

JPL Publication 95-21

OK

Improved CDMA Performance Using Parallel Interference Cancellation

Dariusz Divsalar
Marvin Simon

October 1995

JPL
Jet Propulsion Laboratory
California Institute of Technology

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, and was equally funded through the Director's Research and Discretionary Fund and Qualcomm, Inc.

Improved CDMA Performance Using Parallel Interference Cancellation

Dariusz Divsalar
Marvin Simon

Abstract

This report considers a general parallel interference cancellation scheme that significantly reduces the degradation effect of user interference but with a lesser implementation complexity than the maximum-likelihood technique. The scheme operates on the fact that parallel processing simultaneously removes from each user the interference produced by the remaining users accessing the channel in an amount proportional to their reliability. The parallel processing can be done in multiple stages. The proposed scheme uses tentative decision devices with different optimum thresholds at the multiple stages to produce the most reliably received data for generation and cancellation of user interference. The 1-stage interference cancellation is analyzed for three types of tentative decision devices, namely, hard, null zone, and soft decision, and two types of user power distribution, namely, equal and unequal powers. Simulation results are given for a multitude of different situations, in particular, those cases for which the analysis is too complex.



CONTENTS

| | | |
|-------|--|----|
| 1.0 | INTRODUCTION | 1 |
| 2.0 | MULTIUSER COMMUNICATION SYSTEM MODEL | 5 |
| 3.0 | BRUTE FORCE PARALLEL INTERFERENCE CANCELLATION | 6 |
| 3.1 | Tentative Hard Decisions – Equal Power, Synchronous Users | 6 |
| 3.1.1 | Comparison with Conventional CDMA and Successive Interference Cancellation | 12 |
| 3.1.2 | Numerical Results | 13 |
| 3.2 | Tentative Hard Decisions – Unequal Power, Synchronous Users | 14 |
| 3.3 | Null Zone Tentative Decisions – Synchronous Users | 16 |
| 3.4 | Multiple Stage Interference Cancellation | 19 |
| 3.5 | Linear (Infinitely Quantized Soft) Tentative Decisions – Equal Power, Synchronous Users | 20 |
| 4.0 | IMPROVED PARALLEL INTERFERENCE CANCELLATION BASED ON JOINT MAXIMUM-LIKELIHOOD CONSIDERATIONS | 29 |
| 4.1 | Tentative Hard Decisions – Equal Power, Synchronous Users | 30 |
| 4.2 | Tentative Hard Decisions – Unequal Power, Synchronous Users | 31 |
| 4.3 | Null Zone Tentative Decisions – Synchronous Users | 31 |
| 4.4 | Linear (Infinitely Quantized Soft) Tentative Decisions – Equal Power, Synchronous Users | 32 |
| 5.0 | PARALLEL INTERFERENCE CANCELLATION FOR CDMA WITH CODED USERS | 36 |

| | | |
|------------|--|-----|
| 6.0 | CONCLUSIONS | 38 |
| | REFERENCES | 38 |
| APPENDICES | | |
| A. | ANALYTICAL MODELS FOR SIMULATION OF MULTIUSER COMMUNICATION RECEIVERS | A-1 |
| B. | DECISION METRICS FOR MULTIUSER COMMUNICATION RECEIVERS | B-1 |
| C. | THE SPECIAL CASE OF LINEAR INTERFERENCE CANCELLATION | C-1 |
| D. | MEAN AND VARIANCE OF NORMALIZED I&D OUTPUT FOR SINGLE STAGE LINEAR INTERFERENCE CANCELLATION | D-1 |
| E. | VARIOUS MOMENTS OF THE USER CROSSCORRELATIONS | E-1 |
| F. | SOME USEFUL CLOSED FORM DEFINITE INTEGRALS . . . | F-1 |

Figures

| | | |
|------|--|----|
| 1. | An L -Stage Interference Cancellation Scheme with Parallel Processing for CDMA..... | 41 |
| 2a. | k th Stage of Brute Force Interference Cancellation..... | 42 |
| 2b. | k th Stage of Partial Interference Cancellation..... | 43 |
| 3a. | Tentative Decision Devices (a, b, c)..... | 44 |
| 3b. | Tentative Decision Devices (d)..... | 45 |
| 4. | Degradation Factor versus Number of Users for CDMA with and without Interference Cancellation..... | 46 |
| 5. | Comparison of the Degradation Factor of the Lowest Power User versus Number of Users for CDMA with Brute Force IC--One Stage, Hard and Null Zone Tentative Decisions, Unequal Power Users..... | 47 |
| 6. | Comparison of the Degradation Factor versus Number of Users for CDMA with Brute Force IC--One Stage, Hard and Null Zone Tentative Decisions, Equal Power Users..... | 48 |
| 7. | A Comparison of the Degradation Factors for One-, Two-, and Three-Stage Brute Force Interference Cancellation..... | 49 |
| 8a. | A Comparison of the Degradation Factors for One-, Two-, and Three-Stage Brute Force Linear Interference Cancellation (a)..... | 50 |
| 8b. | A Comparison of the Degradation Factors for One-, Two-, and Three-Stage Brute Force Linear Interference Cancellation (b)..... | 51 |
| 9. | Degradation Factor Performance for One-Stage Partial Nonlinear IC as a Function of the Partial Cancellation Parameter..... | 52 |
| 10a. | Degradation Factor Performance for One-Stage Partial Linear IC as a Function of the Partial Cancellation Parameter (a)..... | 53 |
| 10b. | Degradation Factor Performance for One-Stage Partial Linear IC as a Function of the Partial Cancellation Parameter (b)..... | 54 |
| 11. | A Comparison of the Degradation Factors for One- and Two-Stage Linear Interference Cancellation..... | 55 |

Figures

12. A Comparison of the Degradation Factors for One,- Two,- and Three-Stage Nonlinear Interference Cancellation56
13. Degradation Factor versus Number of Users for Convolutionally Coded CDMA with and without Interference Cancellation57

Improved CDMA Performance Using Parallel Interference Cancellation

Dariusz Divsalar
Marvin K. Simon

1.0 Introduction

Multiuser communications systems that employ code division multiple access (CDMA) exhibit a user capacity limit in the sense that there exists a maximum number of users that can simultaneously communicate over the channel for a specified level of performance per user. This limitation is brought about by the ultimate domination of the other user interference over the additive thermal noise. Over the years researchers have sought ways to extend the user capacity of CDMA systems either by employing optimum [maximum-likelihood (ML)] detection or interference cancellation methods [1-14]. In what follows, we briefly summarize the highlights of these two approaches as well as their advantages and disadvantages relative to the conventional approach wherein each user is demodulated assuming that the other users are not sharing the channel. A nice summary of some of these considerations as well as an extensive list of references on the subject can be found in the article by Poor [15].

The conventional receiver for asynchronous multiuser communication referred to above is comprised of a parallel bank of M (the number of users) receivers each of which is optimum for that particular user communicating over the additive white Gaussian noise (AWGN) channel alone.¹ As such, data detection for each user is accomplished independently of that performed on the other users. Although this structure is simple to implement, it is only optimum against the AWGN background but far from optimum against the highly structured nonGaussian interference produced by the simultaneous

¹As an example, if each user is transmitting an uncoded direct sequence BPSK modulation, then each element of the parallel receiver bank would be composed of a coherent carrier demodulator, matched filter (chip matched filter, code despreaders, and accumulator), and hard decision device.

presence of the other $M - 1$ users in the channel. Other disadvantages of the conventional scheme are: (1) it can become multiple access interference limited when the signal-to-background (Gaussian) noise is sufficiently large, and (2) it suffers from the *near-far* problem, i.e., high power users destroying the communication of low power users. The reasoning here is that when the received powers of the users are dissimilar, then in addition to the desired component, the output of each matched filter contains a component which is linear in the amplitude of each of the $M - 1$ interfering users. Thus, high power users contribute a disproportionate amount of interference to low power users making the conventional detector unable to recover their messages.

In view of the above disadvantages, it became natural to ask whether significant performance improvement could be achieved by investigating receivers structures that were optimum against the background Gaussian noise as well as the multiuser interference. Of the multitude of contributions that exist in this area, the work of Verdu [1,2] is perhaps the most cited in the literature and the one upon which much of the other work is based. In Verdu's work, the receiver structure is derived based on minimizing the squared Euclidean distance between the received signal and the *sum* of the M asynchronous user signals, i.e, the total transmitted signal. As such, the presence of all M users simultaneously sharing the channel is accounted for in arriving at the ML receiver. The primary difference between the structure that evolves from such an approach and the conventional structure is that *joint sequence* decisions are made on the set of M matched filter outputs as opposed to individual bit-by-bit decisions on each matched filter output alone.

While indeed such optimum multiuser algorithms offer significantly improved performance by alleviating the disadvantages associated with the conventional scheme, they unfortunately suffer from the fact that their complexity grows exponentially with the number of users and the length of the sequence. This follows directly from the fact that the optimum ML decision algorithm can be implemented as a dynamic program with time complexity per binary decision that is $O(2^M)$ [16]. While in many practical applications such performance complexity prohibits implementation of the Verdu algorithm, its performance is still very much of interest since it serves as a benchmark against which to compare other schemes with less implementation complexity such as those that employ interference

cancellation to be discussed shortly. Another disadvantage of the Verdu algorithm is the necessity of knowing the relative amplitudes of the various user signals present at the input to the receiver. One possibility around this disadvantage is to perform multiuser amplitude estimation [17]. An alternative scheme is to employ *power control* at the receiver which is a common technique used in cellular radio systems to solve the near-far problem. In this case, all received users are assumed to have the same power.

The most obvious solution to the multiuser interference problem would be to design the user codes to have more stringent crosscorrelation properties since indeed if the signals were truly orthogonal this interference would not exist. Unfortunately, however, the near-far problem mentioned above still exists even for well-designed quasiorthogonal signal constellations. Thus, the multiuser interference problem must be dealt with and tackled from another viewpoint.

One popular approach is to employ interference cancellation (IC), i.e., to attempt removal of the multiuser interference from each user's received signal before making data decisions. In principle, the IC schemes considered in the literature fall into two categories, namely, *serial (successive)* and *parallel* cancellation. With regard to the former, Viterbi [6] (see also Dent [7]) suggested coordinated processing of the received signal with a successive cancellation scheme in which the interference caused by the remaining users is removed from each user in succession. One disadvantage of this scheme is the fact that a specific geometric power distribution must be assigned to the users in order that each see the same signal power to background plus interference noise ratio. This comes about because of the fact that with successive cancellation the first user to be processed sees all the interference from the remaining $M-1$ users whereas each user downstream sees less and less interference as the cancellation progresses. Another disadvantage of this scheme has to do with the required delay necessary to fully accomplish the interference cancellation for all users in the system. Since the interference cancellation proceeds serially, a delay on the order of M bit times is required to complete the job. Nevertheless, Viterbi showed that the successive IC scheme could approach channel capacity for the aggregate Gaussian noise channel. As such, the scheme does not become multiuser interference limited.

Parallel processing of multiuser interference *simultaneously* removes from *each* user the total interference produced by the remaining users accessing

the channel. In this way, each user in the system receives equal treatment insofar as the attempt is made to completely cancel his or her multiple user interference. As compared with the serial processing scheme, since the interference cancellation is performed in parallel for all users, the delay required to complete the operation is only a single bit time. The early papers that dealt with parallel interference cancellation recognized the desire to arrive at a structure that could be motivated by the ML approach. In particular, a multistage iterative approach was suggested by Varanasi and Aazhang [8,9] which at a given stage estimated a given user's bit under the assumption that the exact knowledge of the other users' bits in the same transmission interval needed to compute the multiuser interference was replaced by *estimates* of these bits from the previous stage. It was indeed this basic idea which led to the multistage iterative schemes subsequently proposed by Yoon, Kohno, and Imai [11-13] and Kawabe et al [14]. What was common to all of these schemes was the fact that at each stage of the iteration, a brute force attempt was made for each user to completely cancel the interference caused by all the other users. As we shall see in this paper, this is not necessarily the best philosophy. Rather, when the interference estimate is poor (as in the early stages of interference cancellation), it is preferable not to cancel the entire amount of multiuser interference. As the IC operation progresses, the estimates of the multiuser interference improve and thus in the later stages of the iterative scheme, it becomes desirable to cancel more of the interference. The motivation behind this approach can also be derived from maximum-likelihood considerations as was done for the brute force approach previously considered.

With the above discussion in mind, this report presents a general parallel interference cancellation scheme that significantly reduces the degrading effect of multiuser interference but with a lesser implementation complexity than the maximum-likelihood technique of Verdu and with improved performance over the previously considered parallel and serial processing techniques. When compared with classical CDMA without interference cancellation, the improvement in performance is even more dramatic. Our scheme is suitable to the case of a nonuniform power distribution as well as a uniform power distribution among the users. However, we shall focus more on the latter since in parallel interference cancellation schemes, because each user sees the same amount of interference, a uniform power distribution achieves the best overall level of performance. In addition, although our scheme is suitable to asynchronous transmission, we shall assume here that

all users have synchronous data streams. This case results in worst case performance, i.e., if the data transition instants of the various users are not aligned, then on the average they have less of an interfering effect on one another.

2.0 Multiuser Communication System Model

We consider a CDMA communication system in which M users are communicating simultaneously at the same rate over a common AWGN channel each with a BPSK data modulation and their own pseudonoise (PN) code. As such the received signal is the sum of M direct sequence BPSK signals each with power S_i , bit time T_b , and PN chip time T_c , and additive white Gaussian noise with single-sided power spectral density (PSD) N_0 w/Hz. At baseband, this signal can be written in the complex form²

$$r(t) = \sum_{i=1}^M s_i(t) + n(t) = \sum_{i=1}^M \sqrt{S_i} m_i(t) PN_i(t) e^{j\phi_i} + n(t) \quad (1)$$

where for the i^{th} user $PN_i(t)$ is the PN code waveform, $m_i(t) = \sum_{k=-\infty}^{\infty} a_{ik} p(t - kT_b)$ is the data modulation with k^{th} bit a_{ik} taking on equiprobable values ± 1 and unit power rectangular pulse shape $p(t)$ of duration T_b , and ϕ_i is the carrier phase. For the equal user power case, one would have, $S_i = S; i = 1, 2, \dots, M$.

We shall assume for the purpose of analysis and simulation that the users have purely random PN codes³ assigned to them. It is to be emphasized, however, that the IC schemes to be discussed in what follows apply equally well to any appropriate set of PN codes chosen for the users provided that the crosscorrelation matrix of these codes is known. In view of our assumption, over the zeroth bit interval, the i^{th} user's PN waveform can be expressed in the form

$$PN_i(t) = \sum_{k=1}^{\eta} c_{ik} p(t - kT_c) \quad (2)$$

where $p(t)$ is again a unit power rectangular pulse shape now of duration T_c , $\eta = T_b/T_c$ is the number of PN code chips per data bit, i.e., the spreading ratio, and $\{c_{ik}\}$ is a random binary (± 1) sequence. The user codes are thus specified

²For convenience, we shall use complex notation to represent the various signals in the receiver.

³For very long linear feedback shift registers, PN codes can be assumed to be purely random.

in terms of their normalized crosscorrelation matrix $\Gamma = [\gamma_{ij}]$ where

$$\gamma_{ij} = \frac{1}{\eta} \sum_{k=1}^{\eta} c_{ik} c_{jk} = \frac{1}{T_b} \int_0^{T_b} PN_i(t) PN_j(t) dt; \quad i, j = 1, 2, \dots, M \quad (3)$$

with $\gamma_{ii} = 1; \quad i = 1, 2, \dots, M$.

3.0 Brute Force Parallel Interference Cancellation

3.1 Tentative Hard Decisions – Equal Power, Synchronous Users

In Appendix B we discuss several decision metrics for parallel interference cancellation each of which is motivated by ML considerations. All of these metrics suggest IC schemes in the form of iterative structures in which each successive stage simultaneously processes (removes multiuser interference from) all of the users' received signals so as to further refine the quality of their data estimates. The simplest of these schemes is the so-called *brute force* approach discussed in Section II of Appendix B in which an attempt is made at each IC stage to *completely* remove from each user's decision variable the multiuser interference contributed by the remaining users. This scheme is typical of what is found in the literature on parallel IC schemes [11-14].

We begin our discussion by considering the performance of a single stage brute force parallel IC scheme such as that illustrated in Fig. 1 with $L = 1$ together with Fig. 2a. We further assume that the tentative decision devices associated with each user are one bit quantizers (hard decisions) as shown in Fig. 3a. This particular case directly corresponds to the scheme proposed in [11-13]. Although the results derived in Appendix B allow for arbitrary user powers, we assume here that all users have the same power; thus, it is sufficient to characterize only the performance of any one user, say the first, he or she being typical of all the others. Furthermore, as previously stated, we assume that all users have synchronous data streams and purely random PN codes.

The first step is to process the received signal with a matched filter which consists of despreading and demodulating⁴ $r(t)$ with user 1's PN code and carrier reference signal (both of these operations are assumed to be ideal) and then passing the result through a normalized I&D circuit, the output of

⁴Since we are working with a baseband model, the term "remodulation" or "demodulation" refers to complex multiplication by the particular user's carrier phase or its complex conjugate, respectively.

which is given by

$$x_{0i} = a_{i0}\sqrt{E_b} + n_i e^{-j\theta} + \sum_{i=2}^M a_{i0} n_i e^{j(\theta-\theta_i)} \triangleq a_{i0}\sqrt{E_b} + n_i e^{-j\theta} + \sqrt{E_b} \sum_{i=2}^M a_{i0} \gamma_{1i} e^{j(\theta-\theta_i)} \quad (4)$$

where $E_b = ST_b$ denotes the bit energy, a_{i0} is the polarity of user i 's bit in the interval $0 \leq t \leq T_b$, $n_i = \frac{1}{\sqrt{T_b}} \int_0^{T_b} n(t) PN_1(t) dt$ is a zero mean complex Gaussian random variable with variance $E\{|n_i|^2\} = N_0$ representing the thermal noise, and $n_i = \sqrt{\frac{S}{T_b}} \int_0^{T_b} PN_1(t) PN_i(t) dt \triangleq \sqrt{E_b} \gamma_{1i}$; $i = 2, 3, \dots, M$ are the interference noises contributed by the other $M-1$ users which are modeled as independent zero mean Gaussian random variables each with variance ST_c .⁵ Also, the first subscript on x denotes the stage at which we are observing the I&D output while the second subscript denotes the particular user. This notation will be useful later on in our discussion of multiple stage cancellation schemes. The foregoing model of user interference as additive Gaussian noise follows from the assumptions made in similar analyses of CDMA systems [18,19], namely, a large spreading ratio $\eta = T_b/T_c$, and purely random PN codes.

Tentative hard decisions are made on the signals x_{0i} ; $i = 1, 2, \dots, M$ and are used in an attempt to completely cancel the other user interference. If a correct tentative decision is made on a particular other user's bit, then the interference from that user can be completely cancelled. On the other hand, if an incorrect tentative decision is made, then the interference from that user will be enhanced rather than cancelled. A quantitative description of this will be given when we model the signal upon which *final* decisions are made. As we shall see, the performance analysis associated with this model is complicated by the fact that *the tentative decisions are not independent of one another*. More about this shortly.

Defining the Gaussian noise variable n_i and PN crosscorrelation γ_{ij} for the i th user analogous to the above definitions for user 1 (see (A.4) and (A.6) of Appendix A), then after resspreading/remodulation, interference cancellation, and despreading/demodulation, the normalized output of the I&D corresponding to the *final* decisions is given by

⁵The normalized interference noises γ_{1i} ; $i = 2, 3, \dots, M$ have variance equal to the reciprocal of the spreading ratio, i.e., $\eta^{-1} = T_c/T_b$.

$$x_{11} = a_{10}\sqrt{E_b} + n_1 e^{-j\theta} + \overbrace{\sqrt{E_b} \sum_{i=2}^M \beta_i \gamma_{1i} e^{j(\theta_i - \theta)}}^{I_1} \quad (5)$$

where

$$\beta_i = a_{i0} - \operatorname{sgn} \left[\operatorname{Re} \left\{ \sqrt{E_b} \left(a_{i0} + \sum_{\substack{m=1 \\ m \neq i}}^M a_{m0} \gamma_{im} e^{j(\theta_m - \theta)} \right) + n_i e^{-j\theta} \right\} \right] \quad (6)$$

is a three-valued (0, ± 2) indicator random variable whose magnitude represents whether or not a correct tentative decision is made on the i^{th} user's bit. It is tempting to model the β_i 's as independent random variables. Unfortunately, this leads to optimistic results (when compared with the true performance results obtained from simulation). In addition to the fact that the β_i 's are not themselves independent, they are also dependent on the PN crosscorrelations, i.e., the γ_{li} 's. Fortunately, however, the β_i 's are not strongly dependent, i.e., aside from the weak dependence between the Gaussian noises n_i and n_j (see (A.5) of Appendix A), the only signal terms that preclude *complete* independence of say β_i and β_j are $a_{j0}\gamma_{ji}$ in β_i and $a_{i0}\gamma_{ji} = a_{i0}\gamma_{ij}$ in β_j . Hence, for sufficiently large M , it is reasonable to assume a Gaussian model for the total residual (after cancellation) interference term I_1 in (5). The accuracy of this model will improve as M increases (actually as the number of nonzero terms in I_1 increases which implies a high tentative decision error rate). Later on we shall numerically compare the performance results derived from this analytical model with those obtained from a true computer simulation of the receiver to establish the accuracy of this model.

Assuming then a Gaussian model for I_1 (note that I_1 is not zero mean), then the average probability of error associated with the final decisions is given by

$$\begin{aligned} P_b(E) &= \frac{1}{2} \Pr \left\{ \operatorname{Re} \{ x_{11} > 0 \mid a_{10} = -1 \} \right\} + \frac{1}{2} \Pr \left\{ \operatorname{Re} \{ x_{11} < 0 \mid a_{10} = 1 \} \right\} \\ &= \Pr \left\{ \operatorname{Re} \{ x_{11} > 0 \mid a_{10} = -1 \} \right\} = \Pr \left\{ N_e > \sqrt{E_b} - \sqrt{E_b} \sum_{i=2}^M \beta_i \gamma_{1i} \cos(\phi_i - \phi_1) \right\} \end{aligned} \quad (7)$$

where⁶

⁶To simplify the notation here and in what follows, it is understood that the statistical mean $\overline{\beta_i \gamma_{1i} \cos(\phi_i - \phi_1)}$ is computed under the hypothesis $a_{10} = -1$.

$$N_s = \text{Re}\{n_1 e^{-j\phi_1} + I_1 - \bar{I}_1\} = N_1 + \sqrt{E_b} \sum_{i=2}^M \beta_i \gamma_{1i} \cos(\phi_i - \phi_1) - \sqrt{E_b} \sum_{i=2}^M \overline{\beta_i \gamma_{1i} \cos(\phi_i - \phi_1)} \quad (8)$$

is the effective noise seen by user 1 after cancellation which in view of the above is modeled as a real zero mean Gaussian noise random variable whose thermal noise component N_1 has variance $\sigma_{N_1}^2 = N_0/2$. It is straightforward to compute the variance of N_s as

$$\begin{aligned} \sigma_{N_s}^2 &= N_0/2 + E_b(M-1) \overline{\beta_i^2 \gamma_{1i}^2 \cos^2(\phi_i - \phi_1)} - E_b(M-1)^2 \overline{(\beta_i \gamma_{1i} \cos(\phi_i - \phi_1))^2} \\ &\quad + E_b(M-1)(M-2) \overline{\beta_i \gamma_{1i} \beta_j \gamma_{1j} \cos(\phi_i - \phi_1) \cos(\phi_j - \phi_1)} \\ &\quad + 2\sqrt{E_b}(M-1) \overline{N_1 \beta_i \gamma_{1i} \cos(\phi_i - \phi_1)} \end{aligned} \quad (9)$$

where i can take on any value from the set $2, 3, \dots, M$. Hence, from (7), the average probability of error can be obtained as

$$P_b(E) = Q\left(\sqrt{\frac{2E_b}{N_0}} \Lambda\right) \quad (10)$$

where

$$\Lambda \triangleq \frac{(1 + (M-1)\overline{\xi_{1i}})^2}{1 + 2\frac{E_b}{N_0}(M-1)\left[\overline{\xi_{1i}^2} - (M-1)(\overline{\xi_{1i}})^2 + (M-2)\overline{\xi_{1i}\xi_{1j}} + \frac{2}{\sqrt{E_b}}\overline{N_1\xi_{1i}}\right]}; \quad \xi_{1i} \triangleq \beta_i \gamma_{1i} \cos(\phi_i - \phi_1) \quad (11)$$

is an SNR degradation factor (relative to the performance of a single BPSK user transmitting alone) and $Q(x)$ is the Gaussian probability integral defined by

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy \quad (12)$$

Thus, the evaluation of $P_b(E)$ reduces to the evaluation of the various statistical averages involving ξ_{1i} required in (11). These statistical averages, which must be performed over the Gaussian noise and interference random variables as well as the uniformly distributed carrier phases, are not trivial to compute in closed form. The primary difficulty lies in the averages that must be performed over the interference crosscorrelations, γ_{ij} , which from their definition in (3) are binomially distributed in the interval $(-1,1)$. In particular,

for the assumption of random user codes, γ_{ij} is a discrete random variable that takes on values $1 - 2k/\eta$, $k = 0, 1, 2, \dots, \eta$ with probabilities $P_k = \binom{\eta}{k} \left(\frac{1}{2}\right)^\eta$. A further complication lies in the fact that although the γ_{ij} 's are pairwise uncorrelated (see Appendix E), they are not independent. All is not lost however. Since the variance of γ_{ij} is equal to $1/N_c = 1/\eta$, then for large η (the typical case of interest), the p.d.f. of γ_{ij} will be quite narrow with respect to its domain and hence for analysis purposes we can approximate the γ_{ij} 's as being Gaussian distributed. Also in view of this Gaussian assumption on the γ_{ij} 's, we can now further invoke their independence because of their pairwise uncorrelated property. Even with the assumption of independent Gaussian random variables for the γ_{ij} 's, evaluation of the statistical averages required in (11) is still quite complex. Nevertheless, with the help of some useful closed form integrals (see Appendix F), the needed moments can be obtained in the form of definite integrals of tabulated functions with the following results:⁷

$$\overline{\xi_{li}} = -\frac{1}{2\pi} \int_0^{2\pi} \sqrt{\frac{2\sigma^2 \cos^2 \phi}{\pi} \left(\frac{\alpha^2 \sigma^2 \cos^2 \phi}{1 + \alpha^2 \sigma^2 \cos^2 \phi} \right)} \exp\left\{-\frac{\alpha^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)}\right\} d\phi \quad (13a)$$

$$\begin{aligned} \overline{\xi_{li}^2} &= \frac{1}{2\pi} \int_0^{2\pi} \sqrt{\frac{8}{\pi}} \sigma^4 \cos^4 \phi \left(\frac{\alpha^2}{1 + \alpha^2 \sigma^2 \cos^2 \phi} \right)^{3/2} \exp\left\{-\frac{\alpha^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)}\right\} d\phi \\ &+ \frac{1}{2\pi} \int_0^{2\pi} (4\sigma^2 \cos^2 \phi) Q\left(\frac{\alpha}{\sqrt{1 + \alpha^2 \sigma^2 \cos^2 \phi}}\right) d\phi \end{aligned} \quad (13b)$$

⁷Despite the Gaussian assumption on γ , the integration on this random variable in (13c) cannot extend from $-\infty$ to ∞ . The reason for this is that the normalized correlation, ρ , which is directly proportional to γ in accordance with (14) must be restricted to the interval (-1,1). Hence, in the evaluation of (13c) we actually replaced the $-\infty$ and ∞ limits for the integration on γ by the limits -1 and 1 thereby treating it as a truncated Gaussian random variable.

$$\begin{aligned}
\overline{\xi_{1i}\xi_{1j}} &= \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \int_{-\pi}^{\pi} \frac{2\sigma^4 \cos^2 \phi_1 \cos^2 \phi_2 \sqrt{(1-\rho^2)B_1 B_2}}{\sqrt{1-\rho^2 \sigma^4 B_1 B_2 \cos^2 \phi_1 \cos^2 \phi_2}} \\
&\quad \times \exp\left\{-\left(\frac{1-\rho}{2}\right)\left[\frac{B_1 + B_2 + \rho B_1 B_2 \sigma^2 (\cos^2 \phi_1 + \cos^2 \phi_2)}{1-\rho^2 \sigma^4 B_1 B_2 \cos^2 \phi_1 \cos^2 \phi_2}\right](1+\gamma \cos(\phi_1 - \phi_2))\right\} \\
&\quad \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\gamma^2}{2\sigma^2}\right\} d\gamma d\phi_1 d\phi_2
\end{aligned} \tag{13c}$$

$$\frac{2}{\sqrt{E_b}} \overline{N_1 \xi_{1i}} = -\frac{1}{2\pi} \int_0^{2\pi} \frac{\sqrt{2} \alpha \sigma^2 \cos^2 \phi}{\sqrt{\pi} (E_b/N_0)_R} \left[\frac{1+(1+\alpha^2)\alpha^2 \sigma^2 \cos^2 \phi}{(1+\alpha^2 \sigma^2 \cos^2 \phi)^{3/2}} \right] \exp\left\{-\frac{\alpha^2}{2(1+\alpha^2 \sigma^2 \cos^2 \phi)}\right\} d\phi \tag{13d}$$

with

$$\alpha \triangleq \sqrt{\frac{2(E_b/N_0)_R}{1+\left(\frac{M-2}{\eta}\right)(E_b/N_0)_R}}, \quad \alpha' \triangleq \sqrt{\frac{2(E_b/N_0)_R}{1+\left(\frac{M-3}{\eta}\right)(E_b/N_0)_R}}, \quad \sigma^2 = \frac{1}{\eta} = \frac{T_c}{T_b} \tag{14}$$

$$B_i \triangleq \frac{\alpha^2}{1-\rho^2 + \alpha^2 \sigma^2 \cos^2 \phi_i}, \quad i=1,2; \quad \rho \triangleq \frac{\alpha'^2 \gamma \cos(\phi_1 - \phi_2)}{2(E_b/N_0)_R}$$

where $(E_b/N_0)_R$ denotes the *required* bit energy-to-noise spectral density ratio for M users communicating simultaneously, each of which operates at an average bit error rate $P_b(E)$.

