# Iterative Multiuser Detection

## Iterating between multiuser detection and channel decoding to achieve low-complexity, near-optimal demodulation in coded multiple-access channels.



*H. Vincent Poor*

Communication channels that involve both error-control coding and multiple-access signaling are of increasing interest in applications such as cellular telephony, wireless computer networks, and broadband local access. Optimal data detection and decoding in such channels generally requires a level of computational complexity that is prohibitive for these types of applications. Turbo multiuser detection (MUD) addresses this problem by applying the turbo principle of iteration among constituent decision algorithms, with intermediate exchanges of soft information (i.e., posterior probabilities) about tentative decisions. Here this principle is applied by considering MUD (which exploits the multiple-access signaling structure) and error-control decoding as the two consituent decision algorithms.


1053-5888/04/$20.00©2004IEEE

The resulting iteration between soft MUD and soft channel decoding yields quite good results. This article reviews this area, outlining both the basic principles involved and the basis for low-complexity turbo multiuser detectors that require minimal increased complexity over that of the standard channel decoder.

## Introduction

MUD refers to the detection of data from multiple terminals in a communication network when observed in a nonorthogonal multiplex, that is, when derived from a multiple-access channel. This problem arises naturally, for example, in code-division multiple-access (CDMA) systems using nonorthogonal spreading codes. It also arises in orthogonally multiplexed wireless channels, such as time-division multiple-access channels, due to effects such as multipath or nonideal frequency channelization, and in wireline channels such as those arising in digital subscriber line (DSL) systems or powerline communications (PLC) in which crosstalk and other types of interference are major impairments. The basic idea of MUD is to exploit the cross-correlations among the signals to be demodulated to improve the data detection process. Considerable progress has been made on this problem over the past two decades. (See, e.g., [31] and [34].) Among other things, it has been shown that the use of MUD can provide very significant performance advantages in interference-limited channels.

There are many types of MUD techniques. Optimal techniques, based on maximum-likelihood (ML) or maximum a posteriori probability (MAP) criteria, can often achieve performance very close to that of a system that is free of interference. However, these methods tend to be quite complex, particularly when compared with the processing resources available in most communications receivers. Consequently, a considerable amount of effort has been devoted to the development of lower-complexity techniques that can achieve some of the benefits of the optimal procedures. One class of such methods are the linear multiuser detectors, which use linear processing to suppress interference, followed by simple memoryless quantization to perform data detection. Another class of lower-complexity multiuser detectors are the iterative multiuser detectors, which make use of tentative channel-symbol decisions (either soft or hard) to provide feedback that can improve the capabilities, in terms of complexity or performance, of optimal or linear MUD methods.

When channel coding is considered in addition to nonorthogonal signaling, the complexity of optimal receiver processing is further exacerbated. In particular, the complexity of optimal (ML or MAP) joint MUD and channel decoding tends to be extremely high. However, this combination also lends itself very well to the use of iterative MUD methods in which the tentative channel-symbol decisions are produced by the channel decoders. Similarly, MUD can be used to provide tentative channel-symbol decisions to the channel decoders. Iteration between these two constituent processes, with intermediate exchanges of soft channel-symbol information, is known as turbo MUD. This idea was originally developed in the context of convolutionally encoded CDMA channels, but has since been applied in a number of other frameworks, including DSL, PLC, space-time coded CDMA, ultra-wideband (UWB), and turbo-coded CDMA channels.

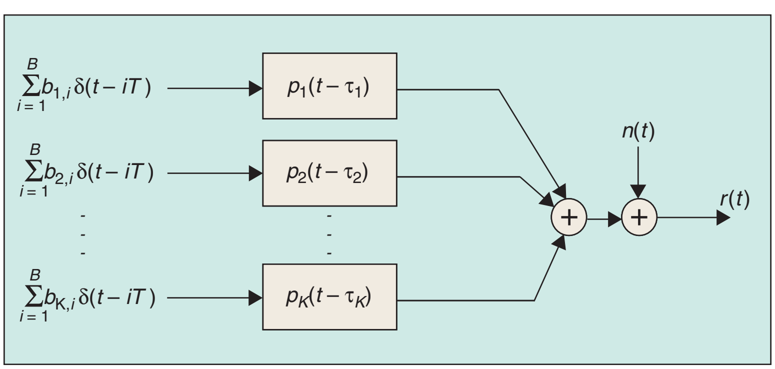
## Elements of MUD

In this section, we review briefly some basic results in MUD. Consider the reception of a multiple-access communication signal as illustrated in Figure 1. Such a signal can be written in the following form:

