNALS USING AN ANTENNA ARRAY

Nicholas D. Sidiropoulos[1], Georgios B. Giannakis[1], Rasmus Bro[2]

[1] Dept. of Electrical Eng., University of Virginia, Charlottesville, VA 22903, USA
[2] Chemometrics Group, Food Technology, Dept. of Dairy and Food Science
The Royal Vet. & Agri. Univ., DK-1958 Frederiksberg C, Denmark

ABSTRACT

We consider deterministic blind separation of DS-CDMA signals using an *uncalibrated* antenna array, subject to: (i) unknown flat fading; (ii) unknown multipath; (iii) quasi-synchronous transmissions; (iv) oversaturation; (v) lack of (or, inexact) knowledge of the spreading codes; and (vi) loose power control. We establish a general identifiability result that allows blind recovery of the source signals without requiring independence or whiteness, nor finite alphabet or constant modulus. The results apply to both deterministic and stochastic sources; knowledge of the spreading codes is not required; and, actually, certain matrices involved may be allowed to loose rank without sacrificing identifiability. A provably monotonically convergent trilinear alternating least squares algorithm is proposed to recover the source signals in the noisy case. Aside from their theoretical generality and practical value, our results have important ramifications in eavesdropping of secure spread-spectrum communications, and the judicious choice of user codes.

1. MOTIVATION AND DATA MODELING

Blind separation of signals impinging on an antenna array is of paramount importance in commercial and military applications, including source localization, sensor calibration, blind signal copy, mitigation of co-channel interference, and eavesdropping, just to name a few. Existing self-recovering approaches separate the inaccessible sources from their mixtures by exploiting: (i) known manifolds [1]; (ii) finite-alphabet, constant-modulus, or cyclostationarity, e.g., [6, 2]; (iii) spatial independence of non-Gaussian sources [7, 8]; or (iv) known orthogonal signatures [3, 4].
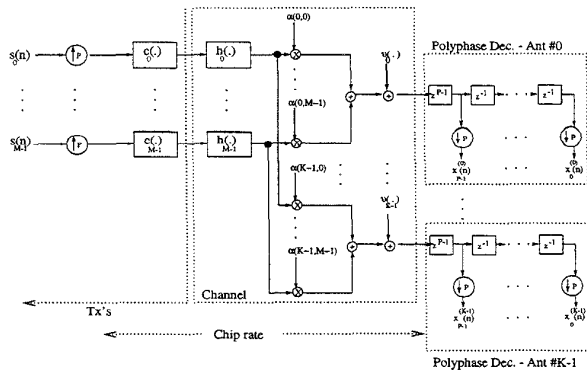

Figure 1: Multiuser/Multirate Discrete-Time Equivalent Baseband CDMA Model (chip rate)

The block diagram in Fig. 1 represents a CDMA system, described in terms of the equivalent discrete-time baseband model, where signals, codes, and channels are represented by samples of their complex envelopes taken at the chip rate. Upsamplers and downsamplers serve the purpose of multiplexing and demultiplexing (spreading and despreading by a factor P). Each of the M users spreads the information sequence $s_m(n)$ with the upsampler and encodes it using the code $c_m(n)$ of length P, before transmission through the *unknown* Lth-order channel $h_m(l)$ which, in addition to multipath, includes the transmit spectral-shaping pulse, the receive-filter, and the mth user's asynchronism in the form of delay factors. The receiver of the system in Fig. 1 employs K antennas. The baseband output of each antenna is sampled at the chip rate, and decomposed into its P polyphase components. We adopt the following assumptions:

(a1) codes have $L_g \geq L$ trailing zeros ('guard chips'), where L is the maximum expected order of all channels $\{h_m(l)\}_{m=0}^{M-1}$ (reasonable for quasi-synchronous systems; see e.g., [5]).

(a2) the multipath/delay channel between user m and antenna k only depends on m; however, each (user,antenna) pair experiences independent fading (e.g., specular multipath).