It is common in analyses of CDMA systems [18] to define a *degradation factor*, D as the ratio (in dB) of the E_b/N_0 required to achieve a given bit error rate in the presence of M users, namely, $(E_b/N_0)_R$ to that which would be required to achieve the same level of performance if only a single user was communicating, namely, $(E_b/N_0)_1$. By the definition of $(E_b/N_0)_1$, we have

$$P_b(E) = Q\left(\sqrt{2(E_b/N_0)_1}\right) \tag{15}$$

To obtain the degradation factor for a given value of $P_b(E)$, we substitute $D \times (E_b/N_0)_1 = D \times \left[\frac{1}{2} [Q^{-1}(P_b(E))]^2 \right]$ for $(E_b/N_0)_R$ in (14) which in turn is substituted in (13). Then using the given value of $P_b(E)$ one can solve for D . Unfortunately, a closed form expression for D cannot be obtained so the results will be obtained numerically. Before presenting these numerical results, however, we briefly review the analogous results for conventional CDMA and the successive (serial) interference cancellation scheme proposed by Viterbi [6] (later patented by Dent [7]) since we shall use these as a basis of comparison to demonstrate the increased effectiveness of parallel cancellation.

3.1.1 Comparison with Conventional CDMA and Successive Interference Cancellation

In a conventional CDMA system, there is no attempt made to cancel the other user interference. Hence, $(E_b/N_0)_1$ is given by

$$\left(\frac{E_b}{N_0} \right)_1 = \frac{E_b}{N_0 + (M-1)ST_c} = \frac{(E_b/N_0)_R}{1 + (M-1)\eta^{-1}(E_b/N_0)_R} \quad (16)$$

Thus, the degradation factor, D , is [18]

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_1} = \frac{1}{1 - (M-1)\eta^{-1}(E_b/N_0)_1} \quad (17)$$

For the successive cancellation scheme [6], Viterbi showed that to guarantee that each user in the system sees the same amount of interference from the other users, the user powers should be assigned as

$$S_k = S_1 \left(1 + \frac{S_1 T_b}{N_0} \eta^{-1} \right)^{k-1}, \quad k = M, M-1, \dots, 2 \quad (18)$$

where S_1 is the power of the user to be processed last (the weakest one) and S_M is the power of the user to be processed first (the strongest one). Distributing the powers as in (18) ideally guarantees that all users see the same ratio of signal power to effective noise spectral density and thus the user to be processed first (the one that sees all the user interference) is not any SNR disadvantage relative to the user to be processed last (the one for which all interference has been removed). In view of the above, the degradation factor

for the k th user is given by

$$D_k = \frac{(E_b/N_0)_{R_k}}{(E_b/N_0)_1} = \frac{S_k}{S_1} = \left(1 + \eta^{-1}(E_b/N_0)_1\right)^{k-1} \quad (19)$$

where $(E_b/N_0)_{R_k}$ denotes the required bit energy-to-noise spectral density ratio for the k th user. The *average* degradation factor, D for the M user system is obtained by averaging (19) over k which yields

$$D = \frac{1}{M} \sum_{k=1}^M D_k = \frac{\left(1 + \eta^{-1}(E_b/N_0)_1\right)^M - 1}{M\eta^{-1}(E_b/N_0)_1} \quad (20)$$

It should be emphasized that the result in (20) ignores the effect of decision errors made at the various successive interference cancellation stages, that is, the interference cancellation is assumed to take perfectly. As a result, numerical results derived from (20) will be optimistic when compared to the actual performance of the scheme.

3.1.2 Numerical Results

To illustrate the significant performance advantage of the brute force parallel interference cancellation scheme in Fig. 1 combined with Fig. 2a, we consider a plot of D versus M for an average bit error probability⁸ $P_b(E) = 10^{-2}$ and a spreading ratio $\eta = 100$. Fig. 4 shows the analytical performance of the three schemes (conventional, successive interference cancellation, parallel interference cancellation) as well as computer simulation results for the latter. The analytical models that describe the simulation are presented in Appendices A and B. We see that for the conventional and parallel interference cancellation schemes there exists a user capacity limit in that regardless of how much one is willing to increase $(E_b/N_0)_R$ (for a given $(E_b/N_0)_1$, or equivalently, a given $P_b(E)$), the required bit error rate cannot be achieved if more than M_{\max} users simultaneously access the system. For conventional CDMA

$$M_{\max} = 1 + \frac{\eta}{\left(\frac{E_b}{N_0}\right)_1} = 1 + \frac{\eta}{\frac{1}{2}[Q^{-1}(P_b(E))]^2} \quad (21)$$

whereas for the parallel interference cancellation scheme the solution is

⁸The value of $P_b(E) = 10^{-2}$ is chosen to allow for obtaining computer simulation results in a reasonable amount of time.

determined from

$$10^{-2} = Q \left(\frac{1 + (M_{\max} - 1) \overline{\xi}_{1i}}{\sqrt{(M_{\max} - 1) \left[\overline{\xi}_{1i}^2 - (M_{\max} - 1) (\overline{\xi}_{1i})^2 + (M_{\max} - 2) \overline{\xi}_{1i} \overline{\xi}_{1j} \right]}} \right) \quad (22)$$

together with the moments in (13) where now

$$\alpha \triangleq \sqrt{\frac{2}{\eta^{-1}(M_{\max} - 2)}}, \quad \alpha' \triangleq \sqrt{\frac{2}{\eta^{-1}(M_{\max} - 3)}} \quad (23)$$

It is emphasized that the user capacity limit for the parallel interference cancellation scheme comes about entirely because of the finite probability of error associated with the tentative decisions. From Figure 4 it appears that the successive interference cancellation does not have a user capacity limit. This is because in [6], it was assumed for this scheme that the interference cancellation is perfect, i.e., the effect of decision errors at the various interference cancellation stages were not accounted for.

Comparing the analytical and simulation results for the parallel interference cancellation scheme, we observe that the two are virtually in perfect agreement. Any discrepancy between the two stems from the analysis assumptions made in connection with (5) of: a) an analytical Gaussian model for the total residual user interference, and b) independent Gaussian code crosscorrelations (γ_{ij} 's), whereas the computer simulation makes no such assumptions and thus predicts the exact performance. For smaller user error probabilities, e.g., $P_b(E) = 10^{-5}$, one is likely to expect a larger discrepancy between the analytical and simulation results since, in this case, the number of nonzero (residual) interference terms could be small in comparison with the number of users and hence the Gaussian residual interference assumption might well become weak. Under these circumstances, one would be best to rely entirely on the simulation results.

3.2 Tentative Hard Decisions – Unequal Power, Synchronous Users

The results of the previous section can be generalized to the case where the users have unequal powers, i.e., $S_i; i = 1, 2, \dots, M$. Let $\alpha_{ij} = S_i/S_j$ denote the ratio of the power of the i th user to that of the j th user who is arbitrarily considered to be the desired user. After interference cancellation, the normalized output of the I&D corresponding to the *final* decisions of user j

is by analogy with (5)

$$x_{1j} = a_{j0}\sqrt{E_{bj}} + n_j e^{-j\phi_j} + \overbrace{\sum_{\substack{i=1 \\ i \neq j}}^M \sqrt{E_{bi}} \beta_i \gamma_{ji} e^{j(\phi_i - \phi_j)}}^{I_j} = a_{j0}\sqrt{E_{bj}} + n_j e^{-j\phi_j} + \sqrt{E_{bj}} \overbrace{\sum_{\substack{i=1 \\ i \neq j}}^M \sqrt{\alpha_{ij}} \beta_i \gamma_{ji} e^{j(\phi_i - \phi_j)}}^{I_j} \quad (24)$$

where $n_j = \frac{1}{\sqrt{T_b}} \int_0^{T_b} n(t) PN_j(t) dt$, $j = 1, 2, \dots, M$ is a zero mean complex Gaussian random variable with variance $E\{|n_j|^2\} = N_0$ representing the thermal noise of the j th user, $\gamma_{ji} \triangleq \frac{1}{T_b} \int_0^{T_b} PN_j(t) PN_i(t) dt$, $i \neq j$ is the normalized interference produced by user i on user j [see (3)], and $E_{bj} \triangleq S_j T_b$ is the bit energy of the j th user. Also, analogous to (6), β_i is now defined by

$$\beta_i = a_{i0} - \text{sgn} \left[\text{Re} \left\{ \sqrt{E_{bi}} \left(a_{i0} + \sum_{\substack{m=1 \\ m \neq i}}^M \sqrt{\alpha_{mi}} a_{m0} \gamma_{im} e^{j(\phi_m - \phi_i)} \right) + n_i e^{-j\phi_i} \right\} \right] \quad (25)$$

Following steps analogous to (7) – (9) we arrive at the desired result for the bit error probability of the desired (the j th) user, namely,

$$P_{bj}(E) = Q \left(\sqrt{\frac{2E_{bj}}{N_0}} \Lambda_j \right) \quad (26)$$

where

$$\Lambda_j \triangleq \frac{\left(1 + \sum_{\substack{i=1 \\ i \neq j}}^M \sqrt{\alpha_{ij}} \xi_{ji} \right)^2}{1 + 2 \frac{E_{bj}}{N_0} \left[\sum_{\substack{i=1 \\ i \neq j}}^M \alpha_{ij} \xi_{ji}^2 - \left(\sum_{\substack{i=1 \\ i \neq j}}^M \sqrt{\alpha_{ij}} \xi_{ji} \right)^2 + \sum_{\substack{i=1 \\ i \neq j}}^M \sum_{\substack{m=1 \\ m \neq j, i}}^M \sqrt{\alpha_{ij} \alpha_{mj}} \xi_{ji} \xi_{jm} + \frac{2}{\sqrt{E_b}} \sum_{\substack{i=1 \\ i \neq j}}^M \sqrt{\alpha_{ij}} N_j \xi_{ji} \right]}; \quad (27)$$

$$\xi_{ji} \triangleq \beta_i \gamma_{ji} \cos(\phi_i - \phi_j)$$

Clearly for all α_{ij} 's = 1 and $j = 1$, (27) reduces to (11).

As an example, consider a group of M users with powers exponentially distributed (linearly distributed on a dB scale) over a range of 10 dB between the minimum and the maximum. This model might correspond to a distribution of mobile users that are exponentially distant from the base station within a cell. Assume that we fix the error probability of the *lowest power user* (assumed to be user 1 for convenience of notation) equal to 10^{-2} (obviously, all others would then have lower error probability.) Then, Fig. 5 illustrates the degradation factor, D_1 , of user 1 versus M . For comparison, the results corresponding to conventional CDMA with the same user power distribution are also shown in this figure. These conventional CDMA results are obtained from Eq. (9) of [18] which is a generalization of (17) for the

unequal power case, namely, $D_1 = \left[1 - \eta^{-1} (E_b/N_0)_1 \left(\sum_{i=2}^M \alpha_i \right) \right]^{-1}$. By comparison with Fig. 4, we observe that in the unequal power case, parallel interference cancellation offers more of a relative advantage compared to conventional CDMA than in the equal power case. The reason behind this observation is that the larger power of the other users (which are producing the multiuser interference to user 1) produces tentative decisions with a smaller error probability which in turn results in a better degree of cancellation with regard to the final decisions.

3.3 Null Zone Tentative Decisions – Synchronous Users

Much like the idea of including erasures in conventional data detection to eliminate the need for making decisions when the SNR is low, one can employ a null zone hard decision device (see Fig. 3d) for the tentative decisions in Fig. 2a to further improve the fidelity of the interference cancellation process. The idea here is that when a given user's signal to interference ratio is low, it is better not to attempt to cancel the interference from that user than to erroneously detect his data bit and thus enhance his interference. Following the development in Section 3.1 for a single stage scheme with equal power synchronous users, then the normalized output of the I&D corresponding to the final decision on user 1's bit a_{i0} is still given by (5) with β_i now defined by

$$\beta_i = a_{i0} - \text{nsgn} \left[\text{Re} \left\{ \sqrt{E_b} \left(a_{i0} + \sum_{\substack{m=1 \\ m \neq i}}^M a_{m0} \gamma_{im} e^{j(\phi_m - \phi_i)} \right) + n_i e^{-j\phi_i} \right\} \right] \quad (28)$$

where "nsgn" denotes the null zone signum function defined by

$$\text{nsgn } x = \begin{cases} 1, & x > \zeta \\ 0, & -\zeta \leq x \leq \zeta \\ -1, & x < -\zeta \end{cases} \quad (29)$$

Here β_i takes on possible values (0, +1, +2) and its magnitude is an indicator of whether a correct decision is made (ith user's interference is perfectly cancelled), no decision is made (ith user's interference is unaltered), or an incorrect decision is made (ith user's interference is enhanced). Once again making a Gaussian assumption on the total residual interference, then since the final decisions are still made as hard decisions, the average bit error probability is still given by (10) together with (11) with now

$$\begin{aligned} \bar{\xi}_{1i} = & -\frac{1}{2\pi} \int_0^{2\pi} \sqrt{\frac{\sigma^2 \cos^2 \phi}{2\pi} \left(\frac{\alpha^2 \sigma^2 \cos^2 \phi}{1 + \alpha^2 \sigma^2 \cos^2 \phi} \right)} \\ & \times \left[\exp\left\{ -\frac{\alpha^2 (1 + \zeta)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} + \exp\left\{ -\frac{\alpha^2 (1 - \zeta)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} \right] d\phi \end{aligned} \quad (30a)$$

$$\begin{aligned} \bar{\xi}_{1i}^2 = & \frac{1}{2\pi} \int_0^{2\pi} \sqrt{\frac{1}{2\pi}} \sigma^4 \cos^4 \phi \left(\frac{\alpha^2}{1 + \alpha^2 \sigma^2 \cos^2 \phi} \right)^{3/2} \\ & \times \left[3(1 + \zeta) \exp\left\{ -\frac{\alpha^2 (1 + \zeta)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} + (1 - \zeta) \exp\left\{ -\frac{\alpha^2 (1 - \zeta)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} \right] d\phi \\ & + \frac{1}{2\pi} \int_0^{2\pi} (\sigma^2 \cos^2 \phi) \left[3Q\left(\frac{\alpha(1 + \zeta)}{\sqrt{1 + \alpha^2 \sigma^2 \cos^2 \phi}} \right) + Q\left(\frac{\alpha(1 - \zeta)}{\sqrt{1 + \alpha^2 \sigma^2 \cos^2 \phi}} \right) \right] d\phi \end{aligned} \quad (30b)$$

$$\begin{aligned}
\overline{\xi_{1i}\xi_{1j}} &= \frac{1}{4} \left(\frac{1}{2\pi} \right)^2 \int_0^{2\pi} \int_0^{2\pi} \int_{-\pi}^{\pi} \frac{2 \sigma^4 \cos^2 \phi_1 \cos^2 \phi_2 \sqrt{(1-\rho^2)B_1B_2}}{\sqrt{1-\rho^2\sigma^4B_1B_2 \cos^2 \phi_1 \cos^2 \phi_2}} \\
&\times \left[\exp \left\{ - \left(\frac{1-\rho}{2} \right) (1 + \gamma \cos(\phi_1 - \phi_2) - \zeta)^2 F^+(B_1, B_2, \phi_1, \phi_2) \right\} \right. \\
&+ \exp \left\{ - \left(\frac{1-\rho}{2} \right) (1 + \gamma \cos(\phi_1 - \phi_2) + \zeta)^2 F^+(B_1, B_2, \phi_1, \phi_2) \right\} \\
&+ \exp \left\{ - \left(\frac{1-\rho}{2} \right) (1 + \gamma \cos(\phi_1 - \phi_2))^2 F^+(B_1, B_2, \phi_1, \phi_2) \right\} \exp \left\{ - \left(\frac{1+\rho}{2} \right) \zeta^2 F^-(B_1, B_2, \phi_1, \phi_2) \right\} \\
&\times 2 \cosh \left\{ \zeta (1-\rho^2) (1 + \gamma \cos(\phi_1 - \phi_2)) G(B_1, B_2, \phi_1, \phi_2) \right\} \\
&\times \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ - \frac{\gamma^2}{2\sigma^2} \right\} d\gamma d\phi_1 d\phi_2
\end{aligned} \tag{30c}$$

$$\begin{aligned}
\frac{2}{\sqrt{E_b}} \overline{N_1 \xi_{1i}} &= - \frac{1}{4\pi} \int_0^{2\pi} \sqrt{\frac{2}{\pi}} \frac{\alpha \sigma^2 \cos^2 \phi}{(E_b/N_0)_R} \left[\frac{1 + (1 + \alpha^2(1 + \zeta^2)) \alpha^2 \sigma^2 \cos^2 \phi}{(1 + \alpha^2 \sigma^2 \cos^2 \phi)^{3/2}} \right] \exp \left\{ - \frac{\alpha^2(1 + \zeta^2)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} d\phi \\
&- \frac{1}{4\pi} \int_0^{2\pi} \sqrt{\frac{2}{\pi}} \frac{\alpha \sigma^2 \cos^2 \phi}{(E_b/N_0)_R} \left[\frac{1 + (1 + \alpha^2(1 - \zeta^2)) \alpha^2 \sigma^2 \cos^2 \phi}{(1 + \alpha^2 \sigma^2 \cos^2 \phi)^{3/2}} \right] \exp \left\{ - \frac{\alpha^2(1 - \zeta^2)^2}{2(1 + \alpha^2 \sigma^2 \cos^2 \phi)} \right\} d\phi
\end{aligned} \tag{30d}$$

where

$$\begin{aligned}
F^\pm(B_1, B_2, \phi_1, \phi_2) &\triangleq \frac{B_1 + B_2 \pm \rho B_1 B_2 \sigma^2 (\cos^2 \phi_1 + \cos^2 \phi_2)}{1 - \rho^2 \sigma^4 B_1 B_2 \cos^2 \phi_1 \cos^2 \phi_2} \\
G(B_1, B_2, \phi_1, \phi_2) &\triangleq \frac{B_1 - B_2}{1 - \rho^2 \sigma^4 B_1 B_2 \cos^2 \phi_1 \cos^2 \phi_2}
\end{aligned} \tag{30e}$$

with B_1, B_2 defined as before in (14) and $\zeta = \zeta / \sqrt{E_b}$ is the normalized decision threshold which should be chosen to minimize D for a given $P_b(E)$ and $(E_b/N_0)_1$ determined from (15).

Numerical results for the case where a null zone limiter is used as the tentative decision device are illustrated in Fig. 6 together with the corresponding results previously given in Fig. 4 for the case of hard limited tentative decisions. For the specified processing gain and average bit error probability (the same values as in Fig. 4), we see that using a null zone limiter

allows the maximum number of users that can be supported to be increased by about 10% – a modest improvement in performance. The normalized threshold has been fixed at $\zeta' = 0.4$ which appears to be the value that gives the best performance improvement. For an unequal (exponentially distributed) power distribution among the users, the corresponding results using null zone tentative decisions are superimposed on the previous discussed results illustrated in Fig. 5. Here, the normalized threshold⁹ has been fixed for each user at $\zeta' = 0.25$ which appears to give the best overall performance improvement for the specified user power distribution. Again we see a modest improvement in performance.

3.4 Multiple Stage Interference Cancellation

The single stage scheme of Fig. 1 combined with Fig. 2a can be improved upon by cascading multiple stages ($L > 1$ in Fig. 1) of parallel interference cancellation.¹⁰ The idea here is to repeatedly improve the fidelity of the M tentative decisions since each successive stage sees less and less interference. Note that in principle this idea is similar to what Viterbi accomplishes in the serial interference cancellation scheme except that here at each stage we simultaneously act on the interference from all M users rather than one user at a time. An analysis of the performance of such a multistage scheme is difficult if not impossible to obtain due to the fact that the tentative decisions at the i^{th} interference cancellation stage depend on the tentative decisions at the $(i-1)^{\text{st}}$ stage. Because of this difficulty, numerical results for the performance of the multi-stage parallel interference scheme must be obtained from computer simulation. Illustrated in Fig. 7 are performance results for two and three stage parallel IC schemes with hard tentative decisions, respectively. The corresponding one stage results are repeated on these figures for the purpose of comparison. We observe that there is significant gain to be achieved by going to more than one stage.

⁹In the unequal user power case, the normalized threshold for the i th user is defined by $\zeta' = \zeta_i / \sqrt{E_{bi}}$ where ζ_i is the actual decision threshold for this user. Thus, a fixed normalized threshold implies a different actual threshold for each user in accordance with his or her energy (power).

¹⁰ For the case where the tentative decisions associated with each user are one bit quantizers (hard decisions) the scheme again reduces to that proposed in [11-13].

3.5 Linear (Infinitely Quantized Soft) Tentative Decisions – Equal Power, Synchronous Users

One disadvantage of the parallel cancellation scheme with hard or null zone limiter tentative decisions is that, in order to perform the resreading and remodulating operations, the receiver must ideally have complete knowledge of each user's power, carrier phase and frequency, and PN code chip timing epoch. Since in practice the receiver does not have knowledge of these parameters, it must estimate them. One simple way of circumventing some of these problems is to use *linear* (infinitely soft quantization) tentative decisions (see Fig. 2c). Since the signal component of the output of the tentative decision devices is now linearly proportional to the square root of the user powers, it is no longer necessary for the receiver to estimate these powers prior to the cancellation operation and thus the $\sqrt{E_b}$ gains following these devices in Fig. 2a may be eliminated. Another simplification that is now possible is that the final decisions can be performed with a differential (rather than coherent) detector thus eliminating the need for carrier synchronization at all stages. There is of course a performance penalty associated with using differential rather than coherent detection; however, it was shown in [20] that by using a multiple symbol (as opposed to the conventional two-symbol) differential detector one can fill the gap between these two performances. The primary disadvantage to using linear tentative decisions is that additive thermal noise from each user is now directly introduced into the interference cancellation process. This will result in a performance that is inferior to the hard and null zone tentative decision schemes but still better than conventional CDMA which employs no interference cancellation at all. Furthermore, one can now in principle *analytically* compute the performance of a multiple stage parallel interference scheme although the analysis becomes quickly complex as the number of stages increases beyond two or three.

We first consider the analysis of a single stage scheme using linear tentative decision devices and with equal power, synchronous users. Analogous to (4), the normalized output of the j th I&D circuit in Fig. 2a is given by

$$x_{0j} = a_{j0}\sqrt{E_b} + n_j e^{-j\phi_j} + \sum_{\substack{i=1 \\ i \neq j}}^M a_{i0} n_{ji} e^{j(\phi_i - \phi_j)} = a_{j0}\sqrt{E_b} + n_j e^{-j\phi_j} + \sqrt{E_b} \sum_{\substack{i=1 \\ i \neq j}}^M a_{i0} \gamma_{ji} e^{j(\phi_i - \phi_j)} \quad (31)$$

After resreading/remodulation, interference cancellation, and despreading/demodulation, the normalized output of the I&D corresponding to the *final* decisions of user 1 (again assumed without loss of generality to be the desired user) is given by

$$x_{11} = a_{10} \left[1 - \sum_{i=2}^M \gamma_{1i}^2 \right] \sqrt{E_b} - \sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\theta_i - \theta_1)} \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{ik} \gamma_{1k} + \left(n_1 - \sum_{i=2}^M \gamma_{1i} n_i \right) e^{-j\theta_1} \quad (32)$$

We note that the decision statistic for user 1 now includes an attenuation of the desired signal component as well as additional thermal noise due to the presence of the other $M-1$ users. The middle term in (32) represents the residual user interference after cancellation.

Exact evaluation of average error probability would be obtained by computing the bit error probability conditioned on a given set of user codes (random sequences) which takes the form of a Gaussian Q-function and then averaging the result over all user codes. An approximation to this is obtained by treating the interference as an additional Gaussian noise source thus allowing a solution for the average error probability in the form of a single Gaussian Q-function. This approach is, in principle, the same as that previously used for the nonlinear IC case except that here we can *exactly* evaluate the first and second moments of the interference using the true binomial distribution of the γ_{ij} 's. In particular, the mean (conditioned on $a_{10} = 1$) and variance of x_{11} are computed in Appendix D. Using these results the average bit error probability for hard decisions made on $x_{R,1} \triangleq \text{Re}\{x_{11}\}$ is

$$P_b(E) = Q \left(\sqrt{\frac{\left(E\{x_{R,1} | a_{10}=1\} \right)^2}{\sigma_{11}^2}} \right) \triangleq Q \left(\sqrt{2 \left(\frac{E_b}{N_0} \right)_1} \right) \quad (33)$$

where σ_{11}^2 denotes the variance of $x_{R,1}$ and

$$\left(\frac{E_b}{N_0} \right)_1 = \frac{\left(\frac{E_b}{N_0} \right)_R \left(1 - \frac{M-1}{\eta} \right)^2}{\left(1 - \frac{M-1}{\eta} + \frac{(M-1)(M-2)}{\eta^2} \right) + \left(\frac{E_b}{N_0} \right)_R \frac{(M-1)}{\eta^2} \left[M+2 + \frac{M^2 - 5M + 2}{\eta} \right]} \quad (34)$$

or, equivalently, the degradation factor is given by

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_I} = \frac{1 - \frac{M-1}{\eta} + \frac{(M-1)(M-2)}{\eta^2}}{\left(1 - \frac{M-1}{\eta}\right)^2 - \left(\frac{E_b}{N_0}\right)_I \left(\frac{M-1}{\eta^2}\right) \left[M+2 + \frac{M^2 - 5M + 2}{\eta}\right]} \quad (35)$$

Analogous to (21), the maximum number of users, M_{\max} , that can be supported with a single stage linear cancellation scheme is obtained by setting $D = \infty$ or equivalently setting the denominator of (35) equal to zero. An explicit solution for M_{\max} is cumbersome. For large M , one has the asymptotic (as $(E_b/N_0)_I$ becomes sufficiently large) solution

$$M_{\max} \cong \frac{\eta}{\sqrt{\left(\frac{E_b}{N_0}\right)_I}} = \frac{\eta}{\sqrt{2} Q^{-1}(P_b(E))} \quad (36)$$

Thus, ignoring the "1+" term on the right hand side of (21), (valid for large M_{\max}), we see that the maximum number of users that can be supported by a CDMA system using a single stage of linear interference cancellation is increased relative to that of a conventional CDMA system with no interference cancellation by the factor $\sqrt{(E_b/N_0)_I}$ in the limit of large $(E_b/N_0)_I$. Figure 8a is an illustration of D versus M obtained from (35) for $\eta = 100$ and $P_b(E) = 10^{-2}$, i.e., the same processing gain and average bit error probability as in Fig. 4. Superimposed on this figure are the results corresponding to conventional CDMA obtained from Eq. (17). We see that for low values of D , improvement is obtained with the brute force linear cancellation method relative to conventional CDMA whereas asymptotically for large D (i.e., where multiuser interference dominates), we indeed obtain a slight loss in performance. Clearly then $(E_b/N_0)_I$ corresponding to $P_b(E) = 10^{-2}$ is not sufficiently large to justify the use of (36). Furthermore, the brute force linear cancellation method is quite inferior to the brute force hard-limited tentative decision case (see Fig. 4). In Figure 8b are the analogous results for $\eta = 100$ and a smaller error probability, namely, $P_b(E) = 10^{-5}$. Superimposed on this figure are again the results corresponding to conventional CDMA. We now see that an improvement relative to conventional CDMA exists for both small and large values of D . Still, however, the linear cancellation method is far inferior to the nonlinear cancellation method using hard tentative decisions.