$$r(t) = \sum_{k=1}^{K}\sum_{i=1}^{B} b_{k,i}\, p_k(t - iT - \tau_k) + n(t), \qquad (1)$$

where $K$ is the number of users active in the channel, $B$ is the number of symbols per user in a received frame to be processed, $T$ is the per-user symbol interval (so, $1/T$ is the per-user signaling rate), $b_{k,i}$ denotes the $i$th symbol of the $k$th user, $p_k(\cdot)$ is the received signaling waveform of user $k$, $\tau_k \in [0, T]$ is the delay with which user $k$'s signal is received, and $n(\cdot)$ is a white Gaussian process with spectral height $N_0/2$.

For the sake of exposition we assume that the data and channel symbols take binary antipodal ($\pm 1$) values, although this is easily relaxed to include any finite alphabet. We also assume that the observations are real valued, although again this assumption is not essential to any of what follows. Note also that channel characteristics such as multipath fading or dispersion can be incorporated into the waveforms $\{p_k(\cdot)\}_{k=1,2,\ldots,K}$. These waveforms can also depend on the symbol index $i$, a situation that occurs, for example, in long-code CDMA systems. The results discussed in this article also apply in this case as well, although we will not explicitly include it here for the sake of simplicity. And, finally, it should be noted that the waveforms can also be vector valued, which corresponds to the case of multiple




▲ *1. Illustration of the multiple-access channel.*

receive antennas. Again, this case is a straightforward generalization of the following discussion (e.g., [6]).

### Optimal MUD

We would like to make decisions about the set of symbols $\{b_{k,i}\}$, of which there are $KB$. Of course, if the waveforms $\{p_k(\cdot)\}_{k=1,2,\ldots,K}$ are orthogonal to one another and do not extend beyond a single symbol interval, this problem is a classical one, solved by the conventional matched filter receiver [21]:

$$\hat{b}_{k,i} = \text{sgn}\{y_{k,i}\}, \; k = 1, \ldots, K, \, i = 1, \ldots, B, \quad (2)$$

where $y_{k,i}$ denotes the output of a filter matched to the waveform of the $k$th user in its $i$th symbol interval; i.e.,

$$y_{k,i} = \int_{-\infty}^{\infty} r(t) \; p_k(t - iT - \tau_k) dt. \quad (3)$$

Here, we are interested in the more general case, in which these waveforms are not orthogonal, and where perhaps there is intersymbol interference (ISI). In general, a sufficient statistic for making inferences about $\{b_{k,i}\}$ is formed by the set of $KB$ matched-filter outputs (see, e.g., [21]):

$$y_{k,i}, \; k = 1, \ldots, K, \, i = 1, \ldots, B. \quad (4)$$

Organizing these observables into a vector $\mathbf{y} \in \mathbb{R}^{KB}$ by sorting them first by symbol number and then by user number, the model (1) can be rewritten as a linear model

$$\mathbf{y} = \mathbf{H}\,\mathbf{b} + \mathcal{N}\left(0, \frac{N_0}{2}\mathbf{H}\right), \quad (5)$$

where $\mathbf{b} \in \{-1, +1\}^{KB}$ denotes a vector containing the symbols $\{b_{k,i}\}$ sorted conformally with $\mathbf{y}$ and where $\mathbf{H}$ denotes a matrix of cross-correlations

$$H_{n,m} = \int_{-\infty}^{\infty} p_k(t - iT - \tau_k) \; p_\ell(t - jT - \tau_\ell) \; dt \quad (6)$$

with the indices $(k, i)$ and $(\ell, j)$ corresponding in the model (1) to the indices $n$ and $m$, respectively, in the vector $\mathbf{y}$. The relationship between $n, m$ and $(k, i), (\ell, j)$ is