(a3) the flat fading / antenna response coefficients, $\alpha(k, m)$, remain constant over N symbol epochs.

(a4) the channels $\{h_m(l)\}_{m=0}^{M-1}$ remain constant over N symbol epochs.

For ease of reference, we now summarize the notation used in the subsequent derivation of the data model (c.f. Fig. 1). M: number of users; P: spreading gain; K: number of receiving antenna array elements; N: number of received symbols (snapshots); $s_m(n)$: the transmitted symbol of user m at time n; $\{c_m(p)\}_{p=0}^{P-1}$: the spreading code of user m; $\{h_m(l)\}_{l=0}^{L-1}$: the multipath/quasi-synchronism induced channel of the m-th user; $\left\{\tilde{h}_m(p)\right\}_{p=0}^{P-1}$: the spreading code of user m convolved at the chip rate with $h_m(l)$: $\tilde{h}_m(\cdot) := (c_m * h_m)(\cdot)$, where $*$ denotes convolution, and the result is confined within one symbol duration due to the trailing zeros (see (a1)); $\alpha(k, m)$: compound flat fading/antenna gain associated with the response of the k-th antenna to the m-th user signal.

Define a $P \times M$ aggregate spreading/multipath channel matrix $\mathbf{H}$, with columns $\mathbf{h}_m^T := [\ \tilde{h}_m(0)\ \ \tilde{h}_m(1)\ \ \dots\ \ \tilde{h}_m(P-1)\ ]$, and consider the p-th polyphase component of the output of the k-th antenna element at time n (cf. Fig. 1), denoted by $x_p^{(k)}(n)$. Let us temporarily ignore the noise term, to obtain under (a1)-(a4):

$$x_p^{(k)}(n) = \sum_{m=0}^{M-1} \alpha(k, m)\tilde{h}_m(p)s_m(n) = \mathbf{a}_k^T \mathbf{D}_p(\mathbf{H})\mathbf{s}(n), \quad (1)$$

where $\mathbf{a}_k^T := \begin{bmatrix} \alpha(k,0) & \alpha(k,1) & \ldots & \alpha(k,M-1) \end{bmatrix}$, $\mathbf{D}_p(\mathbf{H})$ $:= diag\left\{ \tilde{h}_0(p), \cdots, \tilde{h}_{M-1}(p) \right\}$, i.e., an $M \times M$ diagonal matrix holding the $(p+1)$-th row of $\mathbf{H}$ in its diagonal, and $\mathbf{s}(n) := [s_0(n) \ s_1(n) \ \ldots \ s_{M-1}(n)]^T$. Let $\mathbf{S}$ be the $M \times N$ source signal matrix obtained by lining up N consecutive instances of $\mathbf{s}(n)$, and consider the row vector formed by lining up N consecutive instances of $x_p^{(k)}(n)$, for $n = 0, 1, \cdots, N-1$. Next, stack the respective row vectors for all K antenna elements. The result is a $K \times N$ matrix $\mathbf{X}_p$, given by $\mathbf{X}_p = \mathbf{A}\mathbf{D}_p(\mathbf{H})\mathbf{S}$ where $\mathbf{A}$ is the $K \times M$ aggregate flat fading-antenna array response matrix with typical element $\alpha(k,m)$, and $\mathbf{X}_p$ is the p-th polyphase matrix component of the observation data in the noiseless case. In the presence of noise, the observation model becomes:

$$\widetilde{\mathbf{X}}_p = \mathbf{X}_p + \mathbf{V}_p = \mathbf{A}\mathbf{D}_p(\mathbf{H})\mathbf{S} + \mathbf{V}_p, \quad p = 0, 1, \cdots, P-1$$

where the $K \times N$ matrix $\mathbf{V}_p$ is the measurement noise corresponding to the p-th polyphase matrix component.