Before proceeding to a discussion of multiple stage linear IC schemes, we wish to compare the "exact"¹¹ degradation factor found in (35) with that which would be obtained by invoking an independent Gaussian assumption on the γ_{ij} 's as discussed in the latter part of Appendix D. Using (D.3') – (D.7') in place of (D.3) – (D.7), the degradation factor becomes¹²

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_1} = \frac{1 - \frac{M-1}{\eta}}{\left(1 - \frac{M-1}{\eta}\right)^2 - \left(\frac{E_b}{N_0}\right)_1 \left[\frac{(M-1)(M+2)}{\eta^2}\right]} \quad (35')$$

Comparing (35') with (35) we see that the results are similar except for higher order terms in η^{-1} . To quantify this comparison, the degradation factor, D , as computed from (35') is plotted in Fig. 8 along side the "exact" result computed from (35). Also shown here are the simulation results which correspond to the true exact performance. We observe that the degradation factor computation based on the independent Gaussian assumption for the γ_{ij} 's is reasonably close (but somewhat optimistic) to the "exact" result based on binomially distributed γ_{ij} 's that are not necessarily independent. The reasonable similarity of the "exact" and approximate degradation factor results for the single stage linear IC scheme will serve as a justification for using the latter in making comparisons among the performances of multiple stage linear IC schemes where exact computation of D is extremely tedious.

As discussed in Section 3.4 matters can be improved by going to a multiple stage interference cancellation configuration. For two stages of cancellation with linear tentative decisions as illustrated in Fig. 1 combined with Fig. 2a, it can be shown that the normalized output of the I&D corresponding to the *final* decisions of user 1 is (by analogy with (32))

¹¹In referring to the degradation factor in (35) as "exact" we mean that the true binomial (as opposed to a Gaussian) distribution of the γ_{ij} 's is used in computing the moments of these random variables. We hasten to point out, however, that the degradation factor in (35) is still not truly exact (i.e., it will not totally agree with the simulation results) because, as mentioned in the text, the total residual interference is still being analytically modeled as a Gaussian random variable [see the discussion preceding Eq. (33).]

¹²Note that (35') results in the same asymptotic M_{\max} as (35) does, namely, (36).

$$\begin{aligned}
x_{21} = & a_{10} \left[1 + \sum_{l=2}^M \sum_{\substack{k=2 \\ k \neq l}}^M \gamma_{lk} \gamma_{1k} \gamma_{1l} \right] \sqrt{E_b} + \sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\phi_i - \phi_1)} \sum_{\substack{l=2 \\ k=l, i}}^M \sum_{k=1}^M \gamma_{lk} \gamma_{ik} \gamma_{1l} \\
& + \left(n_1 - \sum_{l=2}^M \gamma_{1l} n_l + \sum_{\substack{l=2 \\ k=1 \\ k \neq l}}^M \sum_{k=1}^M \gamma_{lk} \gamma_{1l} n_k \right) e^{-j\phi_1}
\end{aligned} \tag{37}$$

The first term again represents the desired signal component, the middle term the other user interference and the third term the additive thermal noise. Using Eq. (E.3) of Appendix E, it is straightforward to show that the mean of the first term (conditioned on $a_{10}=1$) is equal to $\sqrt{E_b} [1 + (M-1)(M-2)/\eta^2]$ while the other two terms in (37) have zero mean. Thus, for the two stage linear cancellation scheme $E\{x_{21}|a_{10}=1\} = \sqrt{E_b} [1 + (M-1)(M-2)/\eta^2]$ which represents a signal enhancement caused by the presence of the other users. The variance of each term can, in principle, be computed analogous to Eqs. (D.3), (D.4), and (D.5) of Appendix D but with considerably more difficulty which makes this approach impractical. Thus, we shall proceed as in the latter part of Appendix D by invoking an independent Gaussian assumption for the γ_{ij} 's. When this is done, the following results are obtained.

The mean of the desired signal term (conditioned on $a_{10}=1$) is now equal to $\sqrt{E_b}$. Similarly, the means of the other two terms in (37) are still equal to zero. The variance of the self noise of the desired signal component, user interference, and equivalent thermal noise¹³ are given by

$$E \left\{ \left(\sqrt{E_b} a_{10} \sum_{l=2}^M \sum_{\substack{k=2 \\ k \neq l}}^M \gamma_{lk} \gamma_{1k} \gamma_{1l} \right)^2 \right\} = E_b \frac{2(M-1)(M-2)}{\eta^3} \tag{38}$$

$$\frac{1}{2} E \left\{ \left(\sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\phi_i - \phi_1)} \sum_{\substack{l=2 \\ k=l, i}}^M \sum_{k=1}^M \gamma_{lk} \gamma_{ik} \gamma_{1l} \right)^2 \right\} = \frac{E_b}{2} \frac{5(M-1)(M^2 - M + 1)}{\eta^3} \tag{39}$$

¹³Recall that the noise components $\{n_i\}$ are conditionally correlated as per Eq. (A.5) of Appendix A.

$$\frac{1}{2} E \left\{ \left[n_1 - \sum_{l=2}^M \gamma_{1l} n_l + \sum_{l=2}^M \sum_{k=1}^M \gamma_{lk} \gamma_{1l} n_k \right]^2 \right\} = \frac{N_0}{2} \left[1 + \left(\frac{M-1}{\eta} \right) \left(1 - \frac{2M-1}{\eta} \right) \right] \quad (40)$$

Finally, combining (38)–(40) we have analogous to (34)

$$\left(\frac{E_b}{N_0} \right)_1 = \frac{\left(\frac{E_b}{N_0} \right)_R}{1 + \left(\frac{M-1}{\eta} \right) \left(1 - \frac{2M-1}{\eta} \right) + \left(\frac{E_b}{N_0} \right)_R \frac{(M-1)(5M^2 - M - 3)}{\eta^3}} \quad (41)$$

or solving for the degradation factor

$$D = \frac{\left(\frac{E_b}{N_0} \right)_R}{\left(\frac{E_b}{N_0} \right)_1} = \frac{1 + \left(\frac{M-1}{\eta} \right) \left(1 - \frac{2M-1}{\eta} \right)}{1 - \left(\frac{E_b}{N_0} \right)_R \frac{(M-1)(5M^2 - M - 3)}{\eta^3}} \quad (42)$$

Analogous to (36), the maximum number of users that can be supported with a two stage linear cancellation scheme is obtained by setting $D = \infty$ in (42), which for large M_{\max} can be approximated by

$$M_{\max} \cong \frac{\eta}{\sqrt[3]{5 \left(\frac{E_b}{N_0} \right)_1}} \quad (43)$$

Thus, again ignoring the "1+" term on the right hand side of (21), the maximum number of users that can be supported by a CDMA system using a two stage linear interference cancellation is increased relative to that of a conventional CDMA system with no interference cancellation by the factor $\sqrt[3]{\left[\left(\frac{E_b}{N_0} \right)_1 \right]^2 / 5}$. An illustration of D versus M obtained from (42) is superimposed on the results of Figures 8a and b. We observe that the addition of a second stage of cancellation gives us a modest improvement in performance.

Obtaining performance results for more than two stages of linear interference cancellation can in principle be accomplished but requires formidable analytical evaluations. The procedure for K stages and the difficulty in obtaining closed form results are described as follows. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iM})^T$, $i = 0, 1, 2, \dots, K$ denote the column vector of M I&D outputs

at the i th stage of the interference canceller. Furthermore, let $\mathbf{A} = \mathbf{\Gamma} - \mathbf{I}$ where $\mathbf{\Gamma} = \{\gamma_{ij}\}$ is the $M \times M$ matrix of normalized PN crosscorrelations and \mathbf{I} is the $M \times M$ identity matrix. Since $\gamma_{ii} = 1$ for all $i = 1, 2, \dots, M$, and $\gamma_{ij} = \gamma_{ji}$, then \mathbf{A} is a symmetric matrix with zeros along its diagonal. In terms of these quantities, the output of the i th stage is given by the iterative relation

$$\mathbf{x}_i = \mathbf{x}_0 - \mathbf{\Phi}^* \mathbf{A} \mathbf{\Phi} \mathbf{x}_{i-1} \quad (44)$$

where

$$\mathbf{x}_0 = \sqrt{E_b} \mathbf{a} + \sqrt{E_b} \mathbf{\Phi}^* \mathbf{A} \mathbf{\Phi} \mathbf{a} + \mathbf{\Phi}^* \mathbf{n} \quad (45)$$

$\mathbf{a} = (a_{10}, a_{20}, \dots, a_{M0})^T$ and $\mathbf{n} = (n_1, n_2, \dots, n_M)^T$ respectively denote the data and noise column vectors, and $\mathbf{\Phi}$ is the complex carrier phase diagonal matrix defined in Eq. (A.12) of Appendix A. The solution to (44) for the output of the K th stage can be easily shown to be

$$\mathbf{x}_K = \sqrt{E_b} \mathbf{a} - \sqrt{E_b} \mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi} \mathbf{a} + \mathbf{\Phi}^* \sum_{i=0}^K (-\mathbf{A})^i \mathbf{n} \quad (46)$$

Defining the column vector $\mathbf{u} = (1, 0, 0, \dots, 0)^T$, then the output of the K th stage for user 1 is

$$x_{K1} = \mathbf{u}^T \mathbf{x}_K = \sqrt{E_b} a_{10} - \sqrt{E_b} \mathbf{u}^T \mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi} \mathbf{a} + \mathbf{u}^T \mathbf{\Phi}^* \sum_{i=0}^K (-\mathbf{A})^i \mathbf{n} \quad (47)$$

Letting $\mathbf{a} = a_{10} \mathbf{u} + \mathbf{a}'$ where $\mathbf{a}' = (0, a_{20}, a_{30}, \dots, a_{M0})^T$, then (47) can be rewritten as

$$x_{K1} = \mathbf{u}^T \mathbf{x}_K = \sqrt{E_b} a_{10} (1 - \mathbf{u}^T \mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi} \mathbf{u}) - \sqrt{E_b} \mathbf{u}^T \mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi} \mathbf{a}' + \mathbf{u}^T \mathbf{\Phi}^* \sum_{i=0}^K (-\mathbf{A})^i \mathbf{n} \quad (48)$$

Note that for an square matrix $\mathbf{C} = \{c_{ij}\}$, we have $\mathbf{u}^T \mathbf{C} \mathbf{u} = c_{11}$. Thus, since the diagonal elements of $\mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi}$ are identical to the diagonal elements of $(-\mathbf{A})^{K+1}$, we can simplify (48) to

$$x_{K1} = \mathbf{u}^T \mathbf{x}_K = \sqrt{E_b} a_{10} (1 - \mathbf{u}^T (-\mathbf{A})^{K+1} \mathbf{u}) - \sqrt{E_b} \mathbf{u}^T \mathbf{\Phi}^* (-\mathbf{A})^{K+1} \mathbf{\Phi} \mathbf{a}' + \mathbf{u}^T \mathbf{\Phi}^* \sum_{i=0}^K (-\mathbf{A})^i \mathbf{n} \quad (49)$$

For $K = 1$ and $K = 2$, it is straightforward to check that (49) reduces to (32) and (37) respectively.

The first term in (49) is real whereas the second and third terms are

complex for all values of K . To compute the SNR of user 1 after K stages of interference cancellation we need the mean and variance (conditioned on $a_{10} = 1$) of the real part of the decision variable in (49). These are given by

$$E\{\text{Re}\{x_{K1}\}|a_{10} = 1\} = \sqrt{E_b} - \sqrt{E_b} \mathbf{u}^T \overline{(-\mathbf{A})^{K+1}} \mathbf{u} = \sqrt{E_b} \left[1 - \mathbf{u}^T \overline{(-\mathbf{A})^{K+1}} \mathbf{u} \right] \quad (50)$$

and

$$\begin{aligned} \text{var}\{\text{Re}\{x_{K1}\}|a_{10} = 1\} &= E_b \left\{ \overline{(\mathbf{u}^T (-\mathbf{A})^{(K+1)} \mathbf{u})^2} - \left[\overline{\mathbf{u}^T (-\mathbf{A})^{K+1}} \mathbf{u} \right]^2 \right\} \\ &\quad + \frac{E_b}{2} E \left\{ \left| \sqrt{E_b} \mathbf{u}^T \Phi^* (-\mathbf{A})^{K+1} \Phi \mathbf{a} \right|^2 \right\} + \frac{1}{2} E \left\{ \left| \mathbf{u}^T \Phi^* \sum_{i=0}^K (-\mathbf{A})^i \mathbf{n} \right|^2 \right\} \\ &= \frac{E_b}{2} \left\{ \overline{\mathbf{u}^T (-\mathbf{A})^{2(K+1)} \mathbf{u}} + \overline{(\mathbf{u}^T (-\mathbf{A})^{(K+1)} \mathbf{u})^2} - 2 \left[\overline{\mathbf{u}^T (-\mathbf{A})^{K+1}} \mathbf{u} \right]^2 \right\} \\ &\quad + \frac{N_0}{2} \mathbf{u}^T \sum_{i=0}^K \sum_{j=0}^K \left(\overline{(-\mathbf{A})^{i+j}} - \overline{(-\mathbf{A})^{i+j+1}} \right) \mathbf{u} \end{aligned} \quad (51)$$

where for convenience of notation the overbar denotes statistical expectation. Also note that

$$\sum_{i=0}^K \sum_{j=0}^K \overline{(-\mathbf{A})^{i+j}} = \sum_{\substack{i=0 \\ i \text{ even}}}^K (i+j) \overline{(-\mathbf{A})^i} + \sum_{\substack{i=K+1 \\ i \text{ even}}}^{2K} (2K-i+1) \overline{(-\mathbf{A})^i} \quad (52)$$

Squaring (48), dividing by (49) and equating the result to $2(E_b/N_0)_1$ gives a relation from which the degradation factor D can be solved, namely,

$$D = \frac{\mathbf{u}^T \sum_{i=0}^K \sum_{j=0}^K \left(\overline{(-\mathbf{A})^{i+j}} - \overline{(-\mathbf{A})^{i+j+1}} \right) \mathbf{u}}{\left(1 - \mathbf{u}^T \overline{(-\mathbf{A})^{K+1}} \mathbf{u} \right)^2 - \left(\frac{E_b}{N_0} \right)_1 \left\{ \overline{\mathbf{u}^T (-\mathbf{A})^{2(K+1)} \mathbf{u}} + \overline{(\mathbf{u}^T (-\mathbf{A})^{(K+1)} \mathbf{u})^2} - 2 \left[\overline{\mathbf{u}^T (-\mathbf{A})^{K+1}} \mathbf{u} \right]^2 \right\}} \quad (53)$$

or using (52)

$$D = \frac{\mathbf{u}^T \left(\sum_{i=0}^K \overline{(-\mathbf{A})^i} - \sum_{i=K+1}^{2K+1} \overline{(-\mathbf{A})^i} \right) \mathbf{u}}{\left(1 - \mathbf{u}^T \overline{(-\mathbf{A})^{K+1}} \mathbf{u} \right)^2 - \left(\frac{E_b}{N_0} \right)_1 \left\{ \overline{\mathbf{u}^T (-\mathbf{A})^{2(K+1)} \mathbf{u}} + \overline{(\mathbf{u}^T (-\mathbf{A})^{(K+1)} \mathbf{u})^2} - 2 \left[\overline{\mathbf{u}^T (-\mathbf{A})^{K+1}} \mathbf{u} \right]^2 \right\}} \quad (54)$$

Substituting as before $\frac{1}{2} [Q^{-1}(P_b(E))]^2$ for $(E_b/N_0)_1$ allows D to be evaluated for

a given value of $P_b(E)$.

Although in principle either an "exact" or approximate expression for D can be obtained from (54) depending on whether the elements of the matrix Γ are modeled as binomial or Gaussian, the difficulty in arriving at a closed form result for D lies in the evaluation of the first element of the mean of A^i , i integer and also the mean of the square of the first element of A^i . Using *Mathematica* software and a Gaussian model for the γ_{ij} 's, we have been successful in carrying out these evaluations for $K = 1, 2$, and 3 leading to the following results:

$$\begin{aligned}
\mathbf{u}^T \overline{(-A)^n} \mathbf{u} &= 0, \quad n \text{ odd} \\
\mathbf{u}^T \overline{(-A)^0} \mathbf{u} &= 1, \quad \mathbf{u}^T \overline{(-A)^2} \mathbf{u} = \frac{M-1}{\eta}, \quad \mathbf{u}^T \overline{(-A)^4} \mathbf{u} = \left(\frac{M-1}{\eta} \right) \left(\frac{2M-1}{\eta} \right), \\
\mathbf{u}^T \overline{(-A)^6} \mathbf{u} &= \left(\frac{M-1}{\eta} \right) \left(\frac{5M^2-3M+1}{\eta^2} \right), \quad \mathbf{u}^T \overline{(-A)^8} \mathbf{u} = \left(\frac{M-1}{\eta} \right) \left(\frac{14M^3-5M^2+17M-21}{\eta^3} \right) \\
\overline{(\mathbf{u}^T \overline{(-A)^2} \mathbf{u})^2} &= \left(\frac{M-1}{\eta} \right) \left(\frac{M+1}{\eta} \right), \quad \overline{(\mathbf{u}^T \overline{(-A)^3} \mathbf{u})^2} = 2 \left(\frac{M-1}{\eta} \right) \left(\frac{M-2}{\eta^2} \right) \\
\overline{(\mathbf{u}^T \overline{(-A)^4} \mathbf{u})^2} &= \left(\frac{M-1}{\eta} \right) \left(\frac{4M^3+12M^2+11M+3}{\eta^3} \right)
\end{aligned} \tag{55}$$

For $K = 1$ and 2, using the results of (55) in (54) gives the closed form results for D found in (35') and (42) respectively. For $K = 3$, the corresponding result is given by

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_1} = \frac{1 + \left(\frac{M-1}{\eta} \right) \left(1 - \frac{2M-1}{\eta} - \frac{5M^2-3M+1}{\eta^2} \right)}{\left(1 - \frac{(M-1)(2M-1)}{\eta^2} \right)^2 - \left(\frac{E_b}{N_0} \right)_1 \frac{(M-1)(10M^3+23M^2+18M-16)}{\eta^4}} \tag{56}$$

An illustration of D versus M as computed from (50) is superimposed on the results of Figs. 8a and 8b. We observe from Fig. 8a that for $P_b(E) = 10^{-2}$, the performance of the three stage linear cancellation scheme is actually worse than that of the two stage scheme. A possible explanation for this seemingly strange result is as follows. Observe that the first term in the denominator of

D in (54) represents the normalized square of the mean of the decision variable [see (50)]. The part of this term corresponding to $\mathbf{u}^T(-\mathbf{A})^{K+1}\mathbf{u}$ represents self-noise which in view of the first equation in (55) is absent when K (the number of stages) is even and present when K is odd. Thus the the first term in the denominator of D is equal to unity when K is even and less than unity when K is odd, i.e., this term alternates as a function of the number of stages between unity and something smaller than unity. If $(E_b/N_0)_1$ is small enough, as would be the case for high user error probability, so that the first term in the denominator of (54) tends to dominate, then for a given number of users the degradation will alternately get better or worse as a function of the number of stages. This alternation has been observed by simulation for 1, 2, 3, and 4 stage IC schemes, i.e., 2 stages is better (smaller D) than 1 stage, 3 stages is worse than 2 stages, but 4 stages is again better than 3 stages but by a smaller amount. At small user error probabilities where $(E_b/N_0)_1$ is large, this alternation does not take place. To demonstrate this, Fig. 8b is the corresponding plot to Fig. 8a for $P_b(E) = 10^{-5}$. We observe here that the performance of the three stage linear cancellation scheme is better than that of the two stage scheme as one might ordinarily expect to be the case. Continuing beyond $K = 3$ is extremely tedious analytically and is not justified from a practical standpoint in view of the limited amount of performance improvement anticipated.

4.0 Improved Parallel Interference Cancellation Based on Joint Maximum-Likelihood Considerations

In the previous sections, we discussed a so-called *brute force* multistage IC scheme in which the receiver attempts at each stage (iteration) to fully cancel the multiuser interference. In Appendix B it is argued that, since in the early stages of interference cancellation the tentative decisions are less reliable than they are in later stages, the brute force philosophy of entirely cancelling the interference at each iteration stage may not result in a receiver with the best performance. Indeed, it is suggested there that a better philosophy might be one which in the early stages only partially cancels the multiuser interference with the amount of cancellation increasing as one continues to iterate toward the ultimate final data decisions, i.e., as the fidelity of the tentative decisions improves. With this philosophy in mind, a decision statistic is derived in Appendix B based upon joint ML-type considerations which leads to the multistage receiver structure of Fig. 1 with k th stage as in Fig. 2b. The modification of Fig. 2a which results in Fig. 2b is the inclusion of a parameter

p_k to allow for *partial* cancellation of the multiuser interference at the k th stage. Intuitively, one would expect that the value of p_k (which depends on the particular stage through the subscript k) would monotonically increase as one progresses toward the final data decision. Indeed, the numerical results to be presented later on bear out this intuition.

We begin our discussion of this modified parallel IC cancellation scheme by once again considering the case of tentative hard decisions with equal power, synchronous users.

4.1 Tentative Hard Decisions – Equal Power, Synchronous Users

For a single stage scheme, it is straightforward to show that the normalized output of the I&D corresponding to the *final* decisions is still given by (5) where β_i now takes on the more general form

$$\beta_i = a_{i0} - p_1 \operatorname{sgn} \left[\operatorname{Re} \left\{ \sqrt{E_b} \left(a_{i0} + \sum_{\substack{m=1 \\ m \neq i}}^M a_{m0} \gamma_{im} e^{j(\theta_m - \theta_i)} \right) + n_i e^{-j\theta_i} \right\} \right] \quad (57)$$

Note that for $p_1 = 1$, (57) reduces to (6). Following the development in Section 3.1, it is straightforward to see that the average probability is still given by (10) together with (11) where β_i of (57) rather than (6) is used in the definition of ξ_{ii} . When this is done, the required moments of (11) can be determined by applying the following replacements to the moments in (13):

$$\overline{\xi_{ii}} \Rightarrow p_1 \overline{\xi_{ii}} \quad (58a)$$

$$\overline{\xi_{ii}^2} \Rightarrow \frac{(1-p_1)^2 \sigma^2}{2} + p_1 \overline{\xi_{ii}^2} \quad (58b)$$

$$\overline{\xi_{ii} \xi_{ij}} \Rightarrow p_1^2 \overline{\xi_{ii} \xi_{ij}} \quad (58c)$$

$$\frac{2}{\sqrt{E_b}} \overline{N_1 \xi_{ii}} \Rightarrow p_1 \frac{2}{\sqrt{E_b}} \overline{N_1 \xi_{ii}} \quad (58d)$$

which reduce, respectively, to (13a-d) when $p_1 = 1$. For $p_1 = 0$, we have

$\overline{\xi_{1i}} = \overline{\xi_{1i}\xi_{1j}} = \overline{N_1\xi_{1i}} = 0$ and $\overline{\xi_{1i}^2} = \sigma^2/2 = 1/2\eta$. Thus, for this case, Λ of (11) simplifies to

$$\Lambda = \frac{1}{1 + (M-1)\eta^{-1}(E_b/N_0)_R} \quad (59)$$

Since from (10) and (15), $(E_b/N_0)_1 = \Lambda(E_b/N_0)_R$, then from (59) and the definition of the degradation factor, D , we have

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_1} = \frac{1}{\Lambda} = 1 + (M-1)\eta^{-1}(E_b/N_0)_R = 1 + (M-1)\eta^{-1}D(E_b/N_0)_1 \quad (60)$$

or solving for D ,

$$D = \frac{1}{1 - (M-1)\eta^{-1}(E_b/N_0)_1} \quad (61)$$

which agrees with (17), i.e., the degradation factor corresponding to conventional CDMA with no interference cancellation.

Figure 9 is a plot (obtained by computer simulation) of the degradation factor D versus the number of users M with p_1 as a parameter and the same values of $P_b(E)$ and η as in Fig. 4. We observe that an optimum (in the sense of maximizing M for a given D) p_1 exists whose value is other than unity. That is, for a specified amount of multiuser degradation, partial interference cancellation allows a considerably larger number of users to be supported than the number permitted by the brute force IC scheme of Section 3.

4.2 Tentative Hard Decisions – Unequal Power, Synchronous Users

The results of Section 4.1 can be generalized to the case where the users have unequal powers in the same manner that the results of Section 3.1 were generalized to those of Section 3.2. In particular, the normalized output of the I&D corresponding to the *final* decisions of user j is still given by (24) with β_j now taking on the more general form in (57). Hence, the bit error probability of the desired (the j th) user is still given by (26) together with (27) where, however, the moments required in (27) are obtained from (58a-d).

4.3 Null Zone Tentative Decisions – Synchronous Users

The results of Section 4.1 can be generalized to the case where a null zone

hard decision device (see Fig. 3b) is employed for the tentative decisions in Fig. 2b in the same manner that the results of Section 3.1 were generalized to those of Section 3.3. Following the development in Section 4.1 for a single stage scheme with equal power synchronous users, then the normalized output of the I&D corresponding to the final decision on user 1's bit a_{10} is still given by (5) with β_i now defined analogous to (57) by

$$\beta_i = a_{i0} - p_1 \text{nsgn} \left[\text{Re} \left\{ \sqrt{E_b} \left(a_{i0} + \sum_{\substack{m=1 \\ m \neq i}}^M a_{m0} \gamma_{im} e^{j(\theta_m - \theta_i)} \right) + n_i e^{-j\theta_i} \right\} \right] \quad (62)$$

where "nsgn" again denotes the null zone signum function defined in (29). Note that for $p_1 = 1$, (62) reduces to (28). Following the development in Section 3.3, it is straightforward to see that the average probability is still given by (10) together with (11) where β_i of (62) rather than (28) is used in the definition of ξ_{1i} . When this is done, the required moments of (11) can be determined by applying the moment replacements of (58a-d) to (30a-d).

A plot of the degradation factor D versus the number of users M with p_1 as a parameter would again reveal that an optimum (in the sense of maximizing M for a given D) p_1 exists whose value is other than unity. That is, for a specified amount of multiuser degradation, partial interference cancellation again allows a considerably larger number of users to be supported than the number permitted by the brute force IC scheme of Section 3.

4.4 Linear (Infinitely Quantized Soft) Tentative Decisions – Equal Power, Synchronous Users

As in Section 4.4, we first consider the analysis of a single stage scheme using linear tentative decision devices and with equal power, synchronous users. Analogous to (32), the normalized output of the I&D corresponding to the *final* decisions of user 1 (again assumed without loss of generality to be the desired user) is given by

$$x_{11} = a_{10} \left[1 - p_1 \sum_{i=2}^M \gamma_{1i}^2 \right] \sqrt{E_b} - \sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\theta_i - \theta_1)} \left[(1 - p_1) \gamma_{1i} - \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{ik} \gamma_{1k} \right] + \left(n_1 - p_1 \sum_{i=2}^M \gamma_{1i} n_i \right) e^{-j\theta_1} \quad (63)$$

As in the brute force IC case, the decision statistic for user 1 includes an

attenuation of the desired signal component, residual user interference after cancellation, and additional thermal noise due to the presence of the other $M-1$ users. The mean (conditioned on $a_{00} = 1$) and variance of x_{11} can be evaluated in a form analogous to the results given in Appendix D. Using the results of these evaluations, the average bit error probability for hard decisions made on the real part of x_{11} is again given by (33) where now

$$\begin{aligned} \left(\frac{E_b}{N_0}\right)_1 &= \left(\frac{E_b}{N_0}\right)_R \left(1 - p_1 \frac{M-1}{\eta}\right)^2 \left\{ 1 - p_1(2-p_1) \left(\frac{M-1}{\eta}\right) + p_1^2 \frac{(M-1)(M-2)}{\eta^2} \right. \\ &\quad \left. + \left(\frac{E_b}{N_0}\right)_R \frac{(M-1)}{\eta} \left[(1-p_1)^2 - 2p_1(1-p_1) \left(\frac{M-2}{\eta}\right) + \frac{p_1^2}{\eta} \left[M+2 + \frac{M^2-5M+2}{\eta} \right] \right] \right\}^{-1} \end{aligned} \quad (64)$$

or, equivalently, the degradation factor is given by

$$D = \frac{1 - p_1(2-p_1) \left(\frac{M-1}{\eta}\right) + p_1^2 \frac{(M-1)(M-2)}{\eta^2}}{\left(1 - p_1 \frac{M-1}{\eta}\right)^2 - \left(\frac{E_b}{N_0}\right)_1 \left(\frac{M-1}{\eta}\right) \left\{ (1-p_1)^2 - 2p_1(1-p_1) \left(\frac{M-2}{\eta}\right) + \frac{p_1^2}{\eta} \left[M+2 + \frac{M^2-5M+2}{\eta} \right] \right\}} \quad (65)$$

For $p_1 = 0$, this result reduces to (17), i.e., the degradation factor corresponding to conventional CDMA with no interference cancellation. Also, for $p_1 = 1$, (65) reduces to (35) corresponding to the case of full (brute force) interference cancellation. Figures 10a, b are plots of D versus M obtained from (65) with p_1 as a parameter and the same processing gain and average bit error probability as in Figures 9a, b. We observe that, as in the hard tentative decision case, an optimum (in the sense of maximizing M for a given D) p_1 exists whose value is other than unity. That is, for a specified amount of multiuser degradation, partial interference cancellation again allows a considerably larger number of users to be supported than the number permitted by the brute force IC scheme. Comparing Figures 10a, b with Figures 9a, b we observe that the optimum value of p_1 obtained at a given value of degradation D is different for the two cases.