$$\begin{aligned} k &= [n-1]_K, \qquad & i &= \left\lfloor \frac{n-1}{K} \right\rfloor, \\ \ell &= [m-1]_K, \qquad & j &= \left\lfloor \frac{m-1}{K} \right\rfloor, \end{aligned} \quad (7)$$

where $[\cdot]_K$ denotes reduction modulo $K$. The term $\mathcal{N}(0, N_0\mathbf{H}/2)$ denotes a noise term having the multivariate Gaussian distribution with zero mean and covariance matrix $N_0\mathbf{H}/2$.

Thus, the basic problem of MUD is, in general, a problem of sequence detection, which involves mapping the vector (or sequence) of observables $\mathbf{y}$ into a vector $\hat{\mathbf{b}}$ of estimates of the symbols $\mathbf{b}$. When this mapping is chosen to satisfy optimality criteria such as ML or MAP, the resulting complexity is nominally quite high: $\mathcal{O}(2^{KB})$. Fortunately, this problem typically can be solved with much lower complexity via dynamic programming [23]. For example, in a dispersive channel in which the waveforms $\{p_k\}$ span at most $\Delta$ symbol intervals, the per-symbol complexity can be reduced to $\mathcal{O}(2^{K\Delta})$. The corresponding dynamic program for the single-user ($K = 1$) dispersive ($\Delta > 1$) case is given in the ML case by the maximum-likelihood sequence detector [25], and for the multiuser ($K > 1$) nondispersive ($\Delta = 1$) case by the ML or the MAP multiuser detector [31]. Although this order of complexity is much better than the nominal complexity of this problem, it is still quite high for many applications. As noted previously, in exchange for this complexity, these detectors can offer a level of performance very close to that which would be observed in an interference-free channel.

### Linear MUD

The basic difficulty with optimal multiuser detectors is their complexity. A considerable amount of research has been devoted to the development of suboptimal multiuser detectors that mitigate this complexity (see, e.g., [31]). One well-studied family of suboptimal multiuser detectors are the linear multiuser detectors, which are of interest in their own right and which also form the basis for many iterative multiuser detectors, including the low-complexity turbo MUD.

The sufficient statistic $\mathbf{y}$ of (4) obeys the linear model (5), and MUD (and equalization as well) can be viewed as the fitting of this model to the observations. The complexity of these problems comes from the fact that the elements of the vector $\mathbf{b}$ take values in a finite alphabet. Without this constraint, the fitting of linear models such as (5) is of relatively low complexity. The basic idea of linear MUD is to take advantage of this relatively low complexity of unconstrained linear model-fitting by first estimating $\mathbf{b}$ in (5) as if it were a vector with real components, and then to project these real estimates onto the finite alphabet of the actual symbols. This, of course, will not yield ML or MAP symbol decisions, but it often works quite well.

Note that the matched filter detector (3)–(4) is a very simple example of a linear multiuser detector, in which the vector $\mathbf{y}$ itself is used to estimate $\mathbf{b}$ before quantization. As noted above, this choice is optimal against the white background noise in the absence of signal cross correlations. On the other hand, referring to (5), we see that this choice essentially ignores the off-diagonal elements of the cross-correlation matrix $\mathbf{H}$. A key alternative to the matched filter is the linear minimum-mean-square-error (MMSE) detector, which detects $\mathbf{b}$ via

$$\hat{\mathbf{b}} = \text{sgn}\left\{\left(\mathbf{H} + \frac{N_0}{2}\mathbf{I}\right)^{-1} \mathbf{y}\right\} \quad (8)$$

**Multiuser detection refers to the detection of data from multiple terminals in a communication network when observed in a nonorthogonal multiplex.**

where $\mathbf{I}$ denotes the $KB \times KB$ identity matrix. This latter detector uses, as its linear estimation stage, the linear MMSE estimator of $\mathbf{b}$ given $\mathbf{y}$ in (5) under the assumption that the symbols have a prior distribution under which they are uncorrelated with zero means; namely, $(\mathbf{H} + N_0\mathbf{I}/2)^{-1}\mathbf{y}$.