2. THE DIVERSITY DATACUBE

Define a $K \times N \times P$ *three-way array* $\underline{\mathbf{Q}}$, with typical element $q_{k,n,p} := x_p^{(k)}(n)$. $\underline{\mathbf{Q}}$ contains the observation data arranged in (k, n, p)-*diversity space*, reflecting the three different kinds of diversity available: antenna array diversity, temporal diversity, and spreading gain diversity. We shall refer to $\underline{\mathbf{Q}}$ as the *diversity datacube*. The model in (1) affords a scalar view of the diversity datacube as an M-component *trilinear decomposition* of $\underline{\mathbf{Q}}$. Another view, namely:

$$\mathbf{X}_p = \mathbf{A}\mathbf{D}_p(\mathbf{H})\mathbf{S}, \quad p = 0, 1, \cdots, P-1$$

is afforded by the polyphase matrix decomposition of the data. This alternative view can be interpreted as "slicing" the 3-D data into a series of "slabs" (2-D arrays), which are perpendicular to one of the diversity dimensions - in this case, the spreading dimension. Indeed, $\mathbf{X}_p$ is nothing but $[q._{,.,p}]$, the $K \times N$ 2-D slice of $\underline{\mathbf{Q}}$ corresponding to a fixed p. Direct manipulation of the scalar trilinear model yields the following two alternative views of the diversity datacube $\underline{\mathbf{Q}}$:

$$\mathbf{Y}_k = \mathbf{S}^T\mathbf{D}_k(\mathbf{A})\mathbf{H}^T, \quad k = 0, 1, \cdots, K-1$$

where the $N \times P$ matrix $\mathbf{Y}_k := [q_{k,.,.}]$. Recall that $\mathbf{S}^T$ is $N \times M$, $\mathbf{D}_k(\mathbf{A})$ is an $M \times M$ diagonal matrix holding the k-th row of $\mathbf{A}$ in its diagonal, and $\mathbf{H}^T$ is $M \times P$. Similarly,

$$\mathbf{Z}_n = \mathbf{H}\mathbf{D}_n(\mathbf{S}^T)\mathbf{A}^T, \quad n = 0, 1, \cdots, N-1$$

where the $P \times K$ matrix $\mathbf{Z}_n := [q._{,n,.}]$, $\mathbf{H}$ is $P \times M$, $\mathbf{D}_n(\mathbf{S}^T)$ is $M \times M$ diagonal holding the n-th row of $\mathbf{S}^T$, and $\mathbf{A}^T$ is $M \times K$. Note that the multicomponent trilinear model in (1) is *completely symmetric* - thus one may choose the rows of any one of $\mathbf{A}$, $\mathbf{S}^T$, or $\mathbf{H}$, to construct the diagonal matrices that appear in the middle of the decomposition. Alternatively, if two of the three parameters K, P, N are $\geq M$ (e.g., $P \geq M$, and $N \geq M$), one may *always* select a low-rank decomposition, meaning one for which the left factor is "tall" and the right factor is "fat".

3. IDENTIFIABILITY

Theorem 1 *In the noiseless case, the signal matrix $\mathbf{S}$, the compound spreading/multipath/delay channel matrix $\mathbf{H}$, and the compound flat fading/antenna gain matrix $\mathbf{A}$ can be uniquely identified*

(up to the inherently unresolvable permutation and scale ambiguity) from the diversity datacube $\underline{\mathbf{Q}}$, provided that

$$k_{\mathbf{A}} + k_{\mathbf{H}} + k_{\mathbf{S}^T} \geq 2(M+1), \tag{2}$$

where $k_{\mathbf{A}}$ stands for the k-rank of $\mathbf{A}$: the maximum number of linearly independent columns that can be drawn from $\mathbf{A}$ in an arbitrary fashion ($k_{\mathbf{A}} \leq rank(\mathbf{A})$). The result holds for $M = 1$ irrespective of condition (2), so long as $\underline{\mathbf{Q}}$ does not contain an identically zero 2-D slice along any dimension.