Here again one can improve performance by going to a multiple stage scheme. For two stages of cancellation with linear tentative decisions as illustrated in Fig. 1 combined with Fig. 2b, it can be shown that the normalized output of the I&D corresponding to the *final* decisions of user 1 is

(by analogy with (37))

$$\begin{aligned}
x_{21} = & a_{10} \left[1 - [p_2(1-p_1) + p_1(1-p_2)] \sum_{i=2}^M \gamma_{1i}^2 + p_1 p_2 \sum_{i=2}^M \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} \right] \sqrt{E_b} \\
& + \sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\theta_i - \theta)} \left[p_1 p_2 \sum_{i=2}^M \sum_{\substack{k=1 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} - [p_2(1-p_1) + p_1(1-p_2)] \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} + (1-p_1)(1-p_2) \gamma_{1i} \right] \\
& + \left(n_1 - [p_2 + p_1(1-p_2)] \sum_{i=2}^M \gamma_{1i} n_i + p_1 p_2 \sum_{i=2}^M \sum_{\substack{k=1 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} n_k \right) e^{-j\theta}
\end{aligned} \tag{66}$$

Letting

$$P_1 = p_2 + p_1(1-p_2), \quad P_2 = p_1 p_2 \tag{67}$$

$$P_3 = p_2(1-p_1) + p_1(1-p_2), \quad P_4 = (1-p_1)(1-p_2)$$

then the mean of the desired signal term (conditioned on $a_{10} = 1$), the variances of the self noise, the user interference, and the equivalent thermal noise are respectively given by

$$E \left\{ \left[1 - P_3 \sum_{i=2}^M \gamma_{1i}^2 + P_2 \sum_{i=2}^M \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} \right] \sqrt{E_b} \right\} = \sqrt{E_b} \left[1 - P_3 \left(\frac{M-1}{\eta} \right) \right] \tag{68}$$

$$E \left\{ \left(\sqrt{E_b} a_{10} \left[P_2 \sum_{i=2}^M \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} - P_3 \left(\sum_{i=2}^M \gamma_{1i}^2 - \frac{M-1}{\eta} \right) \right] \right)^2 \right\} = E_b \frac{2(M-1)}{\eta^2} \left[P_2^2 \frac{(M-2)}{\eta} + P_3^2 \right] \tag{69}$$

$$\begin{aligned}
& \frac{1}{2} E \left\{ \left[\sqrt{E_b} \sum_{i=2}^M a_{i0} e^{j(\theta_i - \theta)} \left[P_2 \sum_{i=2}^M \sum_{\substack{k=1 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} - P_3 \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{1k} \gamma_{1i} + P_4 \gamma_{1i} \right] \right]^2 \right\} \\
& = \frac{E_b}{2} \left(\frac{M-1}{\eta} \right) \left[P_2^2 \left(\frac{5(M^2 - M + 1)}{\eta^2} \right) + P_3^2 \left(\frac{M-2}{\eta} \right) + P_4^2 + 2P_2 P_4 \left(\frac{2M-1}{\eta} \right) \right]
\end{aligned} \tag{70}$$

$$\frac{1}{2} E \left\{ \left[n_1 - [P_2 + P_1(1 - P_2)] \sum_{l=2}^M \gamma_{1l} n_l + P_1 P_2 \sum_{l=2}^M \sum_{k=1}^M \gamma_{lk} \gamma_{1l} n_k \right]^2 \right\} \quad (71)$$

$$= \frac{N_0}{2} \left\{ 1 + \left(\frac{M-1}{\eta} \right) \left[2P_2 - 2P_1 + P_1^2 - P_2(2P_1 - P_2) \left(\frac{2M-1}{\eta} \right) \right] \right\}$$

Note that (69)–(71) reduce, respectively, to (38)–(40) when $p_1 = p_2 = 1$ since in this case, $P_1 = P_2 = 1$ and $P_3 = P_4 = 0$. The corresponding degradation factor can be expressed in the form

$$D = \frac{(E_b/N_0)_R}{(E_b/N_0)_1} = \frac{1 + \left(\frac{M-1}{\eta} \right) A(M, p_1, p_2)}{\left[1 - P_3 \left(\frac{M-1}{\eta} \right) \right]^2 - \left(\frac{E_b}{N_0} \right)_1 \frac{(M-1)B(M, p_1, p_2)}{\eta}} \quad (72)$$

where

$$A(M, p_1, p_2) \triangleq 2P_2 - 2P_1 + P_1^2 - P_2(2P_1 - P_2) \left(\frac{2M-1}{\eta} \right) \quad (73)$$

$$B(M, p_1, p_2) \triangleq P_2^2 \left(\frac{5M^2 - M - 3}{\eta^2} \right) + P_3^2 \left(\frac{M+2}{\eta} \right) + P_4^2 + 2P_2 P_4 \left(\frac{2M-1}{\eta} \right)$$

The two stage linear partial interference cancellation scheme has also been computer simulated and the values of p_1 and p_2 jointly optimized to provide for a minimum degradation. The performance results are illustrated in Fig. 11 along with the corresponding results from Fig. 10a for a single stage scheme.

Finally, two and three stage nonlinear (hard tentative decisions) partial interference cancellation schemes were computer simulated with optimized values of the partial cancellation parameters. The performance results are illustrated in Fig. 12 along with the corresponding simulation results for a single stage scheme. We observe that, for the parameters considered and uncoded BPSK users, a three stage nonlinear partial IC scheme allows as many as 80 users with a degradation of only 1 dB as compared to 9 users in a conventional CDMA system with the same degradation. This ideally represents an almost nine-fold increase in the user capacity of the system.

5.0 Parallel Interference Cancellation for CDMA with Coded Users

Until now we have considered interference cancellation for CDMA systems in which the users transmit uncoded BPSK modulation. The IC techniques discussed in the previous sections are equally applicable to CDMA systems whose users employ error correction coding. To illustrate this applicability, we shall consider the performance of a CDMA system in which the users transmit identical convolutionally coded BPSK modulations.

When used with a convolutionally coded modulation, there essentially exist two options for handling the "tentative" decisions made by each user in the presence of the interference. In one instance, these decisions can be made ignoring the fact that the each user's data is encoded (that is, treating it as if it was an uncoded symbol stream). These symbol-by-symbol decisions would then be respread and remodulated directly on the carrier of each user. This is the simplest of the two options in that the only delay involved is the time (one coded symbol interval) to make these decisions. The second option is to treat the incoming data stream per user as coded symbol streams (as they truly are) and make the tentative decisions using a convolutional decoder (soft or hard decision) with the resulting bit stream being reencoded before respreading and remodulation on the carrier. This is the better (from a performance standpoint) of the two options but now involves a delay on the order of the decoder buffer size as well as a decoder/reencoder pair for each user at each IC stage. In view of this additional complexity, we shall consider only the first of the two options.

The performance analysis of conventional (no IC) CDMA with convolutionally coded users was given in [18]. In this work, a constraint length $K = 7$, rate $r = 1/2$ convolutional code with Viterbi decoding and eight-level soft decisions was assumed as the error correction coding/decoding scheme employed by each user. A summary of the performance of this code applied to PSK modulation in the presence of AWGN only, i.e., no multiuser interference, was given there in Table I. In particular, at a bit error probability $P_b(E) = 10^{-5}$, the required coded bit energy-to-noise ratio is $E_b/N_0 = 4.5$ dB (as opposed to $E_b/N_0 = 9.6$ dB for uncoded data) which results in a coded symbol energy-to-noise ratio $E_s/N_0 = 1.5$ dB. Since the degradation factor only depends on the additional SNR required to overcome the multiuser interference and not on the specifics of the coding/decoding process, we can use the results of Section 3.1.1 as if we were

dealing with an uncoded system whose required *bit* energy-to-noise ratio is equal to the required *symbol* energy-to-noise ratio in the coded system. This is precisely what was done in [18] to assess the performance of conventional CDMA for coded users. In particular, for the above convolutionally encoded BPSK system, we have analogous to (16)

$$\left(\frac{E_s}{N_0}\right)_1 = \frac{E_s}{N_0 + (M-1)ST_c} = \frac{(E_s/N_0)_R}{1 + (M-1)\eta^{-1}r^{-1}(E_s/N_0)_R} \quad (74)$$

where $\eta = T_b/T_c$ is still the ratio of the chip rate to the information bit rate as previously defined and $r = T_s/T_b$ is the code rate with T_s denoting the coded symbol duration. Since, however, the bit energy and coded symbol energy are related by $E_s = rE_b$, then (74) can be rewritten as

$$\left(\frac{rE_b}{N_0}\right)_1 = \frac{(rE_b/N_0)_R}{1 + (M-1)\eta^{-1}r^{-1}(rE_b/N_0)_R} \quad (75a)$$

or

$$\left(\frac{E_b}{N_0}\right)_1 = \frac{(E_b/N_0)_R}{1 + (M-1)\eta^{-1}(E_b/N_0)_R} \quad (75b)$$

which is identical to (16). Hence, the degradation factor given in (17) is still applicable to the coded case. Since, however, in the coded case it is the coded symbols that are spread by the PN modulation, we shall in the numerical evaluations choose a value of η equal to r^{-1} times the value used for the uncoded numerical results, e.g., for the rate 1/2 code, we shall use $\eta = 200$, in order to maintain the analogy with the uncoded system. Once again, this is what was done in [18].

Insofar as the analysis for CDMA with IC is concerned, we can make the same analogy between the uncoded and coded cases as was done above since here again the IC process results from symbol-by-symbol decisions at each stage. Hence, as long as the Gaussian assumptions previously made for modelling the interference are valid, we can directly apply the uncoded user analytical results derived in the previous sections to the coded case. As far as simulation results are concerned, this allows using an encoder and decoder for only the desired user while treating the remainder of the users (the interferers) as uncoded.

Fig. 13 is a plot (obtained by computer simulation) of the degradation

factor D versus the number of users M for a single stage IC scheme with hard tentative decisions. The value of p_1 has been optimized to yield the best performance (minimum D) and, in accordance with the above discussion, the values of $P_b(E)$ and η have been chosen analogous to those in Fig. 4.

6.0 Conclusions

The inclusion of multistage parallel interference cancellation in a CDMA receiver can significantly improve its performance relative to that of a conventional CDMA receiver where no interference cancellation is attempted. A partial interference cancellation philosophy, in which the amount of interference cancelled is related to the fidelity of the tentative decisions involved in forming the interference estimate, is in general superior to a brute force philosophy of entirely cancelling the interference at each stage. Using a null zone type of device for making the tentative decisions at the various stages of the cancellation process is superior to using either a hard limiter or linear device. The linear device on the other hand has the advantage that the receiver implementation does not require knowledge of the user powers nor does it need carrier synchronization at the various stages. The latter implies that the final data decisions can be performed with a differential (rather than a coherent) detector. The technique is equally applicable to coded as well as uncoded modulations.

References

- [1] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Transactions on Information Theory*, vol. IT-32, no. 1, January 1986, pp. 85-96.
- [2] S. Verdu, "Optimum multiuser asymptotic efficiency," *IEEE Transactions on Communications*, vol. COM-34, no. 9, September 1986, pp. 890-897.
- [3] K. S. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 1, January 1979, pp. 181-183.
- [4] R. Lupas and S. Verdu, "Linear multiuser detectors for asynchronous code division multiple access channels," *IEEE Transactions on Information Theory*, vol. IT-35, no. 1, January 1989, pp. 123-136.

- [5] R. Lupas and S. Verdu, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Transactions on Communications*, vol. COM-38, no. 4, April 1990, pp. 496–508.
- [6] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels," *IEEE Transactions on Selected Areas in Communications*, vol. 8, no. 4, May 1990, pp. 641–649. Also presented in part at the 1989 IEEE Communication Theory Workshop.
- [7] P. W. Dent, "CDMA subtractive demodulation", U. S. Patent 5,218,619, June 8, 1993.
- [8] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications." *IEEE Transactions on Communications*, vol. 38, no. 4, April 1990, pp. 509–519.
- [9] M. K. Varanasi and B. Aazhang, "Near optimum detection in synchronous code-division multiple-access systems," *IEEE Transactions on Communications*, vol. 39, no. 5, May 1991, pp. 725–736.
- [10] M. K. Varanasi and S. Vasudevan, "Multiuser detectors for synchronous CDMA communication over non-selective Rician fading channels," *IEEE Transactions on Communications*, vol. 42, no. 2/3/4, February/March/April 1994, pp. 711–722.
- [11] Y. C. Yoon, R. Kohno, and H. Imai, "Cascaded co-channel interference cancelling and diversity combining for spread-spectrum multi-access over multipath fading channels," Symposium on Information Theory and its Applications (SITA '92), Minakami, Japan, September 8–11, 1992.
- [12] Y. C. Yoon, R. Kohno, and H. Imai, "A spread-spectrum multi-access system with a cascade of co-channel interference cancellers for multipath fading channels," International Symposium on Spectrum Techniques and Applications (ISSSTA '92), Yokohama, Japan, November 29–December 2, 1992.
- [13] Y. C. Yoon, R. Kohno, and H. Imai, "A spread-spectrum multiaccess system with cochannel interference cancellation," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 7, September 1993, pp. 1067–1075.

- [14] M. Kawabe, T. Kato, A. Kawahashi, T. Sato and A. Fukasawa, "Advanced CDMA scheme based on interference cancellation," *Proceedings of the 43rd Annual IEEE Vehicular Technology Conference*, May 18–20, 1993, pp. 448–451.
- [15] H. V. Poor, "Signal processing for wideband communications," *IEEE Information Theory Society Newsletter*, vol. 42, no. 2, June 1992.
- [16] S. Verdu and H. V. Poor, "Abstract dynamic programming models under commutativity conditions," *SIAM Journal of Control and Optimization*, vol. 25, July 1987, pp. 990–1006.
- [17] H. V. Poor, "On parameter estimation in DS/SSMA formats," in *Advances in Communications and Signal Processing*, A. Porter and S. C. Kak, Eds., Springer-Verlag, New York, 1989, pp. 59–70.
- [18] C. L. Weber, G. K. Huth, and B. H. Batson, "Performance considerations of code division multiple-access systems," *IEEE Transactions on Vehicular Technology*, vol. VT-30, no. 1, February 1981, pp. 3–10.
- [19] A. J. Viterbi, "When not to spread spectrum – A Sequel," *IEEE Communication Society Magazine*, vol. 23, April 1985, pp. 12–17.
- [20] D. Divsalar and M. K. Simon, "Multiple symbol differential detection of MPSK," *IEEE Transactions on Communications*, vol. 38, no. 3, March 1990, pp. 300–308.
- [21] D. Divsalar and M. K. Simon, "Improved CDMA performance using parallel interference cancellation," *MILCOM '94 Conference Proceedings*, October 3–5, 1994, pp. 911–917.

Figure 1
AN L-STAGE INTERFERENCE CANCELLATION SCHEME WITH PARALLEL PROCESSING FOR CDMA
 (Complex Baseband Model)

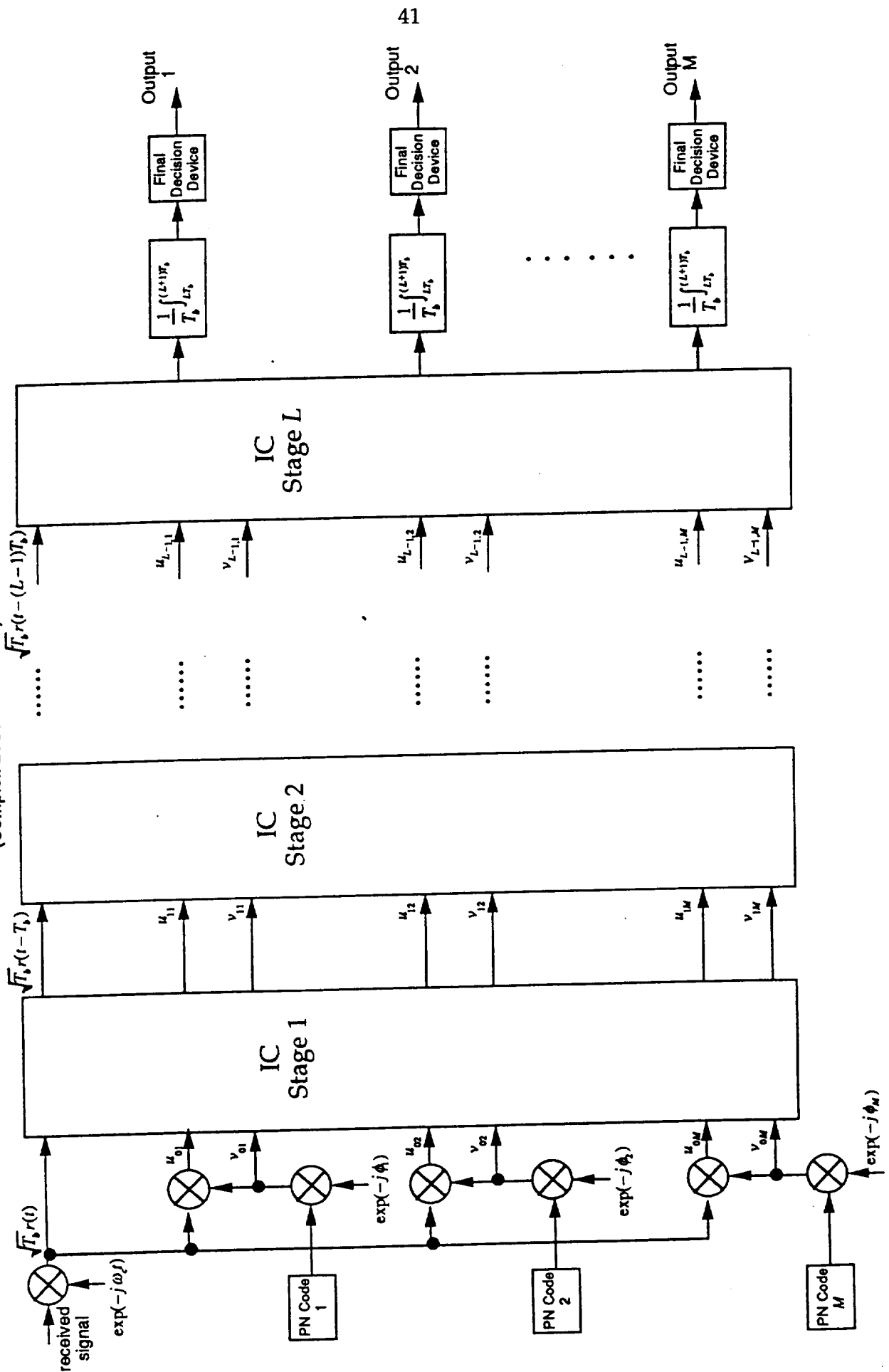


Figure 2a

kth STAGE OF BRUTE FORCE INTERFERENCE CANCELLATION

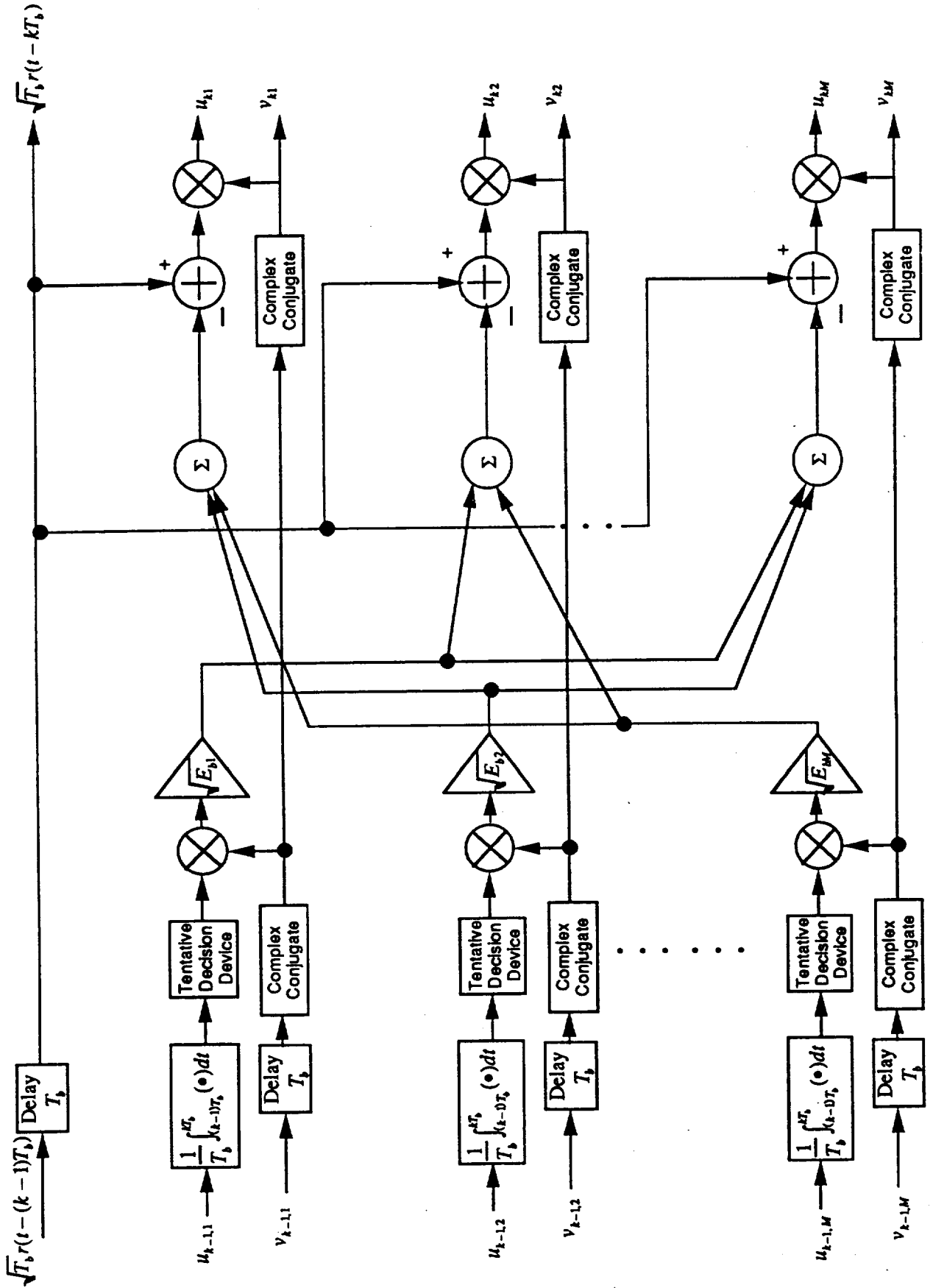
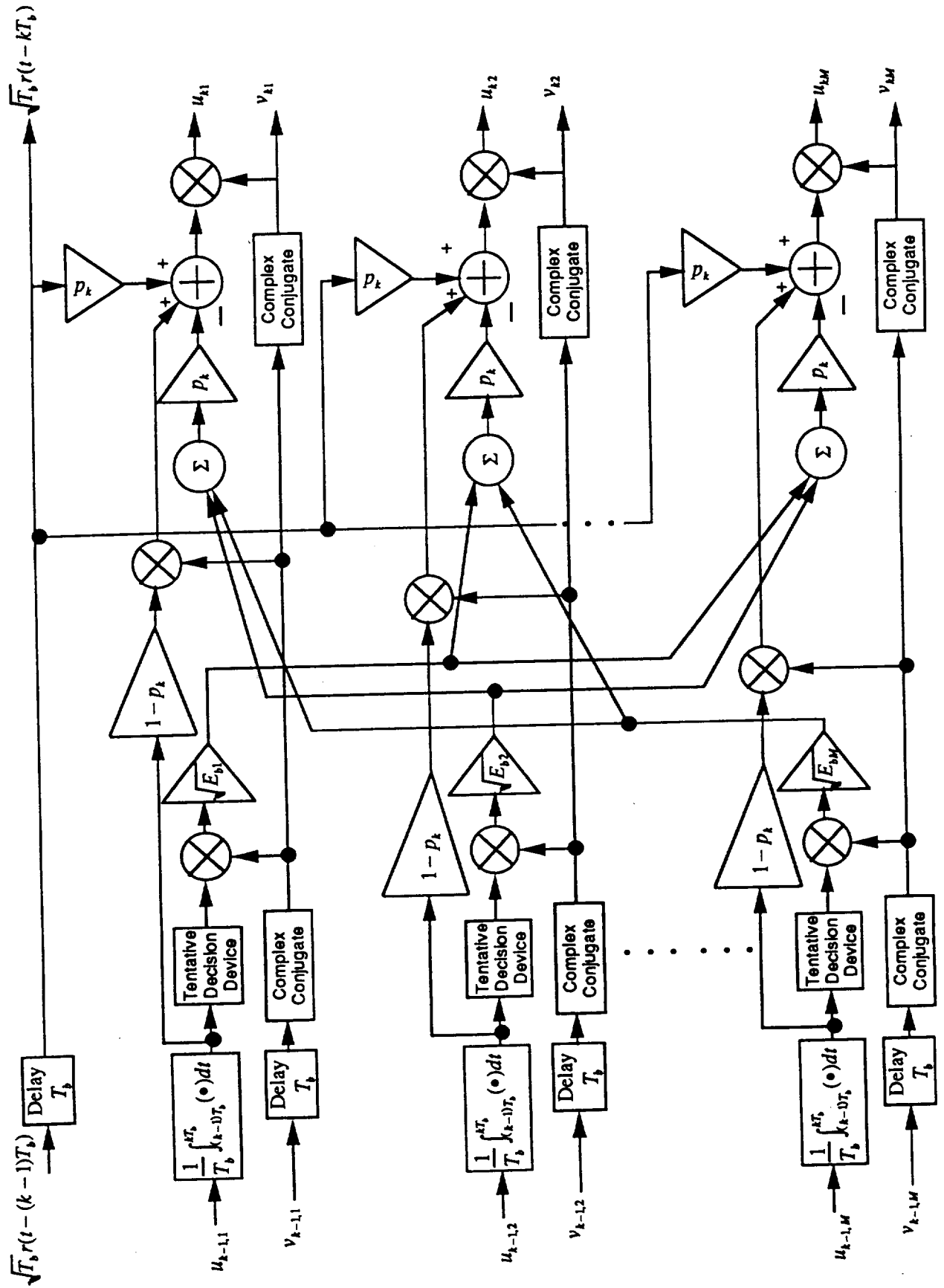
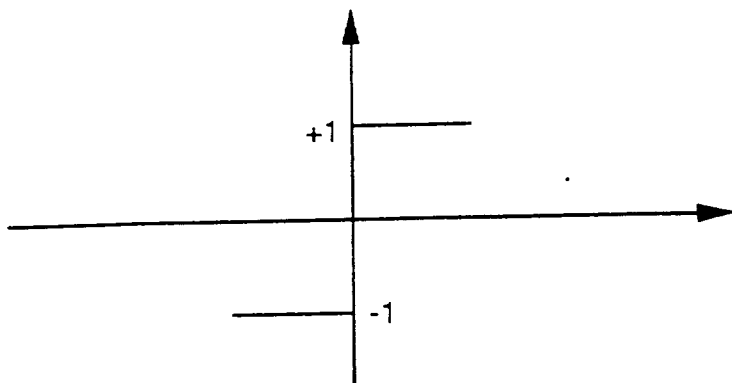
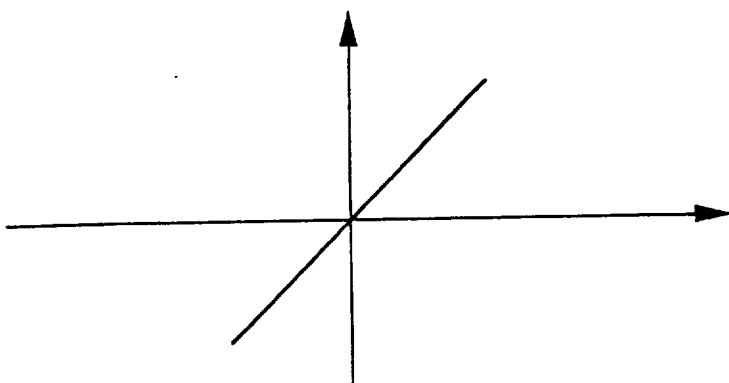