### *Iterative MUD*

Turbo MUD falls within the category of iterative MUD, in which tentative decisions are used iteratively to improve overall data detection. Aside from turbo MUD, iterative detectors include several varieties, including linear and nonlinear interference cancellers, and model-based techniques such as those based on the expectation-maximization (EM) algorithm. We now discuss these very briefly.

Note that the linear detectors discussed above typically require the inversion of a $KB \times KB$ matrix. The complexity of the matrix inversion is, in its worst case, $\mathcal{O}((KB)^3)$. Although simpler in principle than the exponential complexity of ML or MAP MUD, this complexity can still be quite significant. Moreoever, this matrix inversion cannot necessarily be amortized over more than one frame of data, since the channel and/or the signaling waveforms may vary from frame to frame. Thus, it is of interest to use lower-complexity methods for computing the estimates used in linear MUD. This problem is a well-studied problem in linear algebra. In particular, we wish to solve an equation of the form

$$\mathbf{C}\mathbf{x} = \mathbf{y} \tag{9}$$

for the unknown vector $\mathbf{x}$, where $\mathbf{C}$ is, for example, $(\mathbf{H} + N_0\mathbf{I}/2)$. Equation (9) can be solved efficiently by using any of several iterative equation-solving methods, such as the Gauss-Siedel, Jacobi, or congugate-gradient methods. Such methods are known as linear interference cancellers, and in general these techniques can achieve performance comparable to full linear MUD, with complexity of order $\mathcal{O}(K \Delta m_{\max})$, where $m_{\max}$ is the maximum number of iterations used. (See, e.g., [6] and the references cited therein for a discussion of these methods.)

Nonlinear interference cancellers are similar in spirit to linear interference cancellers, in that they use iterative methods to fit the model (5). Unlike their linear counterparts, however, that exploit only the linearity of the model while iterating, nonlinear interference cancellers also exploit the discrete nature of the symbol vector $\mathbf{b}$ at each iteration by making intermediate soft or hard decisions between iterations. As with linear interference cancellers, there are a number of such methods. (See [6] for a discussion of these detectors.)

As noted above, the basic problem of MUD is the accurate fitting of the model (5). Linear interference cancellers perform this fitting by exploiting only the linear structure of the model, while nonlinear interference cancellers seek to improve on this fit by making use of further information about the model, namely that the symbols are elements of a known finite alphabet. Often, further information is known about the symbols, and this can also be exploited to provide further performance improvement. For example, the EM algorithm or Markov-chain Monte Carlo (MCMC) techniques can be used to exploit statistical information about $\mathbf{b}$, leading to several soft-decision iterative nonlinear MUD algorithms. (See, e.g., [34].) Turbo MUD is a further example of such an exploitation, in which the information to be exploited is the set of constraints imposed by channel coding.

## Iterative Joint MUD and Decoding

We now turn to the situation in (1) in which the symbols are constrained by having been produced by an error-correcting code. In principle, this constraint should strengthen our ability to fit the model (5), as it reduces the number of sequences $\mathbf{b}$ that are possible. However, the complexity of including such constraints is quite high, as we will see below. Essentially, turbo MUD is a technique for fitting (5) when the symbols satisfy coding constraints with dramatically lower complexity than optimal algorithms.

Error-control coding is, of course, ubiquitous in wireless and other impaired channels. Similarly to MUD, the decoding of error-control codes exploits the dependencies among successive channel symbols to improve the detection of a single stream of data symbols. Like MUD, channel decoding typically involves very complex optimal algorithms, and so complexity issues often dominate the study of these problems. Notable among coding techniques with this problem are parallel and serially concatenated codes separated by interleavers, which are known to offer considerable performance improvement over traditional codes, exhibiting near-Shannon-limit performance in many cases. However, although the optimal decoding of such codes is of particularly high complexity, iterative or turbo decoding algorithms that involve the iterative exchange of soft information between constituent decoders (separated by interleavers/de-interleavers) have been shown to be very effective approximations to optimal decoding. These well-known ideas are discussed, for example, in [12] and [13].