In the case of real modulations, the elements of all arrays involved are real numbers, and the result can be proven using Kruskal's results in [9],[10, Theorem 4a]. For complex modulations, the elements of all arrays involved are drawn from the complex field, and, under full-rank conditions, a simpler proof can be constructed using generalized eigenanalysis. Interestingly, it turns out that the base field from which the array elements are drawn makes a difference for such basic concepts as *rank* of multi-way arrays, and complex modulations require a generalization of the line of argument in [9],[10, Theorem 4a].

Under spatially independent fading, $\mathbf{A}$ will be full rank, and thus $k_{\mathbf{A}} = \min(K, M)$ with probability one. Under very mild "persistence of excitation" conditions, $\mathbf{S}$ will be full rank, and hence $k_{\mathbf{S}^T} = \min(N, M)$. If the multipath/delay channel coefficients are modeled as i.i.d. random variables, $k_{\mathbf{H}} = \min(P, M)$ with probability one. Given the above, condition (2) becomes:

$$\min(K, M) + \min(N, M) + \min(P, M) \geq 2M + 2$$

which has the following important corollaries:
(i) Datacube should be truly three-dimensional: $K, N, P \geq 2$;
(ii) If $N \geq M$, $P \geq M$ (typical in DS-CDMA), then $K \geq 2$ antennas are sufficient - which implies that one may separate any number of DS-CDMA multiplexed signals in the presence of unknown ISI/delay/MUI/fading, provided that one uses at least two antennas, the MIMO channel is slowly varying, and appropriate spreading/guard time provisions are met at the transmitter;
(iii) If $K \geq M$, $P \geq M$, then $N \geq 2$ symbols are sufficient - which implies that the fading/multipath/delay can vary as fast as half the symbol rate;
(iv) If $K \geq M$, $N \geq M$, then $P \geq 2$ chips are sufficient. Recall that we have assumed L trailing zeros, which means that $P = 2 + L$ chips per symbol are sufficient (where L is the maximum anticipated order of the multipath/delay channels) -this implies that one does not necessarily need orthogonal or quasi-orthogonal spreading sequences, and it is particularly important in oversaturated systems (more users than spreading).

The importance of Theorem 1 lies precisely in the fact that the sufficient condition involves the concept of k-rank ($\leq$ rank), and identifiability may be guaranteed even when the matrices involved *do loose* rank - it is *the sum of k-ranks that counts*. It is important to note here that, even though the k-rank condition (2) has been shown to be *sufficient*, practical experience of R.A. Harshman (as cited in [10]) indicates that identifiability is not possible when the k-rank condition fails. Thus (2) can be used to characterize the *diversity trade-off*, i.e., the intricate balance of different kinds of diversity (signaling/sampling at the chip rate, use of orthogonal spreading codes, independent fading, multiple receiver antennas) that leads to identifiability.

4. TRILINEAR ALTERNATING LS REGRESSION

The (real) trilinear data model is common in the Chemometrics community [12, 13], where it is used for spectrophotometric, chro-

matographic, and flow injection analyses; it is commonly referred to as the PARAFAC (PARAllel FACtor Analysis) model [11]. Trilinear Alternating Least Squares (TALS) can be used to fit the trilinear model on the basis of noisy observations, possibly initialized using generalized eigenanalysis. This has been proven to be the method of choice in applications of PARAFAC in Chemometrics [12]; its appeal rests primarily in its simplicity, guaranteed convergence, and good performance. Experience shows that, with reasonable initialization, convergence to the global minimum is usually achieved [12]. The explanation (and fundamental difference *vis-a-vis* unstructured bilinear ALS) lies in the inherent *uniqueness* of the trilinear model.