Figure 2b
 kth STAGE OF PARTIAL INTERFERENCE CANCELLATION

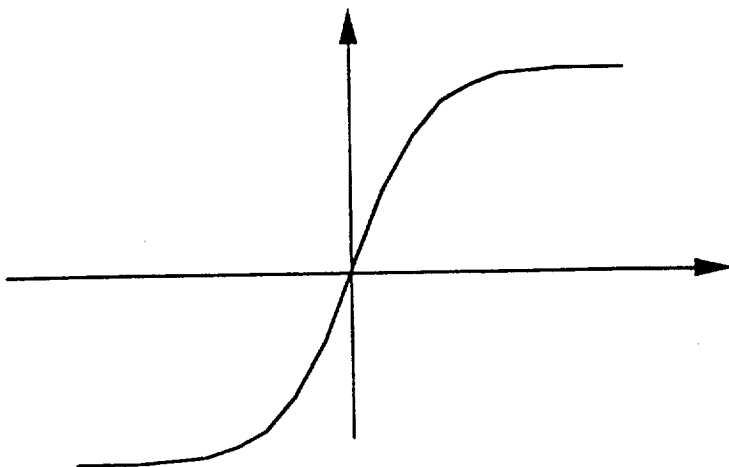




(a) Hard Limiter (One-Bit Quantizer)

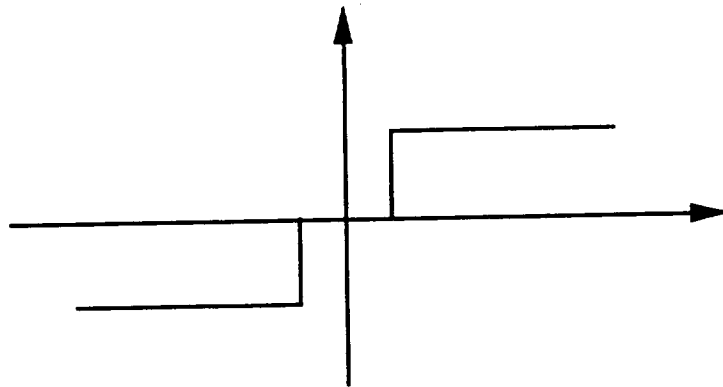


(b) Linear (Infinite-Bit Quantizer)



(c) Hyperbolic Tangent (Soft Quantizer)

Figure 3 Tentative Decision Devices



(d) Null Zone Device

Figure 3 Tentative Decision Devices (continued)

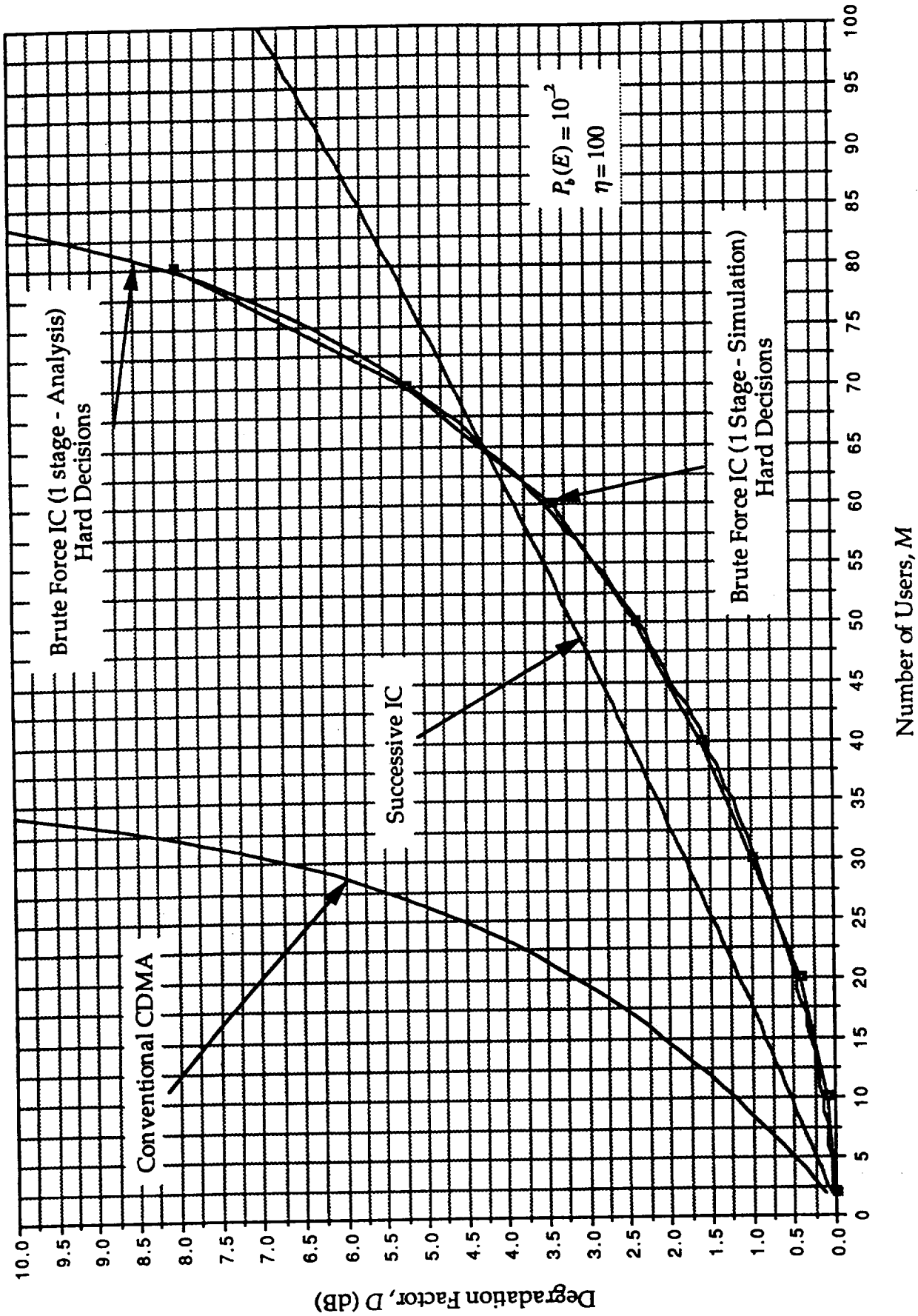


Fig. 4. Degradation Factor versus Number of Users for CDMA with and without Interference Cancellation - One Stage, Brute Force Parallel Cancellation, Hard Tentative Decisions, Equal Power Users

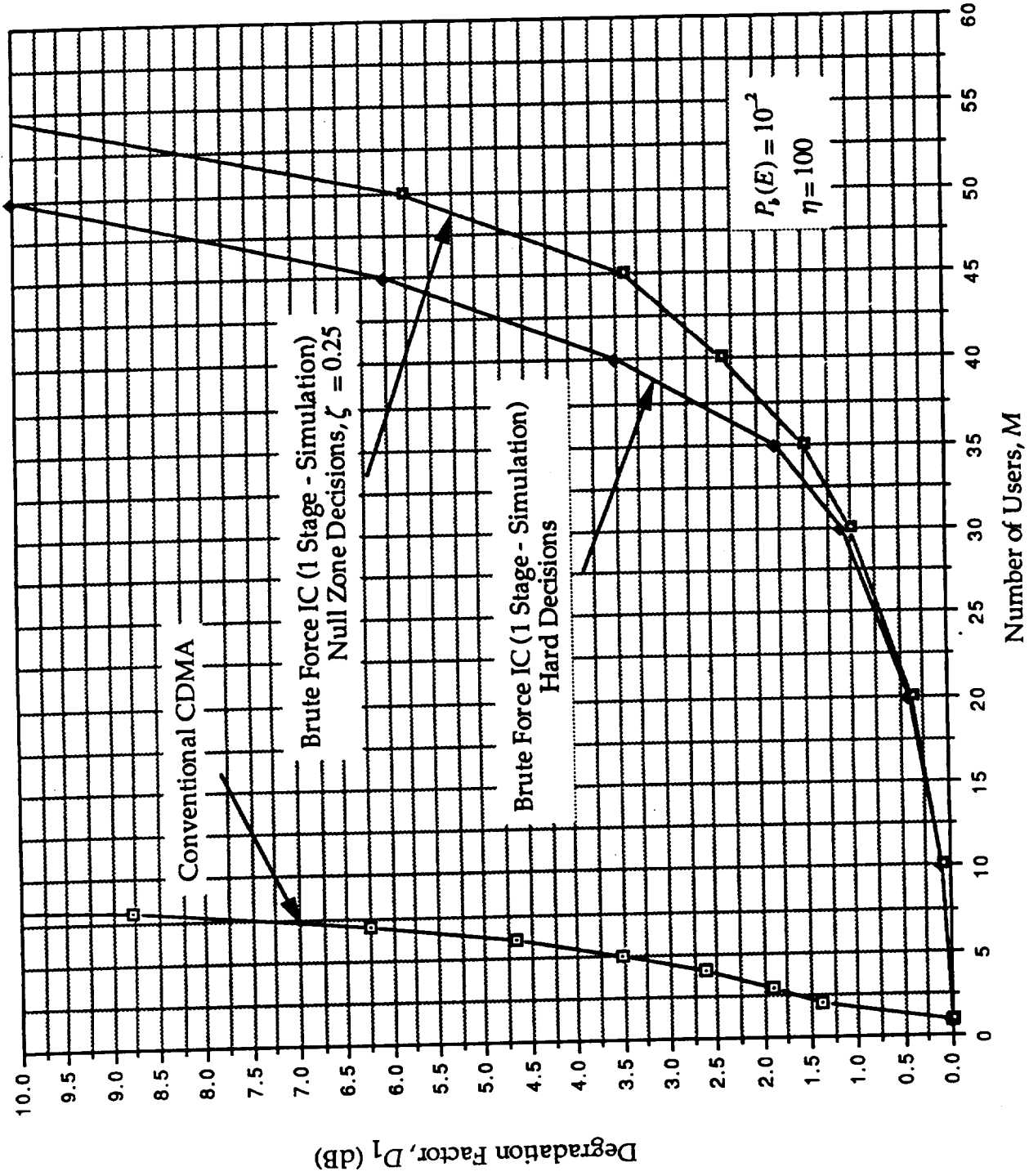


Fig. 5. Comparison of the Degradation Factor of the Lowest Power User versus Number of Users for CDMA with Brute Force IC – One Stage, Hard and Null Zone Tentative Decisions, Unequal Power Users

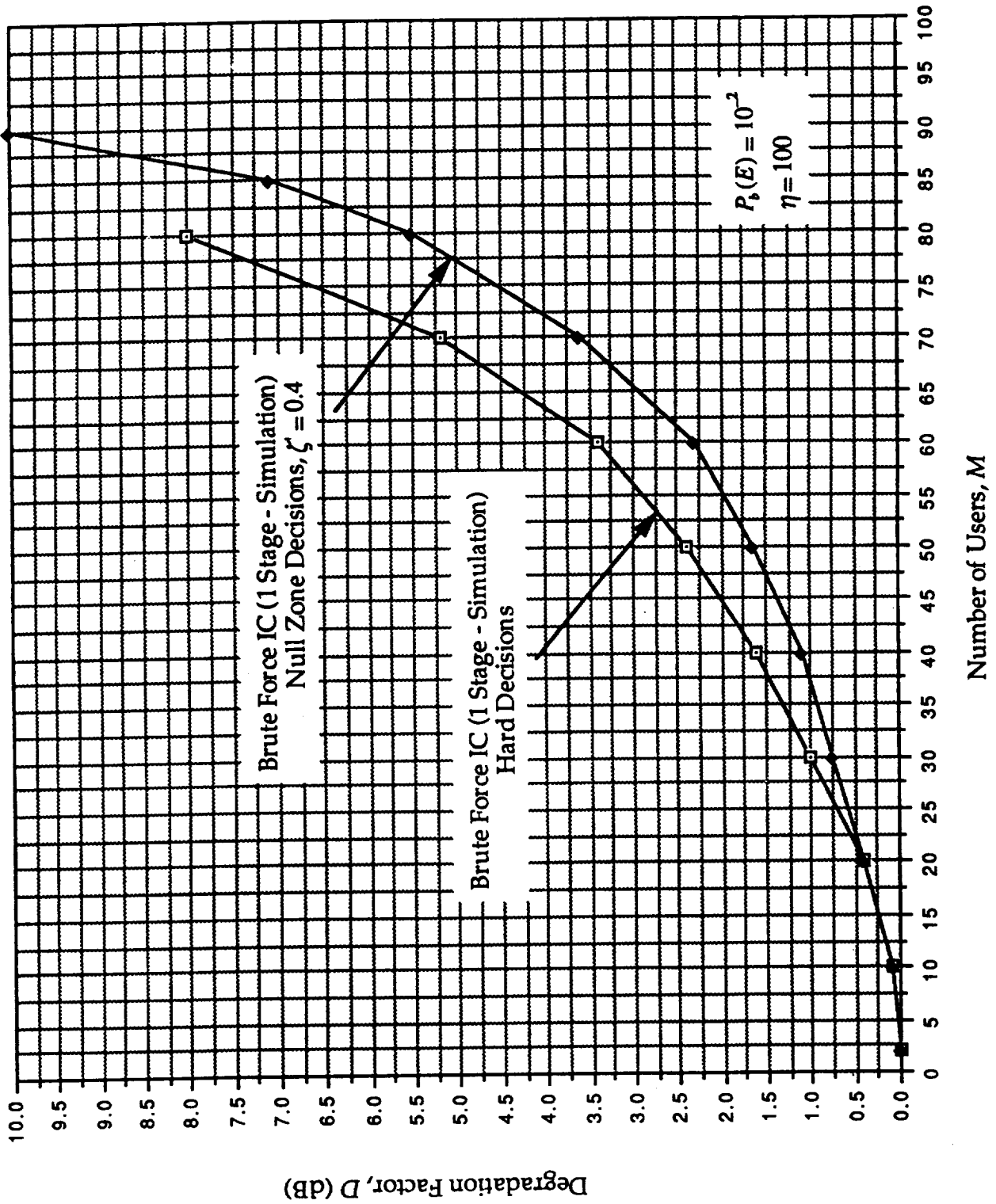


Fig. 6. Comparison of the Degradation Factor versus Number of Users for CDMA with Brute Force IC - One Stage, Hard and Null Zone Tentative Decisions, Equal Power Users

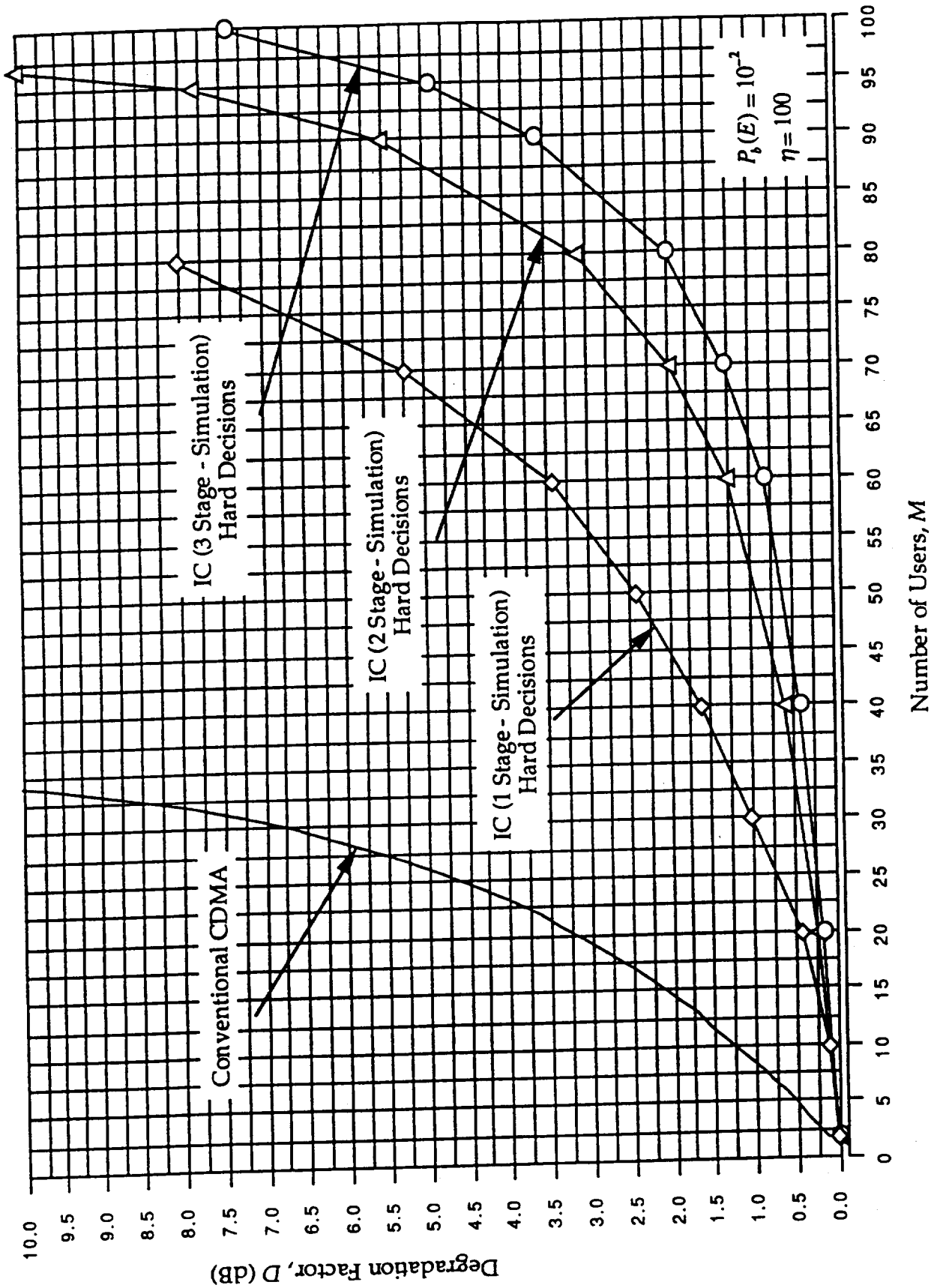


Fig. 7. A Comparison of the Degradation Factors for One, Two and Three Stage Brute Force Interference Cancellation - Hard Tentative Decisions, Equal Power Users

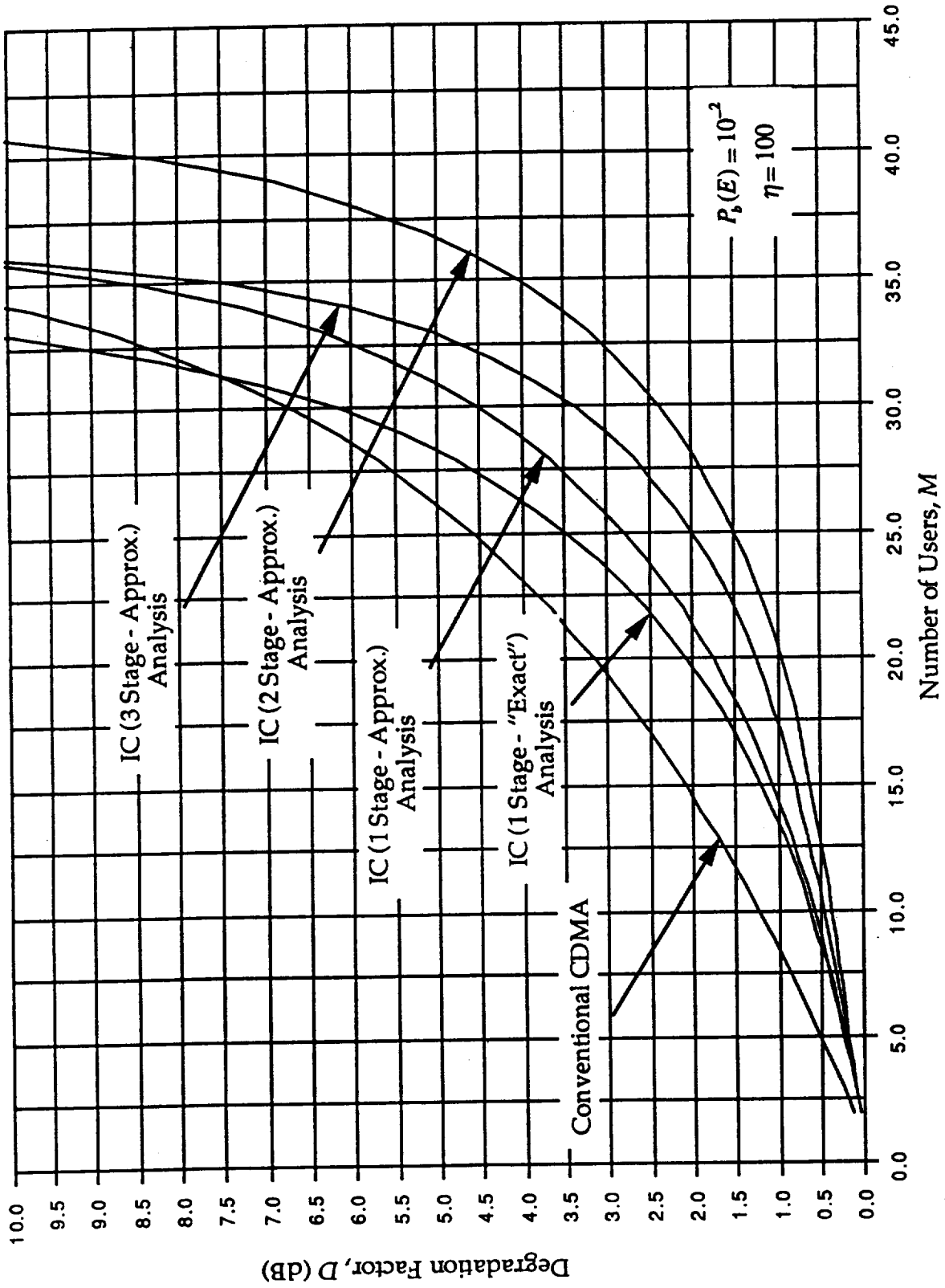


Fig. 8a. A Comparison of the Degradation Factors for One, Two, and Three Stage Brute Force Linear Interference Cancellation - Equal Power Users, $P_b(E) = 10^{-2}$

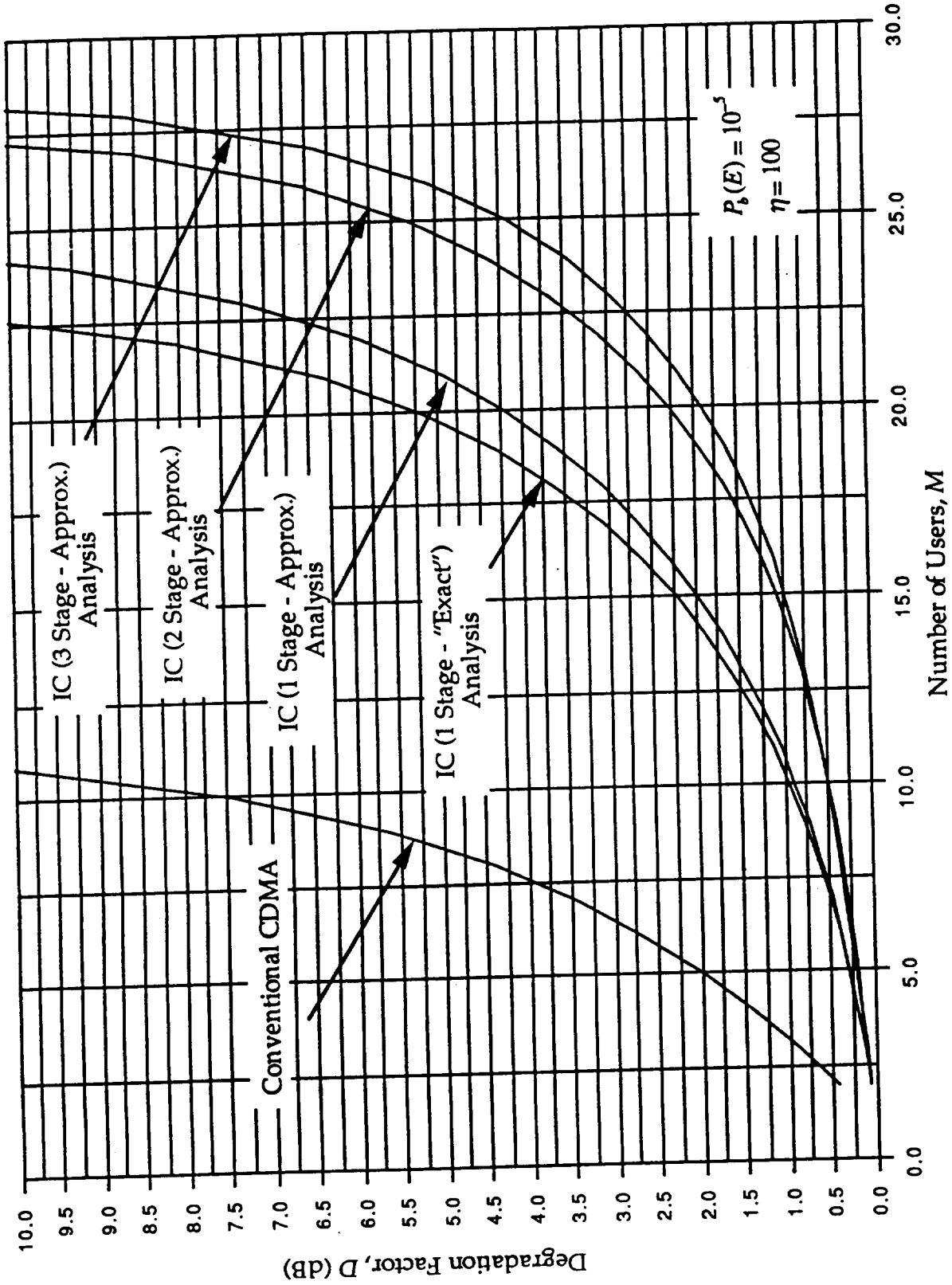


Fig. 8b. A Comparison of the Degradation Factors for One, Two, and Three Stage Brute Force Linear Interference Cancellation - Equal Power Users, $P_b(E) = 10^{-5}$

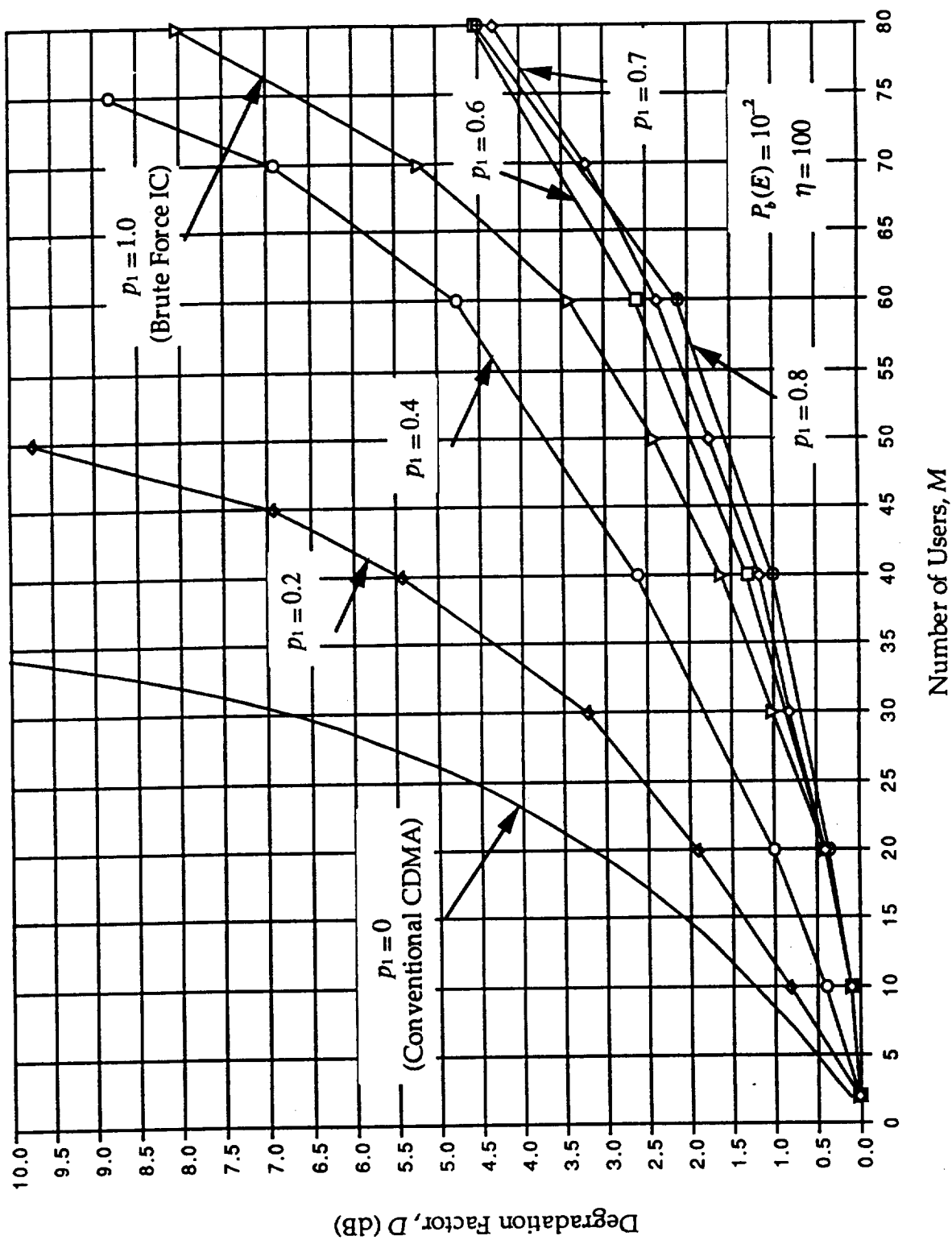


Fig. 9. Degradation Factor Performance for One Stage Partial Nonlinear IC as a Function of the Partial Cancellation Parameter (p_1) - Equal Power Users, $P_b(E) = 10^{-2}$