Many communication systems involve both error-control coding and nonorthogonal multiplexing. A typical configuration is a convolutional encoder mapping data symbols into channel symbols, followed by an interleaver, and then a multiple-access channel, as shown in Figure 2. We will focus on this model, although other applications can also fit within the formalism discussed here. One can view the configuration of Figure 2 as a serially concatenated code, in which the multiple-access channel (e.g., a CDMA spreading code) is the inner code, and the convolutional code is the outer code. A traditional way of decoding this concatenation is to first demodulate the multiple-access signals (using either a conventional matched-filter detector, or a multiuser detector) and then to follow this demodulator by a de-interleaver and a channel decoder.

To seek optimality in such a situation, we could replace this traditional configuration with an overall optimal demodulator/decoder that uses an optimal (say ML or MAP) mapping from the received signal to the original data symbols. The complexity of such a system is potentially quite high. This complexity can be mitigated however, by appealing to the turbo principle for decoding concatenated codes noted above. In particular we can reduce the complexity of joint decoding and MUD by an iterative exchange of soft information, iterating until some kind of convergence is reached. Like turbo decoding, this iterative approach to joint MUD and channel decoding can achieve very good performance (close to the interference-free case).

To consider this problem, we need to modify the model of (1) to include coding. This can be done very simply, by writing the channel symbols $\{b_{k,i}\}$ explicitly as functions of underlying data symbols; i.e., for a rate-$R$ code, we have

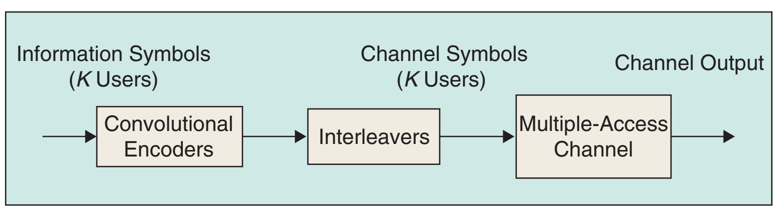
$$r(t) = \sum_{k=1}^{K} \sum_{i=1}^{B} b_{k,i}(\mathbf{d}_k)\; p_k(t - iT - \tau_k) + n(t), \quad (10)$$

where, in addition to the definitions given with (1), we have that $\mathbf{d}_k$ is a set of $RB$ data symbols being transmitted by user $k$. Thus, $b_{k,i}(\mathbf{d}_k)$, $i = 1, \ldots, B$, denote the post-interleaver channel symbols obtained by encoding $\mathbf{d}_k$, $k = 1, 2, \ldots, K$. We assume that the code is a convolutional code with constraint length $\nu$, and for simplicity we consider the nondispersive ($\Delta = 1$) case. (The extension to the dispersive case ($\Delta > 1$) is straightforward. An approach to joint iterative equalization and channel decoding for the single-user ($K = 1$) case has been examined in [9], with further results in [26], [30], and [35].)

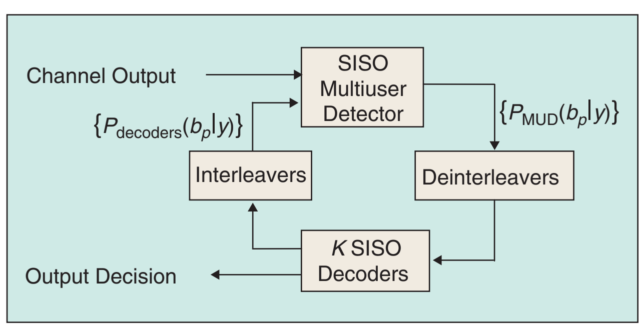
We would like to make inferences about the set of data symbol vectors $\mathbf{d}_1, \ldots, \mathbf{d}_K$, which contain a total of $K\,RB$ symbols. The observation vector $\mathbf{y}$ of (3) and (4) is still a sufficient statistic for such inferences, and thus joint channel decoding and MUD is another problem of sequence detection. Like MUD, the decoding task in this situation can be simplified by dynamic programming. For example, in the single-user ($K = 1$) case, the per-symbol complexity of optimal decoding reduces to $\mathcal{O}(2^{\nu})$, with the corresponding dynamic program being specified by the Viterbi algorithm in the case of ML decoding and by the Bahl, Cocke, Jelinek, and Raviv (BCJR) algorithm in the case of MAP decoding [25]. With multiple users ($K > 1$), optimal detection and decoding in this problem essentially combines the complexity of the constituent problems, to yield a dynamic program with $\mathcal{O}(2^{K\nu})$ complexity [11]. This complexity would typically be too high for most applications, since the constraint length of the code would normally be chosen to meet the limits of the receiver's processing capabilities. Amplifying this constraint length by a factor of $K$ in the exponent will push the processing capability well beyond its limits.