The basic idea behind ALS is simple: each time update *one* of the factors, using least squares *conditioned* on previously obtained estimates of the remaining factors; proceed to update the other factors; repeat until convergence. Effective estimation of the complex trilinear model involves several sub-algorithms and it is of interest in its own right. Algorithmic issues will be presented separately, in a follow-up paper. Here we confine ourselves to (briefly) explaining how one may utilize the three different ways of slicing the datacube to come up with conditional least squares updates for each of the three factors. We are given the noisy datacube $\underline{\widetilde{Q}} = \underline{Q} + \underline{V}$, and wish to estimate $\mathbf{A}$, $\mathbf{H}$, and $\mathbf{S}$. From the first way of slicing the data (polyphase decomposition), it follows that LS fitting amounts to minimizing:

$$\left\| \begin{bmatrix} \widetilde{\mathbf{X}}_0 \\ \vdots \\ \widetilde{\mathbf{X}}_{P-1} \end{bmatrix} - \begin{bmatrix} \mathbf{AD}_0(\mathbf{H}) \\ \vdots \\ \mathbf{AD}_{P-1}(\mathbf{H}) \end{bmatrix} \mathbf{S} \right\|_F^2,$$

where $\widetilde{\mathbf{X}}_p, p = 0, 1, \cdots, P - 1$ are the noisy slabs. It follows that the conditional least squares update for $\mathbf{S}$ is:

$$\widehat{\mathbf{S}} = pinv\left(\begin{bmatrix} \widehat{\mathbf{A}}\mathbf{D}_0(\widehat{\mathbf{H}}) \\ \vdots \\ \widehat{\mathbf{A}}\mathbf{D}_{P-1}(\widehat{\mathbf{H}}) \end{bmatrix} \right) \begin{bmatrix} \widetilde{\mathbf{X}}_0 \\ \vdots \\ \widetilde{\mathbf{X}}_{P-1} \end{bmatrix},$$

where $\widehat{\mathbf{A}}$, $\widehat{\mathbf{H}}$ denote previously obtained estimates of $\mathbf{A}$, and $\mathbf{H}$. Similarly, from the second way of slicing the datacube, the conditional LS update for $\mathbf{H}$ is:

$$\widehat{\mathbf{H}}^T = pinv\left(\begin{bmatrix} \widehat{\mathbf{S}}^T\mathbf{D}_0(\widehat{\mathbf{A}}) \\ \vdots \\ \widehat{\mathbf{S}}^T\mathbf{D}_{K-1}(\widehat{\mathbf{A}}) \end{bmatrix} \right) \begin{bmatrix} \widetilde{\mathbf{Y}}_0 \\ \vdots \\ \widetilde{\mathbf{Y}}_{K-1} \end{bmatrix}.$$

Finally, from the third way of slicing the datacube, it follows that the conditional LS update for $\mathbf{A}$ is:

$$\widehat{\mathbf{A}}^T = pinv\left(\begin{bmatrix} \widehat{\mathbf{H}}\mathbf{D}_0(\widehat{\mathbf{S}}^T) \\ \vdots \\ \widehat{\mathbf{H}}\mathbf{D}_{N-1}(\widehat{\mathbf{S}}^T) \end{bmatrix} \right) \begin{bmatrix} \widetilde{\mathbf{Z}}_0 \\ \vdots \\ \widetilde{\mathbf{Z}}_{N-1} \end{bmatrix}.$$

Note that the conditional update of any given factor, as prescribed above, may either improve, or maintain, but cannot worsen the current fit. Monotone convergence follows directly from this observation.