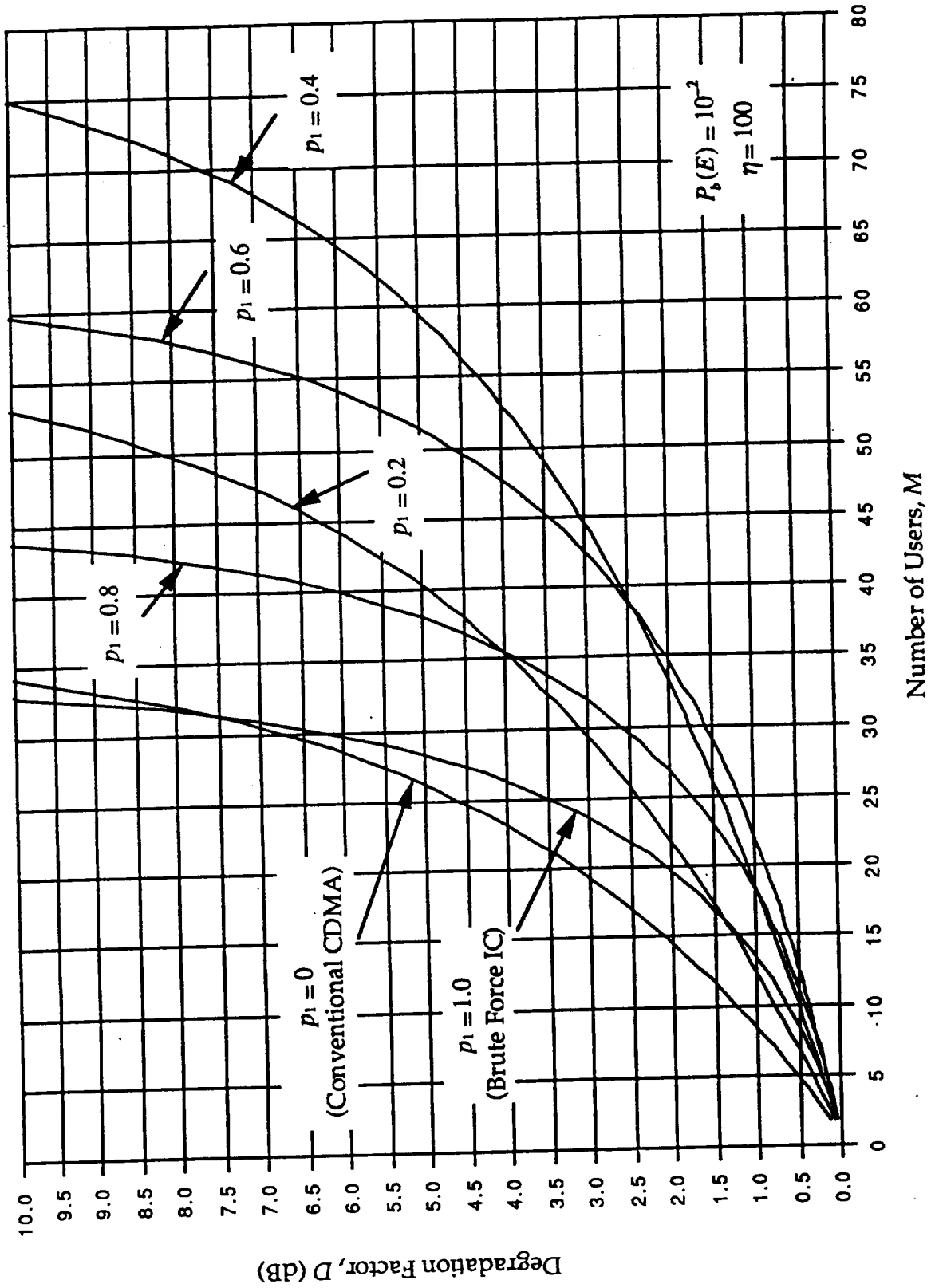


Fig. 10a. Degradation Factor Performance for One Stage Partial Linear IC as a Function of the Partial Cancellation Parameter (p_1) - Equal Power Users, $P_b(E) = 10^{-2}$

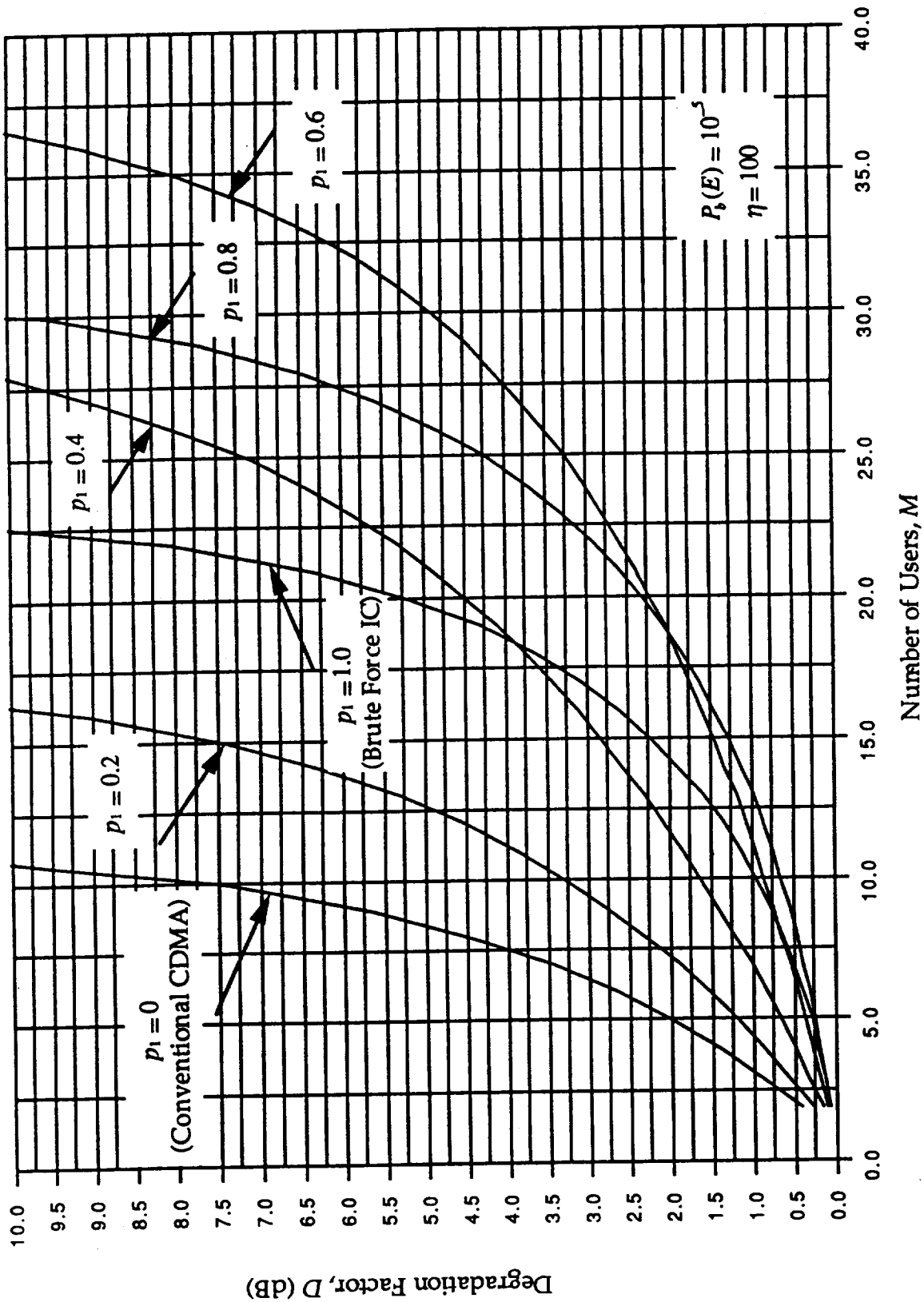


Fig. 10b. Degradation Factor Performance for One Stage Partial Linear IC as a Function of the Partial Cancellation Parameter (p_1) - Equal Power Users, $P_b(E) = 10^{-5}$

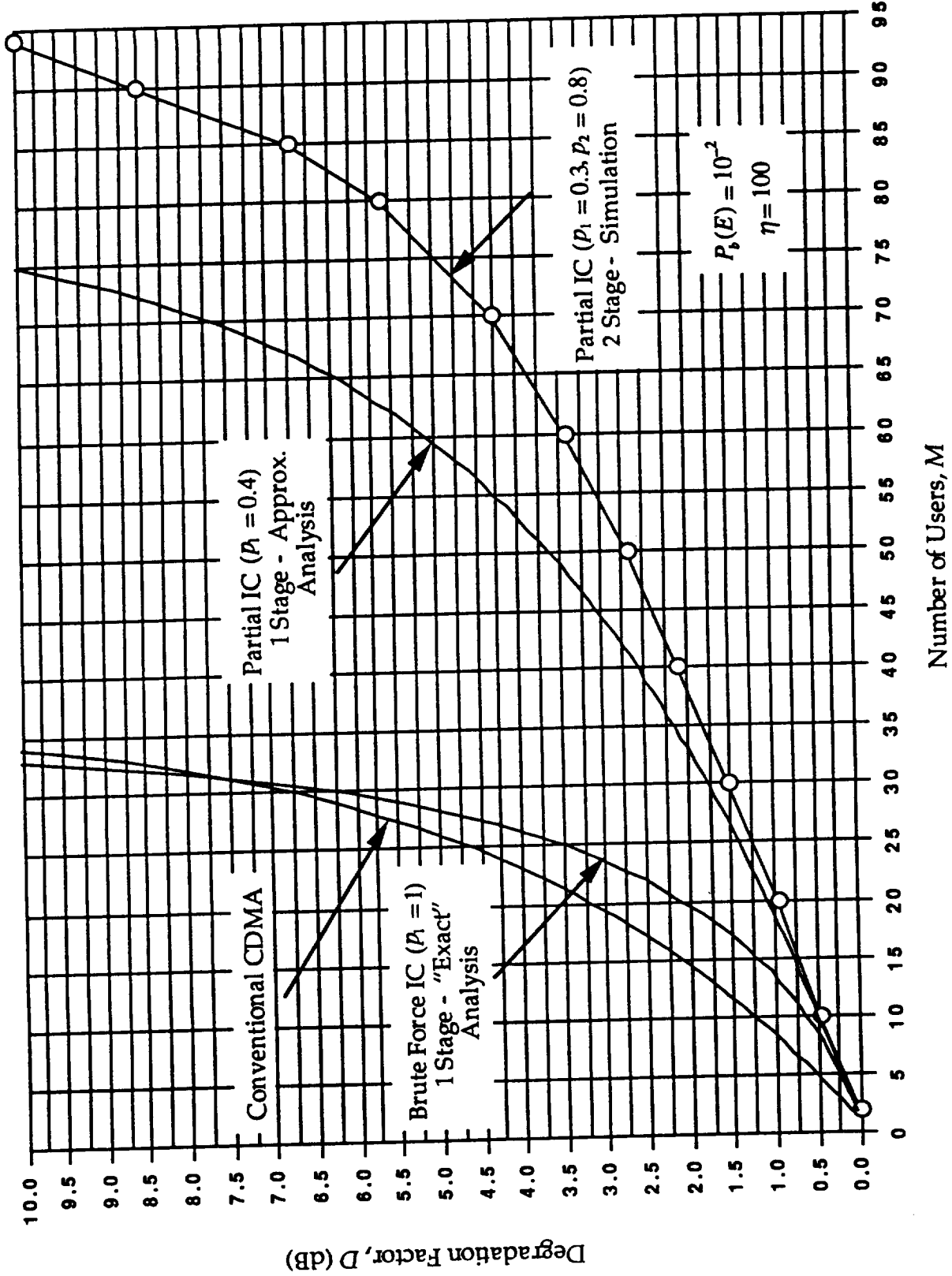


Fig. 11. A Comparison of the Degradation Factors for One and Two Stage Linear Interference Cancellation - Equal Power Users

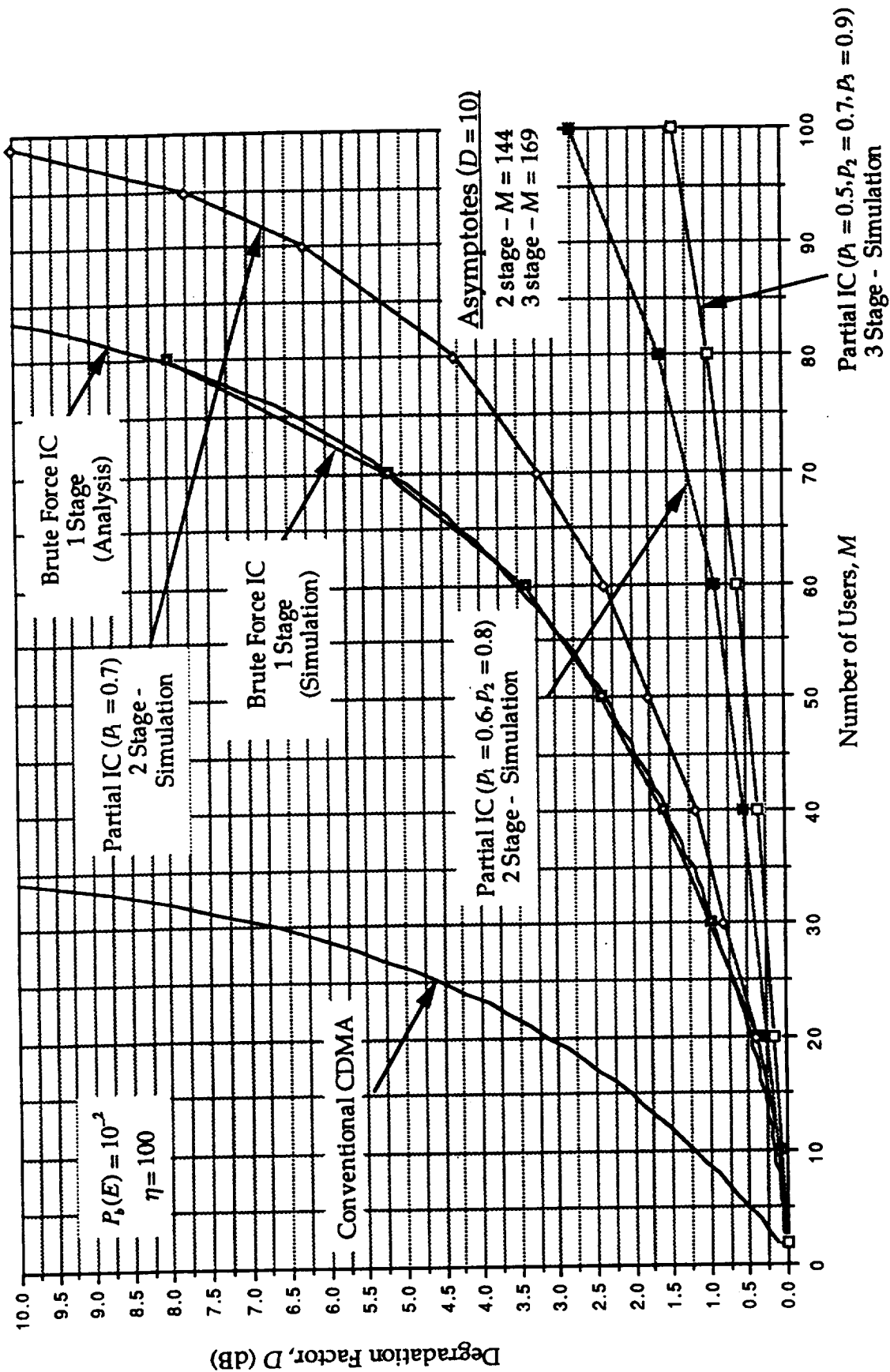


Fig. 12. A Comparison of the Degradation Factors for One, Two, and Three Stage Nonlinear Interference Cancellation - Equal Power Users

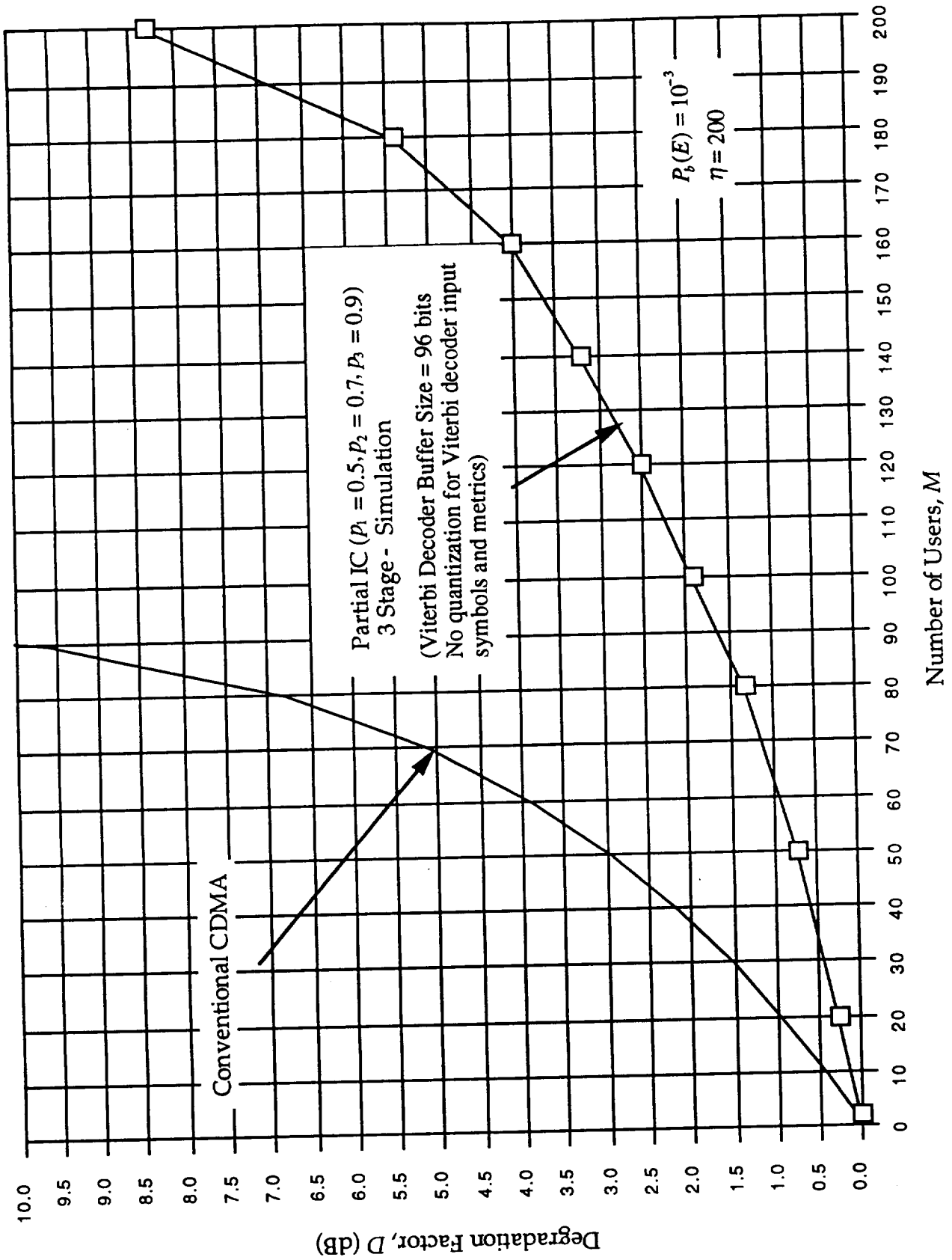


Fig. 13. Degradation Factor versus Number of Users for Convolutionally Coded CDMA with and without Interference Cancellation - Hard Tentative Decisions, Equal Power Users

Appendix A

Analytical Models for Simulation of Multiuser Communication Receivers

Consider the received signal of Eq. (1) corresponding to the transmission of M BPSK users over an AWGN channel in complex baseband form. For convenience this signal is rewritten here as

$$r(t) = \sum_{i=1}^M s_i(t) + n(t) = \sum_{i=1}^M \sqrt{S_i} m_i(t) PN_i(t) e^{j\theta_i} + n(t) \quad (\text{A.1})$$

where $m_i(t) = \sum_{k=-\infty}^{\infty} a_{ik} p(t - kT_b)$ is the data modulation of the i^{th} user with unit power rectangular pulse shape $p(t)$ of duration T_b (the bit time interval) and binary (± 1) data bit a_{ik} in the k^{th} transmission interval $kT_b \leq t \leq (k+1)T_b$, $n(t)$ is a complex white Gaussian noise process with power spectral density N_0 watts/Hz, and $PN_i(t)$ is the i^{th} user's ± 1 PN code waveform. In the zeroth transmission interval $0 \leq t \leq T_b$, Eq. (A.1) becomes

$$r(t) = \sum_{i=1}^M \sqrt{S_i} a_{i0} PN_i(t) e^{j\theta_i} + n(t) \quad (\text{A.2})$$

In deriving the decision metrics associated with the optimum and interference cancellation (IC) receivers, we shall be interested in the projection of (A.2) on the M PN code waveforms treated as a set of basis functions. As such, we define the normalized received vector¹ \mathbf{r} with components

$$r_i = \frac{1}{\sqrt{T_b}} \int_0^{T_b} r(t) PN_i(t) dt; \quad i = 1, 2, \dots, M \quad (\text{A.3})$$

and the normalized AWGN vector \mathbf{n} with components

$$n_i = \frac{1}{\sqrt{T_b}} \int_0^{T_b} n(t) PN_i(t) dt; \quad i = 1, 2, \dots, M \quad (\text{A.4})$$

The noise components $\{n_i\}$ are complex zero mean Gaussian random

¹All vectors defined will be assumed to be *column* vectors.

variables with variance $\sigma_n^2 = E\{|n_i|^2\} = (N_0/T_b) \int_0^{T_b} PN_i^2(t) dt = N_0$ and conditional crosscorrelation

$$E\{n_i n_j^* | \gamma_{ij}\} = N_0 \frac{1}{T_b} \int_0^{T_b} PN_i(t) PN_j(t) dt \triangleq N_0 \gamma_{ij}; \quad i, j = 1, 2, \dots, M \quad (\text{A.5})$$

where

$$\gamma_{ij} = \frac{1}{T_b} \int_0^{T_b} PN_i(t) PN_j(t) dt; \quad i, j = 1, 2, \dots, M \quad (\text{A.6})$$

are the components of the normalized PN crosscorrelation (user interference noise) matrix Γ .

To proceed further, we model the i^{th} PN waveform over the zeroth bit interval $0 \leq t \leq T_b$ as having been generated by a random binary (+1) data sequence $\{c_{ik}\}$ of rate $1/T_c$ (i.e., $\eta = T_b/T_c$ PN code chips per data bit). Thus, we can write

$$PN_i(t) = \sum_{k=1}^{\eta} c_{ik} p(t - kT_c) \quad (\text{A.7})$$

where $p(t)$ is again a unit power rectangular pulse shape now of duration T_c . Defining the user $M \times N_c$ PN code matrix $C = \{c_{ij}\}$ whose k^{th} row is the k^{th} users code sequence in the interval $0 \leq t \leq T_b$, i.e., $C_k^T = (c_{k1}, c_{k2}, \dots, c_{kN_c})$, then using (A.6). the PN crosscorrelation of (A.6) can be expressed as

$$\gamma_{ij} = \frac{1}{\eta} \sum_{k=1}^{\eta} c_{ik} c_{jk} = \sum_{k=1}^{\eta} \dot{c}_{ik} \dot{c}_{jk}, \quad \dot{c}_{jk} = \frac{c_{jk}}{\sqrt{\eta}}; \quad i, j = 1, 2, \dots, M \quad (\text{A.8})$$

and hence the user interference matrix Γ becomes

$$\Gamma = \frac{1}{\eta} CC^T = C' C'^T \quad (\text{A.9})$$

where $C' = \{\dot{c}_{ij}\}$ is the normalized user PN code matrix. Note that the diagonal elements of Γ all have value unity.

The noise components of (A.4) can also be expressed in terms of the user PN code matrix. Using (A.7) we have

$$n_i = \sum_{k=1}^{\eta} \frac{c_{ik}}{\sqrt{T_b}} \int_{(k-1)T_c}^{kT_c} n(t) dt = \sum_{k=1}^{\eta} c_{ik} N_k, \quad N_k = \frac{1}{\sqrt{T_b}} \int_{(k-1)T_c}^{kT_c} n(t) dt; \quad i = 1, 2, \dots, M \quad (\text{A.10})$$

The noise components $\{N_k\}$ are independent zero mean complex Gaussian random variables with variance $\sigma^2 = E\{|N_k|^2\} = N_0 T_c / T_b = N_0 / \eta$. Alternately, defining the normalized noise vector ζ with zero mean, variance two (unit variance per dimension) Gaussian components $\zeta_k \triangleq N_k / \sqrt{N_0 / 2\eta}$, then from (A.10) the normalized noise vector \mathbf{n} can be expressed as

$$\mathbf{n} = \sqrt{\frac{N_0}{2}} \mathbf{C}' \zeta \quad (\text{A.11})$$

Finally, defining the zeroth bit interval data vector \mathbf{a}_0 with components a_{i0} ; $i = 1, 2, \dots, M$, the complex carrier phase diagonal matrix

$$\Phi = \begin{bmatrix} e^{j\theta_1} & 0 & \cdot & \cdot & 0 \\ 0 & e^{j\theta_2} & 0 & \cdot & 0 \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & e^{j\theta_M} \end{bmatrix} \quad (\text{A.12})$$

and the scalar SNR diagonal matrix

$$\mathbf{B} = \begin{bmatrix} \sqrt{\frac{2E_{b1}}{N_0}} & 0 & \cdot & \cdot & 0 \\ 0 & \sqrt{\frac{2E_{b2}}{N_0}} & 0 & \cdot & 0 \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \sqrt{\frac{2E_{bM}}{N_0}} \end{bmatrix} \quad (\text{A.13})$$

with $E_{bi} \triangleq S_i T_b$, we can write (A.1) in the normalized vector form

$$\mathbf{y} = \mathbf{C}' \mathbf{C}'^T \Phi \mathbf{B} \mathbf{a}_0 + \mathbf{C}' \zeta \quad (\text{A.14})$$

where $\mathbf{y} \triangleq \mathbf{r} / \sqrt{N_0 / 2}$ is the normalized received vector. For the case where all users have the same power, then (A.14) simplifies to

$$\mathbf{y} = \sqrt{\frac{2E_b}{N_0}} \mathbf{C}' \mathbf{C}'^T \Phi \mathbf{a}_0 + \mathbf{C}' \zeta \quad (\text{A.15})$$

Appendix B

Decision Metrics for Multiuser Communication Receivers

I. Optimum (Maximum-Likelihood) Decision Metric – 1 Bit Observation Interval

In this section, we derive the optimum decision metric for the multiuser signal of Eq. (1) based on an observation of the received signal $r(t)$ over a single bit interval assumed, without loss in generality, to be the zeroth transmission interval $0 \leq t \leq T_b$. In particular, we derive the metric for making a decision on the data vector $\mathbf{a}_0 = (a_{10}, a_{20}, \dots, a_{M0})^T$ corresponding to the data bits transmitted by the M users in the interval $0 \leq t \leq T_b$. This metric is a special case of that derived by Verdu [1] who considered maximum-likelihood sequence estimation of multiuser signals corresponding to an observation interval greater than a single bit in duration. Although, one can immediately reduce the results in [1] to the special case of interest here, our purpose is to obtain the results in the form of the models and notation set up in Appendix A and also to use them as a motivation for deriving the interference cancellation (IC) metrics that follow in subsequent sections. With this in mind, we shall independently derive the needed results here in this section.