Like turbo-coded systems, this complexity can be mitigated by making use of the turbo principle of iterating between algorithms for the constituent problems, and exchanging soft information between iterations. (See [3], [18], and [22].) The basic building blocks of a turbo multiuser detector are a soft-input/soft-output (SISO) multiuser detector and a bank of single-user SISO channel decoders, as shown in Figure 3. The role of each of these algorithms is to compute posterior probabilities of the channel symbols based on given prior probabilities and on the corresponding signal structure. That is, the SISO multiuser detector uses prior symbol probabilities and the multiuser signaling structure to compute posterior symbol probabilities conditioned on the observations. Similarly, the SISO channel decoders use prior symbol probabilities and the structure imposed by the channel code to compute posterior symbol probabilities. (Of course, the SISO decoders also compute posterior data symbol probabilities, which will ultimately yield the overall output of the combined algorithm.)

The turbo multiuser detector begins with a SISO multiuser detector applied to the frame of $B$ channel symbols ($B$ is assumed to be equal to the interleaver length). This detector particularly computes posterior probabilities, conditioned on the observations $\mathbf{y}$, for




▲ *2. The multiple-access channel with convolutional coding.*




▲ *3. General structure of turbo multiuser detection.*

**This complexity can be mitigated by making use of the turbo principle of iterating between algorithms for the constituent problems and exchanging soft information between iterations.**

each of the channel symbols of each of the users; that is, for each element of the vector **b**. This first set of posterior probabilities is based on the prior assumption that the channel symbols are drawn uniformly from $\{-1, +1\}^{KB}$; that is, that the channel symbols are i.i.d. equiprobably $\pm 1$ random variables. Although this assumption is not correct due to the channel coding (which correlates the channel symbols), it serves as a useful approximation for initializing the algorithm because the interleavers at the transmitter serve to decorrelate the symbols as they appear at the input to the channel.

The posterior probabilities computed by the SISO MUD will then be used as prior probabilities in the next step of the algorithm, which makes use of the bank of single-user channel decoders. Before applying channel decoding, however, the symbols must be de-interleaved to return them to their correct order for decoding. This de-interleaving has the approximate effect of removing any correlations that are introduced into the channel symbols by conditioning on the observations **y** in the SISO MUD. Thus, after SISO MUD and de-interleaving, the channel symbols can again be assumed to be independent of one another, but now having marginal (i.e., individual) probability distributions determined by the probabilities computed by the SISO MUD. This probability model becomes the prior probability model used by the SISO channel decoders, which compute (via, say, the BCJR algorithm) corresponding posterior probabilities for both the channel and data symbols.

These posterior probabilities for the data symbols could, at this point, be used to MAP decode the data symbols. This would correspond to a conventional receiver approach based on MUD followed by decoding. However, a more powerful receiver results by re-interleaving the channel symbols at the output of the decoders and returning to the SISO MUD, now using as a prior distribution the posterior channel-symbol probabilities computed by the SISO decoders. The SISO MUD then refines its estimates of the posterior probabilities of the symbol probabilities and hands them back to the channel decoders after de-interleaving again. This process of soft-information exchange between the SISO MUD and the SISO decoders can continue until the posterior channel-symbol probabilities converge to stable values, at which point the data symbols can be MAP decoded via the data-symbol posterior probabilities computed on the last application of the SISO decoding algorithm. The constituents of this process, namely MAP MUD and MAP decoding, are well known, and thus details are omitted for the sake of brevity. (Explicit equations can be found in [34].)