5. MONTE-CARLO RESULTS

We define the sample SNR (in dB's) at the input of the multiuser receiver as $SNR = 10\log_{10}\frac{||\underline{Q}||_F^2}{||\underline{V}||_F^2}$. Consider $M = 4$ differentially-encoded (DE) BPSK sources, Hadamard(4) codes, $P - M = 2$ trailing zeros, $P = 6$ chips, $L + 1 = 3$-chip Rayleigh multipath/delay channels, $N = 50$ snapshots, and $K = 2$ antennas. For each Monte-Carlo run, the multipath/delay channel coefficients are re-drawn from an i.i.d. Rayleigh generator, and so are the fading coefficients. Note that the channel impulse responses are not normalized to unit norm; this, along with the presence of fading, means that the effective signal power for a given user is usually much different from that of other users for the same run, and varies considerably for the same user from run to run. We compare against the traditional MF receiver (averaging the independently-faded signal copies received by the two antennas), as well as two *non-blind* approaches: the non-blind minimum mean squared error (MMSE) receiver, and the non-blind zero-forcing (ZF) receiver. Both assume full knowledge of spreading codes, multipath channels, relative delays, and fading coefficients. In addition, the non-blind MMSE receiver assumes knowledge of SNR.

The results are presented in Figure 2, which depicts average BER (all users) versus average SNR (both averaged over signal, multipath-delay, fading, and Gaussian noise statistics) results for the proposed method (solid with circles), versus the conventional MF receiver (dashed with stars), the non-blind MMSE method (dash-dot with squares), and the non-blind ZF method (double-dot with triangles), for 1000 Monte-Carlo trials ($O(10^5)$ effective averaging) per datum. Note that the proposed blind receiver achieves results very close to those of the non-blind receivers. Figure 3 presents the corresponding results for DE-QPSK (symbol error rate is reported). Figure 4 presents results for $M = 8$, $N = 50$, DE-BPSK, and everything else as before. Note that the difference between the proposed blind approach and the (completely) non-blind approaches increases as M increases. This is to be expected, for we seek to determine more unknowns from (and a higher-rank model of) the same noisy observation data. However, this can be easily compensated by increasing K (the number of receive antennas) from the present $K = 2$ (minimum possible) to, e.g., $K = 3$.

6. DISCUSSION AND CONCLUSIONS

Blind separation of DS-CDMA signals is of paramount importance in commercial and military spread spectrum wireless communications. The system model considered herein is sufficiently general and flexible to account for a wide array of degradations, including uncalibrated antenna arrays, unknown fading, unknown frequency-selective multipath, quasi-synchronous transmissions, over-saturation, lack of (or, inexact) knowledge of the spreading codes, and loose power control. We have shown that parallel factor analysis provides a unifying framework for blind source separation for DS-CDMA systems utilizing receive antenna arrays, within which the delicate trade-off between different types of diversity required for identifiability can be understood and thoroughly exploited. Multi-way arrays, multi-linear models, and the trilinear model in particular have many additional applications in signal processing and communications, including joint azimuth-elevation and signal-array response estimation, joint angle-delay estimation, and the subspace fitting formulation of multiple-parameter / multiple-invariance ESPRIT. Work on the above problems is reported in companion submissions [14, 15].

7. REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research", *IEEE Signal Processing Magazine*, pp 67–95, July, 1996.

[2] A-J van der Veen and A. Paulraj, "An Analytical Constant-Modulus Algorithm", *IEEE Trans. Signal Processing*, 44:1136–1155, May 1996.

[3] M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. on Info. Theory*, pp. 944-96, July 1995.

[4] M. K. Tsatsanis and Z. Xu, "On Minimum Output Energy CDMA Receivers in the presence of Multipath", *Proc. 31st Conf. on Info. Sciences and Systems*, pp. 724-729, Johns Hopkins Univ., Baltimore MD, March 19-21, 1997.

[5] R. A. Iltis and L. Mailaender, "Multiuser detection of quasi-synchronous CDMA signals using linear decorrelators," *IEEE Trans. on Communications*, pp. 1561-1571, Nov. 1996.

[6] S. Talwar, M. Viberg, and A. Paulraj, "Blind Separation of Synchronous Co-Channel Digital Signals Using an Antenna Array - Part I: Algorithms", *IEEE Trans. Signal Processing*, 44(5):1184–1197, May 1996.