Using the well-known maximum-likelihood (ML) approach, the optimum decision rule for \mathbf{a}_0 based on an observation of $r(t)$ over the interval $0 \leq t \leq T_b$ is to choose the particular sequence $(a_{10}, a_{20}, \dots, a_{M0})$ that minimizes the squared Euclidean distance

$$\begin{aligned}
 d^2 &= \int_0^{T_b} \left| r(t) - \sum_{i=1}^M s_i(t) \right|^2 dt = \int_0^{T_b} \left| r(t) - \sum_{i=1}^M \sqrt{S_i} a_{i0} PN_i(t) e^{j\phi_i} \right|^2 dt \\
 &= \int_0^{T_b} |r(t)|^2 dt - 2 \operatorname{Re} \left\{ \int_0^{T_b} r(t) \sum_{i=1}^M \sqrt{S_i} a_{i0} PN_i(t) e^{-j\phi_i} dt \right\} \\
 &\quad + \int_0^{T_b} \left(\sum_{i=1}^M \sqrt{S_i} a_{i0} PN_i(t) e^{j\phi_i} \right) \left(\sum_{l=1}^M \sqrt{S_l} a_{l0} PN_l(t) e^{-j\phi_l} \right) dt
 \end{aligned} \tag{B.1}$$

Ignoring the first term in (B.1) since it does not depend on $(a_{10}, a_{20}, \dots, a_{M0})$ and dividing the remaining two terms by two, we get the equivalent decision, namely, choose the particular sequence $(a_{10}, a_{20}, \dots, a_{M0})$ that maximizes

$$\begin{aligned} \Omega = \operatorname{Re} \left\{ \int_0^{T_b} r(t) \sum_{i=1}^M \sqrt{S_i} a_{i0} P N_i(t) e^{-j\kappa t} dt \right\} \\ - \frac{1}{2} \int_0^{T_b} \left(\sum_{i=1}^M \sqrt{S_i} a_{i0} P N_i(t) e^{j\kappa t} \right) \left(\sum_{i=1}^M \sqrt{S_i} a_{i0} P N_i(t) e^{-j\kappa t} \right) dt \end{aligned} \quad (\text{B.2})$$

Rewriting (B.2) in terms of the notation introduced in Appendix A, we get

$$\begin{aligned} \Omega = \operatorname{Re} \left\{ \sum_{i=1}^M \sqrt{E_{bi}} a_{i0} r_i e^{-j\kappa t} \right\} - \frac{1}{2} \sum_{i=1}^M \sum_{l=1}^M \sqrt{E_{bi} E_{bl}} a_{i0} a_{l0} \gamma_{il} e^{j(\kappa - \kappa) t} \\ = \frac{N_0}{2} \left[\operatorname{Re} \left\{ \sum_{i=1}^M \sqrt{\frac{2E_{bi}}{N_0}} a_{i0} y_i e^{-j\kappa t} \right\} - \frac{1}{2} \sum_{i=1}^M \sum_{l=1}^M \sqrt{\frac{2E_{bi}}{N_0} \frac{2E_{bl}}{N_0}} a_{i0} a_{l0} \gamma_{il} e^{j(\kappa - \kappa) t} \right] \end{aligned} \quad (\text{B.3})$$

where as before $E_{bi} \triangleq S_i T_b$ is the i^{th} user's energy per bit. Alternately, in vector notation, we wish to maximize¹

$$\begin{aligned} \Omega = \operatorname{Re} \left\{ \mathbf{a}_0^T \mathbf{B} \Phi^* \mathbf{y} \right\} - \frac{1}{2} \mathbf{a}_0^T \mathbf{B} \Phi \Gamma \Phi^* \mathbf{B} \mathbf{a}_0 \\ = \operatorname{Re} \left\{ \mathbf{a}_0^T \mathbf{B} \left[\Phi^* \mathbf{y} - \frac{1}{2} \Phi \Gamma \Phi^* \mathbf{B} \mathbf{a}_0 \right] \right\} \end{aligned} \quad (\text{B.4})$$

where the matrices Φ and \mathbf{B} are defined in (A.12) and (A.13), respectively. Because of the symmetry of the Γ matrix [see (A.6) and (A.9)], we can rewrite (B.4) in the simpler form

$$\begin{aligned} \Omega = \operatorname{Re} \left\{ \mathbf{a}_0^T \mathbf{B} \Phi^* \mathbf{y} \right\} - \frac{1}{2} \mathbf{a}_0^T \mathbf{B} \Phi^* \Gamma \Phi \mathbf{B} \mathbf{a}_0 \\ = \operatorname{Re} \left\{ \mathbf{a}_0^T \mathbf{B} \Phi^* \left[\mathbf{y} - \frac{1}{2} \Gamma \Phi \mathbf{B} \mathbf{a}_0 \right] \right\} \end{aligned} \quad (\text{B.5})$$

Again for the case where all users have the same power, maximizing (B.5) is equivalent to maximizing

$$\Omega = \operatorname{Re} \left\{ \mathbf{a}_0^T \Phi^* \left[\mathbf{y} - \frac{1}{2} \sqrt{\frac{2E_b}{N_0}} \mathbf{C} \mathbf{C}^T \Phi \mathbf{a}_0 \right] \right\} \quad (\text{B.6})$$

II. Interference Cancellation (IC) Decision Metric Based Only on Observation of the Received Signal in a Single Bit Interval

Suppose now that instead of making a *joint* decision on the M data bits

¹We shall continue to use the notation Ω for all normalized forms of the decision metric.

transmitted in the zeroth transmission interval, we instead choose to *individually* decide on each user's data bit in this same interval. The motivation for doing this is to reduce the complexity of the metric derived in the previous section. Clearly, in deriving such an ML metric for any one user, one would theoretically have to exact knowledge of the $M - 1$ data bits corresponding to all the other users. Since indeed this information is unknown, the above theoretical assumption is practically invalid. Nevertheless, let us proceed to find the appropriate ML metric under this assumption after which we shall modify it to allow for a practical realization in terms of an IC decision algorithm. For simplicity of notation, we shall assume that user #1 is the user of interest, i.e., we seek an ML metric for deciding on data bit a_{10} .

If once again we were to start by trying to minimize (B.1), assuming now that the sequence $(a_{20}, a_{30}, \dots, a_{M0})$ is known a priori, then one would again arrive at the desire to maximize (B.3) with the terms that do not involve a_{10} eliminated. Eliminating such terms from (B.3) gives the simplified decision rule: Choose the particular a_{10} that maximizes

$$\Omega = \text{Re} \left\{ \sqrt{\frac{2E_{b1}}{N_0}} a_{10} y_1 e^{-j\theta} \right\} - \frac{1}{2} a_{10} \sqrt{\frac{2E_{b1}}{N_0}} \sum_{i=2}^M a_{i0} \gamma_{1i} e^{j(\theta - \phi_i)} - \frac{1}{2} a_{10} \sqrt{\frac{2E_{b1}}{N_0}} \sum_{i=2}^M a_{i0} \gamma_{ii} e^{j(\theta - \phi_i)} \quad (\text{B.7})$$

or, equivalently, since $\gamma_{1i} = \gamma_{ii}$

$$\Omega = \text{Re} \left\{ a_{10} e^{-j\theta} \left[y_1 - \sum_{i=2}^M \sqrt{\frac{2E_{b1}}{N_0}} a_{i0} \gamma_{1i} e^{j\theta} \right] \right\} \quad (\text{B.8})$$

Note that as predicted above, a decision on a_{10} based on maximizing (B.8) theoretically requires exact knowledge of the data sequence $(a_{20}, a_{30}, \dots, a_{M0})$. Before proceeding we rewrite (B.8) in the vector form

$$\Omega = \text{Re} \left\{ a_{10} e^{-j\theta} \left[y_1 - (\Gamma - \mathbf{I})_1 \Phi \mathbf{B} \mathbf{a}_0 \right] \right\} = \text{Re} \left\{ a_{10} e^{-j\theta} \left[y_1 - (\mathbf{C} \mathbf{C}^T - \mathbf{I})_1 \Phi \mathbf{B} \mathbf{a}_0 \right] \right\} \quad (\text{B.9})$$

where \mathbf{I} is an $M \times M$ identity matrix and the "1" subscript on $\Gamma - \mathbf{I}$ and $\mathbf{C} \mathbf{C}^T - \mathbf{I}$ denotes the first row of these matrices. Since for a single user communication system, the optimum ML decision metric for a_{10} would be to choose the particular a_{10} that maximizes

$$\Omega = \text{Re} \left\{ a_{10} e^{-j\theta} y_1 \right\} \quad (\text{B.10})$$

then the subtractive term $(C'C^T - I)_1 \Phi B a_0$ in (B.9) represents the receiver's attempt at cancelling the interference produced by the presence of the other $M - 1$ users.

If instead of user #1, the user of interest was user # k , then the decision on a_{k0} would analogously be made by maximizing

$$\Omega = \text{Re}\{a_{k0}e^{-j\theta_k}[y_k - (\Gamma - I)_k \Phi B a_0]\} = \text{Re}\{a_{k0}e^{-j\theta_k}[y_k - (C'C^T - I)_k \Phi B a_0]\} \quad (\text{B.11})$$

Here again this decision on a_{k0} would require exact knowledge of the $M - 1$ data bits $a_{i0}; i \neq k$.

Since a_{10} in (B.9) is binary (± 1) valued, then the decision rule based on (B.9) can be formulated as a comparison of $d^2|_{a_0=1}$ with a zero threshold, or equivalently,

$$\hat{a}_{10} = \text{sgn}\left\{\text{Re}\left\{e^{-j\theta_1}\left[y_1 - (C'C^T - I)_1 \Phi B a_0\right]\right\}\right\} \quad (\text{B.12})$$

where \hat{a}_{10} denotes the decision on a_{10} and "sgn" denotes the signum function. An analogous decision statistic based on (B.11) would exist for each of the other user data bits in the zeroth transmission interval. Combining all these decisions into a vector \hat{a}_0 we have the final desired result

$$\hat{a}_0 = \text{sgn}\left\{\text{Re}\left\{\Phi^*\left[y - (C'C^T - I)\Phi B a_0\right]\right\}\right\} \quad (\text{B.13})$$

where it is understood that the signum function is applied term by term in the vector. Note that although the complete data vector $a_0 = (a_{10}, a_{20}, \dots, a_{M0})^T$ appears on the right hand side of (B.13), because of the subtraction of the identity matrix from $C'C^T$, the decision \hat{a}_{i0} , i.e., the i th component of \hat{a}_0 , is, in reality, only affected by the vector $a_0 = (a_{10}, a_{20}, \dots, a_{i-1,0}, a_{i+1,0}, \dots, a_{M0})^T$.

In order to transform (B.13) into a practically implementable decision statistic, we must overcome the lack of exact knowledge of a_0 in so far as deciding on a_{i0} . One method of approximating this knowledge is to apply an *iterative* (in time) approach as follows. Suppose that at the first stage of the iteration we replace a_0 on the right hand side of (B.13) by an estimate of it, namely, $\hat{a}_0(0)$, obtained from the solution to the single user communication problem. In particular, analogous to (B.10), we would have

$$\hat{\mathbf{a}}_0(0) = \text{sgn}\{\Phi^* \mathbf{y}\} \quad (\text{B.14})$$

where we have introduced the parenthetical notation “(k-1)” to denote the values of the corresponding variables at the beginning of the k^{th} stage of iteration. (In terms of the present discussion, we have $k = 1$.) We refer to $\hat{\mathbf{a}}_0(0)$ as a “tentative decision” vector.² Note that each component of $\hat{\mathbf{a}}_0(0)$ requires no knowledge of the remaining user data bits and as such represents a decision made in the presence of the *total* interference from all other users. Needless to say these decisions will be poor relative to our final goal and it is for that reason that we refer to them as *tentative* decisions. Using $\hat{\mathbf{a}}_0(0)$ as a replacement for \mathbf{a}_0 on the right hand side of (B.13) we obtain after one stage of iteration the decision statistic³

$$\hat{\mathbf{a}}_0(1) = \text{sgn}\left\{\text{Re}\left\{\Phi^* \left[\mathbf{y} - (\mathbf{C}'\mathbf{C}'^T - \mathbf{I})\Phi\mathbf{B}\hat{\mathbf{a}}_0(0)\right]\right\}\right\} \quad (\text{B.15})$$

Continuing this iteration for more than one stage will further improve our final decision on \mathbf{a}_0 since now the tentative decision vector for the second stage will be $\hat{\mathbf{a}}_0(1)$ which has the benefit of one stage of interference cancellation. Thus, after k stages of iteration we obtain the final decision⁴

$$\hat{\mathbf{a}}_0(k) = \text{sgn}\left\{\text{Re}\left\{\Phi^* \left[\mathbf{y} - (\mathbf{C}'\mathbf{C}'^T - \mathbf{I})\Phi\mathbf{B}\hat{\mathbf{a}}_0(k-1)\right]\right\}\right\} \quad (\text{B.16})$$

It is important to understand that the iterative decision statistic described by (B.16) or any variation thereof is not optimum in the ML sense and thus its performance will be inferior relative to that corresponding to the ML approach taken in Section I of this appendix. Nonetheless, as will be demonstrated in the main text, by proper design, iterative IC receivers based on the notions leading up to (B.16) can be made to yield performance comparable with the optimum ML receiver as suggested by Verdu *but with considerably reduced implementation complexity*.

²Shortly, we shall modify this “tentative decision” to allow for a more general form of decision device.

³In reality, \mathbf{y} here corresponds to the normalized received vector in the zeroth transmission interval delayed by one bit interval. Since, however, a delay of one or more bit intervals does not affect the evaluation of \mathbf{y} (as given by (A.15)), we shall continue to denote by \mathbf{y} this normalized received vector at the k^{th} stage of iteration, i.e., \mathbf{y} delayed by k bit intervals.

⁴Again, we shall soon allow for modifying the tentative decisions made at *all* stages of interference cancellation prior to the final decision in terms of a more general form of decision device.

Returning to the notion of the tentative decisions made at each stage of interference cancellation, one should not conclude that a hard decision device will necessarily yield the best ultimate (after say K stages of IC cancellation) decision on \mathbf{a}_0 . Indeed, it will be shown in the main text that other forms of zero memory nonlinearity⁵ for a tentative decision device can produce improved performance relative to the signum function. In fact, even a linear tentative decision has certain merits as will be discussed in the main text. In view of this, we shall generalize our iterative IC cancellation decision statistic to allow for an arbitrary tentative decision function which we shall denote as $f(x)$. Thus, in this more general context, (B.16) becomes⁶

$$\hat{\mathbf{a}}_0(k) = f\left(\text{Re}\left\{\Phi^*\left[\mathbf{y} - (\mathbf{C}\mathbf{C}^T - \mathbf{I})\Phi\mathbf{B}\hat{\mathbf{a}}_0(k-1)\right]\right\}\right) \triangleq f(\tilde{\mathbf{a}}_0(k)) \quad (\text{B.17})$$

or for the case where all users have equal power

$$\hat{\mathbf{a}}_0(k) = f\left(\text{Re}\left\{\Phi^*\left[\mathbf{y} - \sqrt{\frac{2E_b}{N_0}}(\mathbf{C}\mathbf{C}^T - \mathbf{I})\Phi\hat{\mathbf{a}}_0(k-1)\right]\right\}\right) \triangleq f(\tilde{\mathbf{a}}_0(k)) \quad (\text{B.18})$$

It is understood, of course, that for the final decision (after say K stages of iteration) the signum function should be used as the decision device. Also, $\tilde{\mathbf{a}}_0(k)$ represents the vector of tentative *soft* decisions at the k^{th} stage of iteration. The decision metric in (B.18) suggests the implementation of an IC receiver as illustrated in Figure 1 together with Figure 2a. We refer to this structure as a *brute force interference canceller* since at each stage of the iteration an attempt is made to completely cancel the multiuser interference.

III. Interference Cancellation (IC) Decision Metric Based on Joint Observation of the Received Signal and the Tentative Soft Decision in a Single Bit Interval

The decision metric derived in Section II as described by (B.17) implies that at each stage (iteration) the receiver attempts to entirely cancel the full amount of multiuser interference. Since in the early stages of interference cancellation, the tentative decisions are less reliable than they are in later stages, it is not intuitively clear that the above philosophy of entirely cancelling the full

⁵In the next section of this appendix where the IC metric is generalized still further to include additional observations, we shall derive a specific form of this nonlinearity based upon nonlinear estimation considerations for $\hat{\mathbf{a}}_0(k)$.

⁶Here again we use the notation $f(\mathbf{x})$ to denote the application of the function " f " to each component of the vector \mathbf{x} .

amount of interference at each iteration stage necessarily leads to the best suboptimum decision metric. Rather, a better philosophy is one which in the early stages cancels only a fraction of the multiuser interference with the amount being cancelled increasing as one continues to iterate toward the ultimate final data decisions, i.e., as the fidelity of the tentative decisions improves. To see how a metric motivated by such a philosophy can come about using ML-type considerations, we proceed as follows.

Analogous to Section II, we shall take an iterative approach to a suboptimum decision statistic derived from now *joint* ML-type considerations. In particular, consider first the component of the normalized received signal vector \mathbf{y} corresponding to user #1 which when multiplied by $e^{-j\theta}$ is obtained from Eq. (A.15) as

$$e^{-j\theta} y_1 = \sqrt{\frac{2E_{b1}}{N_0}} a_{10} + \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} a_{i0} e^{j(\theta-\theta_i)} \gamma_{1i} + n_1 e^{-j\theta} \quad (\text{B.19})$$

Taking the real part of (B.19) gives an expression of the form

$$Y_1 = \sqrt{\frac{2E_{b1}}{N_0}} a_{10} + I_1 + W_1 \quad (\text{B.20})$$

Here I_1 denotes the real part of the multiuser interference experienced by user #1 due to the remaining $M - 1$ users and W_1 is a zero mean Gaussian random variable with variance $E\{W_1^2\} = \sigma_1^2$. Next, consider the component of the vector of tentative soft decisions at the k^{th} stage corresponding to user #1. From (B.17) together with (B.19) we have

$$\begin{aligned} \tilde{a}_{10}(k-1) &= \text{Re} \left\{ e^{-j\theta} y_1 - \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} \hat{a}_{i0}(k-2) e^{j(\theta-\theta_i)} \gamma_{1i} \right\} \\ &= \text{Re} \left\{ \sqrt{\frac{2E_{b1}}{N_0}} a_{10} + \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} (a_{i0} - \hat{a}_{i0}(k-2)) e^{j(\theta-\theta_i)} \gamma_{1i} + n_1 e^{-j\theta} \right\} \\ &= \sqrt{\frac{2E_{b1}}{N_0}} a_{10} + \text{Re} \left\{ \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} (a_{i0} - \hat{a}_{i0}(k-2)) e^{j(\theta-\theta_i)} \gamma_{1i} \right\} + W_1 \end{aligned} \quad (\text{B.21})$$

For the purpose of what follows we can model the middle term of (B.21), i.e., the residual (uncancelled) interference, as a zero mean Gaussian random variable which when combined with W_1 results in a zero mean Gaussian random variable $W_2(k-1)$ with variance $E\{W_2^2(k-1)\} = \sigma_{2k}^2$. Note that the variance σ_{2k}^2 depends on the iteration stage k . In view of the previous

discussion, we rewrite (B.21) in the form

$$\tilde{a}_{10}(k-1) = \sqrt{\frac{2E_b}{N_0}} a_{10} + W_2(k-1) \quad (\text{B.22})$$

Clearly, W_1 and $W_2(k-1)$ are correlated, i.e., $E\{W_1 W_2(k-1)\} \triangleq \rho_k \sigma_1 \sigma_{2k}$. As we shall see shortly, it is not necessary to be able to specifically evaluate σ_1 , σ_{2k} and ρ_k . Rather a specific combination of these parameters will be used to define a parameter p_k which shall have significance in terms of the amount of interference for which cancellation is attempted at each stage of the receiver.

Using (B.20) and (B.22), a ML estimate of a_{10} based on joint observation of Y_1 and $\tilde{a}_{10}(k-1)$ is obtained by choosing the value of a_{10} that maximizes the conditional probability density function (pdf)

$$p(Y_1, \tilde{a}_{10}(k-1) | a_{10}, I_1) = \frac{1}{2\pi\sigma_1\sigma_{2k}\sqrt{1-\rho_k^2}} \exp \left\{ - \frac{\sigma_{2k}^2 \left(Y_1 - \sqrt{\frac{2E_b}{N_0}} a_{10} - I_1 \right)^2 + \sigma_1^2 \left(\tilde{a}_{10}(k-1) - \sqrt{\frac{2E_b}{N_0}} a_{10} \right)^2}{2\sigma_1^2\sigma_{2k}^2(1-\rho_k^2)} \right. \\ \left. - \frac{2\rho_k\sigma_1\sigma_{2k} \left(Y_1 - \sqrt{\frac{2E_b}{N_0}} a_{10} - I_1 \right) \left(\tilde{a}_{10}(k-1) - \sqrt{\frac{2E_b}{N_0}} a_{10} \right)}{2\sigma_1^2\sigma_{2k}^2(1-\rho_k^2)} \right\} \quad (\text{B.23})$$

Since from (B.19) the interference term I_1 is a function of the data sequence $(a_{20}, a_{30}, \dots, a_{M0})$ which is unknown, we proceed as before by replacing I_1 with

$$\hat{I}_1(k) \triangleq \text{Re} \left\{ \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} \hat{a}_{i0}(k-1) e^{j(\theta_i - \theta)} \gamma_{1i} \right\} \quad (\text{B.24})$$

which is synonymous with replacing the unknown sequence $(a_{20}, a_{30}, \dots, a_{M0})$ with the k^{th} stage estimate of it $(\hat{a}_{20}(k-1), \hat{a}_{30}(k-1), \dots, \hat{a}_{M0}(k-1))$. Substituting $\hat{I}_1(k)$ for I_1 in (B.23) and simplifying yields

$$\begin{aligned}
p(Y_1, \bar{a}_{10}(k-1) | \hat{I}_1(k)) &= C \exp \left\{ a_{10} \sqrt{\frac{2E_{b1}}{N_0}} \left[\frac{\sigma_{2k}^2 (Y_1 - \hat{I}_1(k)) + \sigma_1^2 \bar{a}_{10}(k-1) - \rho_k \sigma_1 \sigma_{2k} (Y_1 - \hat{I}_1(k) + \bar{a}_{10}(k-1))}{\sigma_1^2 \sigma_{2k}^2 (1 - \rho_k^2)} \right] \right\} \\
&= C \exp \left\{ a_{10} \sqrt{\frac{2E_{b1}}{N_0}} \left[\frac{(\sigma_{2k}^2 - \rho_k \sigma_1 \sigma_{2k})(Y_1 - \hat{I}_1(k)) + (\sigma_1^2 - \rho_k \sigma_1 \sigma_{2k}) \bar{a}_{10}(k-1)}{\sigma_1^2 \sigma_{2k}^2 (1 - \rho_k^2)} \right] \right\}
\end{aligned} \tag{B.25}$$

where the constant C includes terms that do not depend on a_{10} . We now introduce a normalization so as to make the coefficients of $Y_1 - \hat{I}_1(k)$ and $\bar{a}_{10}(k-1)$ in (B.25) sum to unity. Letting⁷

$$p_k \triangleq \frac{\sigma_{2k}^2 - \rho_k \sigma_1 \sigma_{2k}}{\sigma_1^2 + \sigma_{2k}^2 - 2\rho_k \sigma_1 \sigma_{2k}}, \quad 1 - p_k = \frac{\sigma_1^2 - \rho_k \sigma_1 \sigma_{2k}}{\sigma_1^2 + \sigma_{2k}^2 - 2\rho_k \sigma_1 \sigma_{2k}} \tag{B.26}$$

then (B.25) can be written in the desired form

$$p(Y_1, \bar{a}_{10}(k-1) | a_{10}, \hat{I}_1(k)) = C \exp \left\{ a_{10} \sqrt{\frac{2E_{b1}}{N_0}} \left[\frac{\sigma_1^2 + \sigma_{2k}^2 - 2\rho_k \sigma_1 \sigma_{2k}}{\sigma_1^2 \sigma_{2k}^2 (1 - \rho_k^2)} \right] \left[p_k (Y_1 - \hat{I}_1(k)) + (1 - p_k) \bar{a}_{10}(k-1) \right] \right\} \tag{B.27}$$

Since the natural logarithm is a monotonic function of its argument, then taking the natural logarithm of (B.27), we see that the ML value of a_{10} at the k^{th} stage of iteration is simply given by

$$\begin{aligned}
\hat{a}_{10}(k) &= \text{sgn} \left\{ \left[p_k (Y_1 - \hat{I}_1(k)) + (1 - p_k) \bar{a}_{10}(k-1) \right] \right\} \\
&= \text{sgn} \left\{ \text{Re} \left[p_k \left(e^{-j\theta} y_1 - \sum_{i=2}^M \sqrt{\frac{2E_{bi}}{N_0}} \hat{a}_{i0}(k-1) e^{j(\theta - \theta_i)} \gamma_{li} \right) \right] + (1 - p_k) \bar{a}_{10}(k-1) \right\}
\end{aligned} \tag{B.28}$$

In terms of the entire vector of decisions for all users in the zeroth transmission interval, we have analogous to (B.16)

⁷Note that based on its definition, the parameter p_k is not necessarily restricted to lie in the range $0 \leq p_k \leq 1$. However, if $\rho_k < \min\left(\frac{\sigma_1}{\sigma_{2k}}, \frac{\sigma_{2k}}{\sigma_1}\right)$ then it can be shown that this restriction is valid. This implies either $\sigma_1 \leq \sigma_{2k}$ or $\sigma_{2k} < \sigma_1$. Intuition, however, would suggest the former.

$$\hat{a}_0(k) = \text{sgn} \left\{ \text{Re} \left\{ p_k \Phi^* \left[y - (C C^T - I) \Phi \hat{a}_0(k-1) \right] \right\} + (1-p_k) \tilde{a}_0(k-1) \right\} \quad (\text{B.29})$$

In applying (B.29), the parameter p_k is to be optimized at each stage of the iteration. The decision metric of (B.29) suggests⁸ the implementation of an IC receiver as illustrated in Figure 1 together with Figure 2b.