From this description, it can be seen that the interpretation of the multiuser detector as a posterior-probability calculator is an essential philosophical underpinning of this approach. Unlike the case with turbo decoding, however, in which the complexity of the constituent decoders is controlled by the system designer, the complexity of the SISO multiuser detector used in this turbo multiuser detector is dependent on the number of users in the channel and is thus beyond the designer's immediate control. Thus, although the $\mathcal{O}(2^{K\nu})$ complexity of optimal joint detection and decoding noted in [11] is reduced to $\mathcal{O}(2^{\nu}) + \mathcal{O}(2^{K})$ via the turbo principle, the second term in this complexity order is prohibitive for most applications, as noted previously.

Because of this complexity issue, some simpler techniques, in which the MUD component of such an iterative scheme is replaced by simpler suboptimal algorithms such as interference cancellers, etc., have been considered by several authors. (See, e.g., [1], [2], [4], [16], [19], and [20].) Moreover, an alternative approach based on an approximate posterior-probability calculator that significantly simplifies the SISO MUD has been developed in [32]. This approach is described briefly in the following section.

## Low-Complexity Turbo MUD

The basic difficulty with the turbo multiuser detector described in the preceding section is the $\mathcal{O}(2^{K})$ complexity of the MAP MUD stages. As discussed earlier, this complexity can be mitigated by using, for example, linear MUD.

Although linear detectors are of considerable interest in the implementation of practical MUD systems, they do not immediately appear to be useful in the context of turbo MUD since they are based on linear regression type criteria rather than on posterior probability computation. So, these detectors do not appear on their face to be amenable to posterior probability calculation. However, it happens that the linear MMSE detector, in fact, can be used as an approximate posterior probability calculator. This is due to the property, exposed in [24], that the residual error in the linear MMSE estimator used by the MMSE detector is approximately Gaussian. From this property, we can obtain posterior probability estimates for the channel symbols conditioned on the output of the linear MMSE transformation (11) straightforwardly via Bayes' formula. Thus, a very useful turbo multiuser detector arising from the use of the linear MMSE detector to create a low-complexity SISO stage. Such a detector was proposed in [32].

To implement this detector, we must first modify the linear MMSE detector to allow for the incorporation of prior channel-symbol probabilities via (11). The standard linear MMSE detector of (8) can easily be modified to account for a prior distribution with nonzero mean, which results in the linear estimator

$$\left(\mathbf{H} + \sigma^2 \Sigma^{-1}\right)^{-1} \left(\mathbf{y} - \mathbf{H}\tilde{\mathbf{b}}\right) \qquad (11)$$

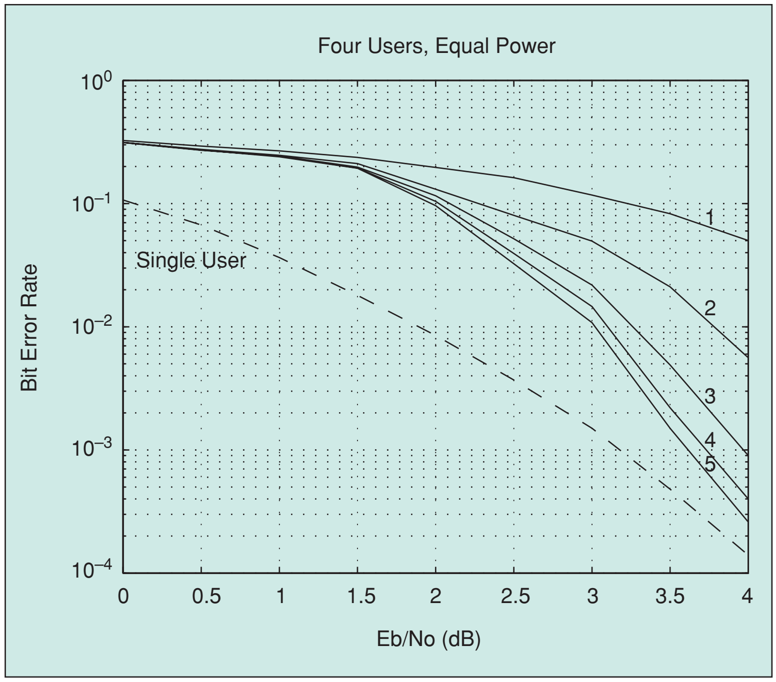
where $\Sigma$ and $\tilde{\mathbf{b}}$ denote, respectively, the prior covariance and mean of $\mathbf{b}$. The elements of $\tilde{\mathbf{b}}$ are thus given by