[7] J.F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals", *IEE Proc.*, Part F, 140:362–370, 1993.

[8] L. Tong, Y. Inouye, and R.-W. Liu, "Waveform-preserving blind estimation of multiple independent sources", *IEEE Trans. Signal Processing*, 41:2461–2470, 1993.

[9] J.B. Kruskal, "Three-Way Arrays: Rank and Uniqueness of Trilinear Decompositions, with Application to Arithmetic Complexity and Statistics", *Linear Algebra and Its Applications*, 18:95–138, 1977.

[10] J.B. Kruskal, "Rank Decomposition, and Uniqueness for 3-way and N-way Arrays", In *Multiway Data Analysis*, R. Coppi and S. Bolasco (*Eds.*), pp. 7–18, North-Holland, Amsterdam, 1988 (Proc. MULTIWAY'88).

[11] R.A. Harshman, "Foundations of the PARAFAC procedure: Model and conditions for an 'explanatory' multi-mode factor analysis", *UCLA Working Papers in phonetics*, 16:1–84, (1970).

[12] R. Bro, "PARAFAC: Tutorial & Applications", *Chemometrics and Intelligent Laboratory Systems*, 38:149–171, 1997.

[13] R. Bro, and C.A. Andersson, "Improving the speed of multi-way algorithms. Part II: Compression", *Chemometrics and Intelligent Laboratory Systems*, in press.

[14] R. Bro, N. D. Sidiropoulos, and G. B. Giannakis, "Optimal Joint Azimuth-Elevation and Signal-Array Response Estimation Using Parallel Factor Analysis", *32nd Asilomar Conf. on Signals, Systems, and Computers*, Nov. 1998, Monterey, CA, to appear.

[15] R. Bro, N. D. Sidiropoulos, and G. B. Giannakis, "A Fast Least Squares Algorithm for Separating Trilinear Mixtures", *ICA'99 - Int. Workshop on Independent Component Analysis and Blind Signal Separation*, Jan. 11-15, 1999, Aussois, France, submitted.
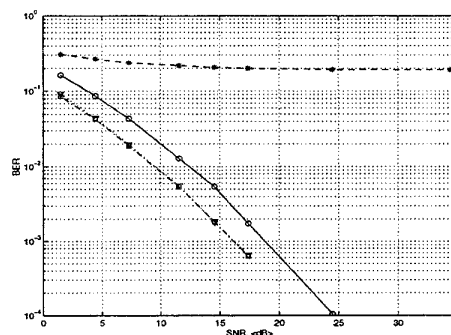
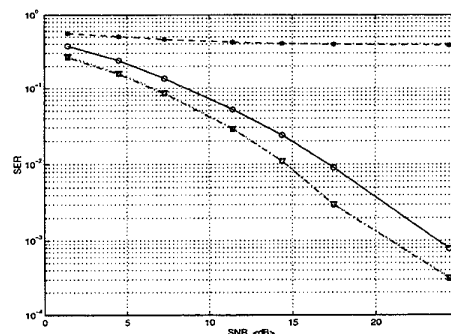Figure 2: BER vs. SNR: $M = 4$ DE-BPSK users (see text).
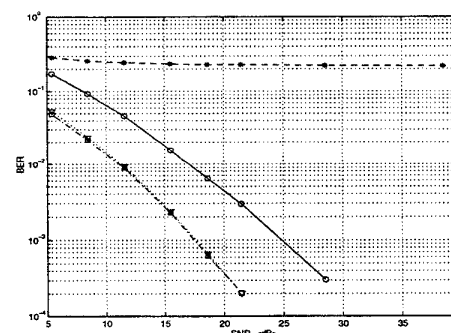

Figure 3: SER vs. SNR: $M = 4$ DE-QPSK users (see text).

Figure 4: BER vs. SNR: $M = 8$ DE-BPSK users (see text).