The iterative joint ML approach that leads to (B.29) produces a decision metric in which the tentative decisions at each stage of the iteration are hard decisions. We now consider a modification of this approach which allows for other than hard tentative decisions. (Of course, the final decision on a_0 will, for uncoded data, always be a hard decision.) In particular, we consider using for $\hat{a}_0(k)$ the nonlinear estimate $E\{a_0|Y_1, \tilde{a}_0(k-1), I_1\}$ which is given by

$$E\{a_0|Y_1, \tilde{a}_0(k-1), I_1\} = (1)p(1|Y_1, \tilde{a}_0(k-1), I_1) + (-1)p(-1|Y_1, \tilde{a}_0(k-1), I_1) \quad (\text{B.30})$$

with $p(a_0|Y_1, \tilde{a}_0(k-1), I_1)$ the a posteriori probability of the user #1's bit in the zeroth transmission interval given the observations and interference. Using Bayes' rule, this probability can be determined in terms of the conditional probability $p(Y_1, \tilde{a}_0(k-1)|a_0, I_1)$ of (B.23) as

$$p(a_0|Y_1, \tilde{a}_0(k-1), I_1) = \frac{\Lambda(a_0)q(a_0)}{\Lambda(1)q(1) + \Lambda(-1)q(-1)} \quad (\text{B.31})$$

where

$$\Lambda(a_0) \triangleq p(Y_1, \tilde{a}_0(k-1)|a_0, I_1), \quad q(a_0) = p(a_0, I_1) \quad (\text{B.32})$$

Evaluating (B.31) at $a_0 = 1$ and $a_0 = -1$ and substituting the results into (B.30) gives

$$E\{a_0|Y_1, \tilde{a}_0(k-1), I_1\} = \frac{\Lambda(1)q(1) - \Lambda(-1)q(-1)}{\Lambda(1)q(1) + \Lambda(-1)q(-1)} \quad (\text{B.33})$$

As before, evaluation of (B.33) ideally requires knowledge of the interference I_1 which in turn requires that user #1 have knowledge of the sequence of data bits corresponding to the other users, i.e., $(a_{20}, a_{30}, \dots, a_{M0})$. In the absence of this exact knowledge we do as before and replace I_1 by the estimate $\hat{I}_1(k)$ as given by (B.24). As such we redefine the functions in (B.32) by

⁸As in Section II, we replace the *hard* tentative decisions with an arbitrary zero memory nonlinearity $f(\mathbf{x})$.

$$\Lambda(a_{10}) \triangleq p(Y_1, \tilde{a}_{10}(k-1) | a_{10}, \hat{I}_1(k)), \quad q(a_{10}) = p(a_{10}, \hat{I}_1(k)) \quad (\text{B.34})$$

Because of the symmetry of the problem, i.e., the equiprobable properties of the data streams, we have that $q(1) = q(-1)$. Hence, (B.33) becomes

$$\hat{a}_{10}(k) \triangleq E\{a_{10} | Y_1, \tilde{a}_{10}(k-1), \hat{I}_1(k)\} = \frac{\Lambda(1) - \Lambda(-1)}{\Lambda(1) + \Lambda(-1)} \quad (\text{B.35})$$

Referring to (B.27) for the evaluation of $\Lambda(a_{10})$ we have that

$$\Lambda(a_{10}) = C \exp(a_{10} \alpha_k \lambda_k) \quad (\text{B.36})$$

where

$$\alpha_k \triangleq \sqrt{\frac{E_{b1}}{N_0} \left[\frac{\sigma_1^2 + \sigma_{2k}^2 - 2\rho_k \sigma_1 \sigma_{2k}}{\sigma_1^2 \sigma_{2k}^2 (1 - \rho_k^2)} \right]} \quad (\text{B.37})$$

$$\lambda_k \triangleq p_k (Y_1 - \hat{I}_1(k)) + (1 - p_k) \tilde{a}_{10}(k-1)$$

Finally, substituting (B.36) into (B.35) we obtained the desired result as

$$\begin{aligned} \hat{a}_{10}(k) &= \frac{C \exp(\alpha_k \lambda_k) - C \exp(-\alpha_k \lambda_k)}{C \exp(\alpha_k \lambda_k) + C \exp(-\alpha_k \lambda_k)} = \tanh(\alpha_k \lambda_k) \\ &= \tanh \left\{ \sqrt{\frac{2E_{b1}}{N_0} \left[\frac{\sigma_1^2 + \sigma_{2k}^2 - 2\rho_k \sigma_1 \sigma_{2k}}{\sigma_1^2 \sigma_{2k}^2 (1 - \rho_k^2)} \right]} \left[p_k (Y_1 - \hat{I}_1(k)) + (1 - p_k) \tilde{a}_{10}(k-1) \right] \right\} \end{aligned} \quad (\text{B.38})$$

Comparing (B.38) with (B.28) we see that the hard tentative decisions have been replaced by soft tentative decisions in the form of hyperbolic tangent functions. Furthermore, the slope of these functions (which is proportional to α_k) is another parameter to be optimized at each stage of the iteration.

Appendix C

The Special Case of Linear Interference Cancellation

We consider here the special case of the interference cancellation methods discussed in Sections II and III of Appendix B when the tentative decision devices are all *linear*, i.e., the tentative decisions are infinite bit soft decisions. In particular, we shall be interested in comparing the limiting (as the number of iterations approaches infinity) performance of such schemes with other linear IC cancellation methods such as those considered by Verdu [4].

For the brute force IC scheme as described in Section II of Appendix B, if we linearize the tentative decision devices, then (B.18) becomes¹

$$\begin{aligned}\hat{\mathbf{a}}_0(k) &= \left(\sqrt{\frac{2E_b}{N_0}} \right)^{-1} \left[\mathbf{y} - \sqrt{\frac{2E_b}{N_0}} (\mathbf{C}'\mathbf{C}'^T - \mathbf{I})\hat{\mathbf{a}}_0(k-1) \right] = \left(\sqrt{\frac{2E_b}{N_0}} \right)^{-1} \mathbf{y} - (\mathbf{\Gamma} - \mathbf{I})\hat{\mathbf{a}}_0(k-1) \quad (\text{C.1}) \\ &= \mathbf{y}' - (\mathbf{\Gamma} - \mathbf{I})\hat{\mathbf{a}}_0(k-1)\end{aligned}$$

where \mathbf{y}' denotes further normalization of the normalized received vector \mathbf{y} . This equation when iterated can be expressed in the form

$$\hat{\mathbf{a}}_0(k+l) = \left[\mathbf{I} + \sum_{j=1}^{l-1} (-1)^j (\mathbf{\Gamma} - \mathbf{I})^j \right] \mathbf{y}' + (-1)^l (\mathbf{\Gamma} - \mathbf{I})^l \hat{\mathbf{a}}_0(k) \quad (\text{C.2})$$

In the limit as $l \rightarrow \infty$, (C.2) becomes

$$\hat{\mathbf{a}}_0 = [\mathbf{I} + \mathbf{\Gamma} - \mathbf{I}]^{-1} \mathbf{y}' = \mathbf{\Gamma}^{-1} \mathbf{y}' \quad (\text{C.3})$$

assuming of course that the inverse correlation matrix $\mathbf{\Gamma}^{-1}$ exists (more about this constraint shortly.) It is straightforward to show that the estimator in (C.3) is also the minimum mean-square error (MSE) linear estimator of the zeroth transmission interval data vector \mathbf{a}_0 , i.e., if the normalized observation \mathbf{y} is

$$\text{given by [see (A.15)] } \mathbf{y} = \sqrt{\frac{2E_b}{N_0}} \mathbf{C}'\mathbf{C}'^T \mathbf{a}_0 + \mathbf{C}'\boldsymbol{\zeta} \text{ or equivalently } \mathbf{y}' = \mathbf{\Gamma} \mathbf{a}_0 + \left(\sqrt{\frac{2E_b}{N_0}} \right)^{-1} \mathbf{C}'\boldsymbol{\zeta},$$

¹For simplicity of the presentation, we shall assume the case where all users have equal power and also ignore the effects of the user carrier phases. As such, the $\text{Re}\{\}$ notation in (B.18) can be ignored since the normalized received vector \mathbf{y} is now real. Also, under these assumptions, the tentative decision transfer function becomes $f(x) = \left(\sqrt{2E_b/N_0} \right)^{-1} x$.

then for $\hat{\mathbf{a}}_0$ of the form $\hat{\mathbf{a}}_0 = \mathbf{A}\mathbf{y}$, the choice of $\mathbf{A} = \mathbf{\Gamma}^{-1}$ minimizes $E\{\|\mathbf{a}_0 - \hat{\mathbf{a}}_0\|^2\}$. Thus, we observe that the brute force IC scheme with linear tentative decisions asymptotically (as the number of stages becomes infinite) approaches the best (in the mean-square sense) linear estimate of the data vector.

Since the final decision must be a hard decision, then (C.3) is replaced by

$$\hat{\mathbf{a}}_0 = \text{sgn}\{\mathbf{\Gamma}^{-1}\mathbf{y}\} \quad (\text{C.4})$$

In [4], the authors recognize the fact that $\mathbf{\Gamma}^{-1}$ may not exist and as such propose a decision algorithm based on (C.4) where $\mathbf{\Gamma}^{-1}$ is replaced by a member of the set of *generalized inverses* of the crosscorrelation matrix $\mathbf{\Gamma}$. The resulting detector is referred to as a *decorrelating detector*. Since the bracketed term in (C.2), namely, $\mathbf{I} + \sum_{j=1}^{l-1} (-1)^j (\mathbf{\Gamma} - \mathbf{I})^j$, always exists for finite l even when $\mathbf{\Gamma}^{-1}$ may not exist, we can view the brute force IC scheme with linear tentative decisions and a finite number of stages as an approximation to the linear MSE estimator of (C.3).

For the generalized IC scheme discussed in Section III of Appendix B, if we once again linearize the tentative decisions then (B.29b) becomes

$$\begin{aligned} \hat{\mathbf{a}}_0(k) &= \left(\frac{2E_b}{N_0}\right)^{-1} \left[\mathbf{y} - \sqrt{\frac{2E_b}{N_0}} (\mathbf{C}'\mathbf{C}'^T - \mathbf{I}) [p_k \hat{\mathbf{a}}_0(k-1) + (1-p_k) \hat{\mathbf{a}}_0(k-2)] \right] \\ &= \left(\frac{2E_b}{N_0}\right)^{-1} \mathbf{y} - (\mathbf{\Gamma} - \mathbf{I}) [p_k \hat{\mathbf{a}}_0(k-1) + (1-p_k) \hat{\mathbf{a}}_0(k-2)] \\ &= \mathbf{y} - (\mathbf{\Gamma} - \mathbf{I}) [p_k \hat{\mathbf{a}}_0(k-1) + (1-p_k) \hat{\mathbf{a}}_0(k-2)] \end{aligned} \quad (\text{C.5})$$

When iterated, (C.5) can be expressed in the form

$$\hat{\mathbf{a}}_0(k+2l) = \left[\mathbf{I} + \sum_{j=1}^{l-1} (-1)^j (\mathbf{\Gamma} - \mathbf{I})^j \right] \mathbf{y} + (-1)^l (\mathbf{\Gamma} - \mathbf{I})^l \sum_{j=0}^l \alpha_j \hat{\mathbf{a}}_0(k+j) \quad (\text{C.6})$$

where the α_j 's are coefficients that nonlinearly depend on the set $p_{k+2}, p_{k+3}, \dots, p_{k+2l}$. In the limit as $l \rightarrow \infty$, (C.6) also becomes equal to (C.3) independent of the values of the p_k 's. Thus, the generalized IC scheme with linear tentative decisions and a finite number of stages can be viewed as yet another approximation to the linear MSE estimator of (C.3). The difference between how (C.2) and (C.6) approximate the linear MSE solution is the manner in which they converge toward this asymptotic solution as $l \rightarrow \infty$, i.e., for any fixed finite l ,

with proper choice of the p_k 's, the generalized IC scheme will offer the better performance of the two.

Appendix D

Mean and Variance of Normalized I&D Output for Single Stage Linear Interference Cancellation

The real part of the normalized output x_{11} of (32), namely, x_{R_1} , consists of three terms. The first is the desired signal component which has mean (conditioned on $a_{10} = 1$)

$$E\left\{\left(1 - \sum_{i=2}^M \gamma_{1i}^2\right)\sqrt{E_b}\right\} = \sqrt{E_b}\left(1 - \frac{M-1}{\eta}\right) \quad (\text{D.1})$$

and variance

$$\begin{aligned} E\left\{\left[\left(1 - \sum_{i=2}^M \gamma_{1i}^2\right)\sqrt{E_b} - \sqrt{E_b}\left(1 - \frac{M-1}{\eta}\right)\right]^2\right\} &= E\left\{\left[\left(\frac{M-1}{\eta}\right) - \sum_{i=2}^M \gamma_{1i}^2\right]^2\right\} E_b \\ &= E_b\left[\left(\frac{M-1}{\eta}\right)^2 - 2\left(\frac{M-1}{\eta}\right)E\left\{\sum_{i=2}^M \gamma_{1i}^2\right\} + E\left\{\sum_{i=2}^M \sum_{j=2}^M \gamma_{1i}^2 \gamma_{1j}^2\right\}\right] \\ &= E_b\left[(M-1)E\{\gamma_{1i}^4\} - \frac{M-1}{\eta^2}\right] \end{aligned} \quad (\text{D.2})$$

Since for purely random PN codes, we have $E\{\gamma_{1i}^4\} = 3/\eta^2 - 2/\eta^3$ (see Appendix E). then

$$E\left\{\left[\left(1 - \sum_{i=2}^M \gamma_{1i}^2\right)\sqrt{E_b} - \sqrt{E_b}\left(1 - \frac{M-1}{\eta}\right)\right]^2\right\} = E_b\left[\frac{2(M-1)}{\eta^2}\left(1 - \frac{1}{\eta}\right)\right] \quad (\text{D.3})$$

The middle term in (32) represents the residual user interference after cancellation which has zero mean and variance (see Appendix E)

$$\begin{aligned}
\frac{1}{2} E \left\{ \left| \sqrt{E_b} \sum_{i=2}^M a_{0i} e^{j(\phi_i - \phi_1)} \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{ik} \gamma_{1k} \right|^2 \right\} &= \frac{E_b}{2} \sum_{i=2}^M \sum_{j=2}^M E \{ a_{0i} a_{0j} \} e^{j(\phi_i - \phi_j)} \sum_{\substack{k=2 \\ k \neq i}}^M \sum_{\substack{l=2 \\ l \neq j}}^M E \{ \gamma_{ik} \gamma_{1k} \gamma_{jl} \gamma_{1l} \} \\
&= \frac{E_b}{2} \sum_{i=2}^M \sum_{k=2}^M \sum_{\substack{l=2 \\ l \neq i}}^M E \{ \gamma_{ik} \gamma_{1k} \gamma_{il} \gamma_{1l} \} \\
&= \frac{E_b}{2} \frac{(M-1)(M-2)}{\eta^2} + \frac{E_b}{2} \frac{(M-1)(M-2)(M-3)}{\eta^3}
\end{aligned} \tag{D.4}$$

The equivalent thermal noise has zero mean and variance

$$\begin{aligned}
\frac{1}{2} E \left\{ \left| n_1 - \sum_{i=2}^M \gamma_{1i} n_i \right|^2 \right\} &= \frac{1}{2} E \{ |n_1|^2 \} - \operatorname{Re} \left\{ E \left\{ \sum_{i=2}^M n_i n_i^* \gamma_{1i} \right\} \right\} + \frac{1}{2} E \left\{ \sum_{i=2}^M |n_i|^2 \gamma_{1i}^2 \right\} + \frac{1}{2} E \left\{ \sum_{\substack{i=2 \\ i \neq j}}^M \sum_{j=2}^M n_i n_j^* \gamma_{1i} \gamma_{1j} \right\} \\
&= \frac{N_0}{2} - N_0 \sum_{i=2}^M E \{ \gamma_{1i}^2 \} + \frac{N_0}{2} \sum_{i=2}^M E \{ \gamma_{1i}^2 \} + \frac{N_0}{2} \sum_{\substack{i=2 \\ i \neq j}}^M \sum_{j=2}^M E \{ \gamma_{ij} \gamma_{1i} \gamma_{1j} \} \\
&= \frac{N_0}{2} \left[1 - \frac{M-1}{\eta} + \frac{(M-1)(M-2)}{\eta^2} \right]
\end{aligned} \tag{D.5}$$

Finally, then the mean of x_{R_1} (conditioned on $a_{10} = 1$) is given by (D.1), i.e.,

$$E \{ x_{R_1} | a_{10} = 1 \} = \sqrt{E_b} \left(1 - \frac{M-1}{\eta} \right) \tag{D.6}$$

and the variance of x_{R_1} is given by the sum of (D.3), (D.4) and (D.5), namely,

$$\sigma_{11}^2 = \frac{N_0}{2} \left[1 - \frac{M-1}{\eta} + \frac{(M-1)(M-2)}{\eta^2} \right] + \frac{E_b}{2} \frac{(M-1)}{\eta^2} \left[M+2 + \frac{M^2 - 5M + 2}{\eta} \right] \tag{D.7}$$

Although in principle the above exact analysis can be applied to a multiple stage linear IC scheme, the difficulty associated with evaluating the expectation of higher order products of the γ_{ij} 's makes this approach impractical. Thus, to allow for analytical evaluation in the multiple stage case, we shall need to apply some simplification with regard to the statistical characterization of the γ_{ij} 's, in particular, for large values of η it is reasonable to approximate them as independent zero mean Gaussian random variables each with variance $1/\eta$. The argument for justifying this assumption is the same as that given in the discussion following Eq. (12) of the main body of the

paper which relates to an IC scheme with nonlinear (hard-limiting) tentative decisions. Thus, while we do not need to invoke this Gaussian assumption for the single stage linear IC scheme, i.e., the results in (D.1) – (D.7) are exact for random PN codes assigned to the users, we shall nonetheless recompute these results based on the above assumption to allow their similarity to be used as a justification for employing the latter in the multiple stage linear IC analysis.

Since evaluation of (D.1) depends only on the variance γ_{1i} , this result remains unchanged. In going from (D.2) to (D.3), the only difference lies in the evaluation of $E\{\gamma_{1i}^4\}$. For γ_{1i} a Gaussian (rather than binomial) random variable, we would have $E\{\gamma_{1i}^4\} = 3(E\{\gamma_{1i}^2\})^2 = 3/\eta^2$ and thus (D.3) would become

$$E\left\{\left[\left(1 - \sum_{i=2}^M \gamma_{1i}^2\right)\sqrt{E_b} - \sqrt{E_b}\left(1 - \frac{M-1}{\eta}\right)\right]^2\right\} = E_b \left(\frac{2(M-1)}{\eta^2}\right) \quad (D.3')$$

In (D.4), the expectation of the fourth-order product of the γ_{ij} 's partitions into two expectations of second-order products of the γ_{ij} 's resulting in

$$\begin{aligned} \frac{1}{2} E\left\{\left|\sqrt{E_b} \sum_{i=2}^M a_{0i} e^{j(\phi_i - \phi_1)} \sum_{\substack{k=2 \\ k \neq i}}^M \gamma_{ik} \gamma_{1k}\right|^2\right\} &= \frac{E_b}{2} \sum_{i=2}^M \sum_{\substack{k=2 \\ k \neq i}}^M \sum_{\substack{l=2 \\ l \neq i}}^M E\{\gamma_{ik} \gamma_{1k} \gamma_{il} \gamma_{1l}\} = \frac{E_b}{2} \sum_{i=2}^M \sum_{\substack{k=2 \\ k \neq i}}^M E\{\gamma_{ik}^2\} E\{\gamma_{1k}^2\} \\ &= \frac{E_b}{2} \frac{(M-1)(M-2)}{\eta^2} \end{aligned} \quad (D.4')$$

Similarly, for the independent Gaussian assumption on the γ_{ij} 's, (D.5) now becomes

$$\begin{aligned} \frac{1}{2} E\left\{\left|n_1 - \sum_{i=2}^M \gamma_{1i} n_i\right|^2\right\} &= \frac{N_0}{2} - N_0 \sum_{i=2}^M E\{\gamma_{1i}^2\} + \frac{N_0}{2} \sum_{i=2}^M E\{\gamma_{1i}^2\} \\ &= \frac{N_0}{2} \left(1 - \frac{M-1}{\eta}\right) \end{aligned} \quad (D.5')$$

Finally, (D.6) remains unchanged and (D.7) now reflects the sum of (D.3'), (D.4') and (D.5') which gives

$$\sigma_{11}^2 = \frac{N_0}{2} \left(1 - \frac{M-1}{\eta} \right) + \frac{E_b}{2} \frac{(M-1)(M+2)}{\eta^2} \quad (\text{D.7'})$$

Appendix E

Various Moments of the User Crosscorrelations

In this appendix we evaluate various moments of the user crosscorrelation γ_{ij} defined in (A.6) and expressed in terms of the user PN code sequences in (A.8), i.e.,

$$\gamma_{ij} = \frac{1}{\eta} \sum_{m=1}^{\eta} c_{im} c_{jm} \quad (\text{E.1})$$

where we have made use of the fact that the number of chips per bit N_c is equal to the spreading ratio $\eta = T_b / T_c$ defined in the main text. Then,

$$E\{\gamma_{ij}\} = \frac{1}{\eta} \sum_{m=1}^{\eta} E\{c_{im} c_{jm}\} = 0 \quad (\text{E.2})$$

$$E\{\gamma_{ij}^2\} = \frac{1}{\eta^2} \sum_{m=1}^{\eta} \sum_{n=1}^{\eta} E\{c_{im} c_{jm} c_{in} c_{jn}\} = \frac{1}{\eta^2} \sum_{m=1}^{\eta} E\{c_{im}^2 c_{jm}^2\} = \frac{1}{\eta} \quad (\text{E.3})$$

$$E\{\gamma_{ij} \gamma_{jm} \gamma_{im}\} = \frac{1}{\eta^3} \sum_{l=1}^{\eta} \sum_{k=1}^{\eta} \sum_{n=1}^{\eta} E\{c_{il} c_{jl} c_{ik} c_{mk} c_{in} c_{mn}\} = \frac{1}{\eta^3} \sum_{l=1}^{\eta} E\{c_{il}^2 c_{jl}^2 c_{ml}^2\} = \frac{1}{\eta^2} \quad (\text{E.4})$$

$$E\{\gamma_{ij}^3\} = \frac{1}{\eta^3} \sum_{k=1}^{\eta} \sum_{l=1}^{\eta} \sum_{m=1}^{\eta} E\{c_{ik} c_{jk} c_{il} c_{jl} c_{im} c_{jm}\} = 0 \quad (\text{E.5})$$

$$E\{\gamma_{ij}^2 \gamma_{ik}^2\} = \frac{1}{\eta^4} \sum_{l=1}^{\eta} \sum_{m=1}^{\eta} \sum_{n=1}^{\eta} \sum_{r=1}^{\eta} E\{c_{il} c_{jl} c_{im} c_{jm} c_{in} c_{kn} c_{ir} c_{kr}\} = \frac{1}{\eta^4} \sum_{l=1}^{\eta} \sum_{n=1}^{\eta} E\{c_{il}^2 c_{jl}^2 c_{in}^2 c_{kn}^2\} = \frac{1}{\eta^2} \quad (\text{E.6})$$

$$\begin{aligned} E\{\gamma_{ij}^4\} &= \frac{1}{\eta^4} \sum_{k=1}^{\eta} \sum_{l=1}^{\eta} \sum_{m=1}^{\eta} \sum_{n=1}^{\eta} E\{c_{ik} c_{jk} c_{il} c_{jl} c_{im} c_{jm} c_{in} c_{jn}\} \\ &= 3 \frac{1}{\eta^4} \sum_{\substack{k=1 \\ k \neq m}}^{\eta} \sum_{m=1}^{\eta} E\{c_{ik}^2 c_{jk}^2 c_{im}^2 c_{jm}^2\} + \frac{1}{\eta^4} \sum_{k=1}^{\eta} E\{c_{ik}^4 c_{jk}^4\} \\ &= \frac{3\eta(\eta-1)}{\eta^4} + \frac{1}{\eta^3} = \frac{3}{\eta^2} - \frac{2}{\eta^3} \end{aligned} \quad (\text{E.7})$$

$$\begin{aligned}
E\{\gamma_{ik}\gamma_{jk}\gamma_{il}\gamma_{jl}\} &= \frac{1}{\eta^4} \sum_{m=1}^{\eta} \sum_{n=1}^{\eta} \sum_{r=1}^{\eta} \sum_{s=1}^{\eta} E\{c_{im}c_{kn}c_{jn}c_{km}c_{ir}c_{lr}c_{js}c_{ls}\} \\
&= \frac{1}{\eta^4} \sum_{m=1}^{\eta} \sum_{r=1}^{\eta} E\{c_{im}c_{jm}c_{km}^2c_{ir}c_{jr}c_{lr}^2\} = \frac{1}{\eta^4} \sum_{m=1}^{\eta} E\{c_{im}^2c_{jm}^2\} \\
&= \frac{1}{\eta^3}
\end{aligned} \tag{E.8}$$

Appendix F

Some Useful Closed Form Definite Integrals

In this appendix, we tabulate some closed form definite integrals that are particularly useful in evaluating the moments of (13). Many of these which deal with integrands containing a Gaussian probability density function (pdf) or Gaussian probability integral cannot be found in standard tables of integrals.

Define the one- and two-dimensional zero mean Gaussian pdf's by

$$G_1(x; \sigma_x) \triangleq \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left\{-\frac{x^2}{2\sigma_x^2}\right\} \quad (\text{F.1})$$

$$G_2(x, y; \sigma_x, \sigma_y) \triangleq \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{\sigma_y^2x^2 + \sigma_x^2y^2 - 2\rho\sigma_x\sigma_yxy}{2\sigma_x^2\sigma_y^2(1-\rho^2)}\right\}$$

and the normalized Gaussian probability integral [see Eq. (12)]

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy \quad (\text{F.2})$$

Denote the hard limiter and null zone limiter functions, respectively, by

$$\text{sgn } x = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}; \quad \text{nsgn } x = \begin{cases} 1, & x > \zeta \\ 0, & -\zeta \leq x \leq \zeta \\ -1, & x < -\zeta \end{cases} \quad (\text{F.3})$$

Then, the following one- and two-dimensional integrals are appropriate:

One-dimensional integrals

$$\int_{-\infty}^{\infty} \exp\{-A(1+Bx)^2\} G_1(x; \sigma_x) dx = \frac{1}{\sqrt{1+2\sigma_x^2AB^2}} \exp\left\{-\frac{A}{1+2\sigma_x^2AB^2}\right\} \quad (\text{F.4})$$

$$\int_{-\infty}^{\infty} x \exp\{-Ax^2 \pm Bx\} G_1(x; \sigma_x) dx = \pm \frac{B}{2\sqrt{2a^3\sigma_x^2}} \exp\left\{\frac{B^2}{4a}\right\}; \quad a \triangleq A + \frac{1}{2\sigma_x^2} \quad (\text{F.5})$$

$$\int_{-\infty}^{\infty} x^2 \exp\{-Ax^2 \pm Bx\} G_1(x; \sigma_x) dx = \frac{1}{2\sqrt{2a^3\sigma_x^2}} \left(1 + \frac{B^2}{2a}\right) \exp\left\{\frac{B^2}{4a}\right\}; \quad a \triangleq A + \frac{1}{2\sigma_x^2} \quad (\text{F.6})$$

$$\int_{-\infty}^{\infty} x \operatorname{sgn}(Ax + B) G_1(x; \sigma_x) dx = \sqrt{\frac{2\sigma_x^2}{\pi}} \exp\left\{-\frac{(B/A)^2}{2\sigma_x^2}\right\} \quad (\text{F.7})$$

$$\int_{-\infty}^{\infty} x \operatorname{nsgn}(Ax + B) G_1(x; \sigma_x) dx = \frac{1}{2} \sqrt{\frac{2\sigma_x^2}{\pi}} \left[\exp\left\{-\frac{[(\zeta - B)/A]^2}{2\sigma_x^2}\right\} + \exp\left\{-\frac{[(\zeta + B)/A]^2}{2\sigma_x^2}\right\} \right] \quad (\text{F.8})$$

$$\int_{-\infty}^{\infty} x Q(A(1+x)) G_1(x; \sigma_x) dx = -\sqrt{\frac{\sigma_x^2}{2\pi}} \left(\frac{A^2\sigma_x^2}{1+A^2\sigma_x^2}\right) \exp\left\{-\frac{A^2}{2(1+A^2\sigma_x^2)}\right\} \quad (\text{F.9})$$

$$\int_{-\infty}^{\infty} x^2 Q(A(1+x)) G_1(x; \sigma_x) dx = \sqrt{\frac{\sigma_x^2}{2\pi}} \frac{A^3\sigma_x^3}{(1+A^2\sigma_x^2)^{3/2}} \exp\left\{-\frac{A^2}{2(1+A^2\sigma_x^2)}\right\} + \sigma_x^2 Q\left(\frac{A}{\sqrt{1+A^2\sigma_x^2}}\right) \quad (\text{F.10})$$

Two-dimensional integrals

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left\{-[a(A+x)^2 + b(B+y)^2]\right\} G_2(x; \sigma, \sigma) dx dy \\ &= \sqrt{\frac{1-\rho^2}{(1+a\Gamma)(1+b\Gamma)-\rho^2}} \exp\left\{-\frac{(aA^2 + bB^2)(1-\rho^2) + (A^2 + B^2 - 2\rho AB)ab\Gamma}{(1+a\Gamma)(1+b\Gamma)-\rho^2}\right\}; \quad \Gamma \triangleq 2\sigma^2(1-\rho^2) \end{aligned} \quad (\text{F.11})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \operatorname{sgn}(A+y) G_2(x; \sigma_x, \sigma_y) dx = \sqrt{\frac{2\rho^2\sigma_x^2}{\pi}} \exp\left\{-\frac{A^2}{2\sigma_y^2}\right\} \quad (\text{F.12})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \operatorname{nsgn}(A+y) G_2(x; \sigma_x, \sigma_y) dx = \frac{1}{2} \sqrt{\frac{2\rho^2\sigma_x^2}{\pi}} \left[\exp\left\{-\frac{(A-\zeta)^2}{2\sigma_y^2}\right\} + \exp\left\{-\frac{(A+\zeta)^2}{2\sigma_y^2}\right\} \right] \quad (\text{F.13})$$