$$\tilde{b}_{k,i} \;=\; 2\,\pi_{k,i} \;-\; 1, \qquad (12)$$

where $\pi_{k,i}$ denotes the prior probability that $b_{k,i} = 1$; and $\Sigma$ is a diagonal matrix with diagonal elements

$$\Sigma_{n,n} \;=\; 4\,\pi_{k,i}\,(1 - \pi_{k,i}), \qquad (13)$$

where, as in (6), the indices $(k, i)$ correspond in the model (10) to the index $n$ in the vector $\mathbf{y}$. (The explicit relationship among these variables in given in (7).) Of course (12) and (13) are simply the mean and variance of a binary $\pm 1$ random variable in which $+1$ occurs with probability $\pi_{k,i}$. The diagonality of $\Sigma$ follows from the assumption that interleaving removes the temporal dependence from the data stream. Note that the complexity of the matrix inversion in (11) is at worst $\mathcal{O}((KB)^3)$, but this can be reduced significantly (to linear-in-$K$ per-symbol complexity) by using the linear iterative methods discussed earlier.




▲ 4. Performance of MMSE-based low-complexity turbo MUD: four equicorrelated users ($\rho = 0.7$); rate-1/2 constraint-length-5 convolutional code; length-128 interleaver.

The application of this idea, which is explored more fully in [32], leads to excellent performance with only a few cycles through the turbo algorithm. Typical performance results show that near-interference-free performance can be achieved quite easily when there is sufficient signal-to-noise ratio (SNR) for the initial SISO MUD to gain useful information about the channel symbols. Figure 4 shows an example of such a result, taken from [32], in which there are $K = 4$ users with equal cross correlations of 0.7, and each making use of a rate-1/2, constraint-length-5 convolutional code with a length-128 interleaver. Note that, here, near-single-user performance is achieved after only five iterations with very moderate SNR. Further approximations to this detector with even lower complexity have also been developed in [32], with comparable performance results.

## Remarks

In this article, we have provided a brief review of turbo MUD, and of the more general context of MUD within which turbo MUD arises. As we have seen, turbo MUD can be viewed as a form of iterative MUD in which channel decoders provide intermediate soft decisions that can be exploited by a multiuser detector. We have also noted that this method, and particularly the linear variant, can provide quite good performance with relatively low complexity.

Although we have discussed turbo MUD only in the context of the convolutionally encoded multiple-access channel, there are many other problems that have been addressed using the general principles of turbo MUD. These include methods for turbo-coded CDMA systems [8], [33], space-time coded systems [15], [17], [20], [28], [29], DSL systems [5], PLC systems [7], UWB impulse radio systems [10], and the introduction of adaptivity [14], [27], [34], [36] into the SISO multiuser detector.

For example, in [15] turbo iteration between space-time trellis decoding and MUD is used to enable space-time coded multiple-access systems that exhibit a "separation" principle between the achievement of full spatial diversity and multiple-access operation. In another example, in [10], iteration between pulse-level soft MUD and symbol level soft symbol detection is used to achieve near single-user performance in multiple-access impulse radio systems.

There have been a great many developments in this field in the past few years. Space prohibits a discussion of all of these developments, or the inclusion of an exhaustive bibliography. However, the reader may refer to [34] for a more detailed treatment of this field.

## Acknowledgments

This work was prepared under the support of the National Science Foundation under Grant CCR-99-80590. It was presented at the Third International Symposium on Turbo Codes and Related Topics, Brest, France, 1–3 September 2003.

*H. Vincent Poor* is the George Van Ness Lothrop Professor in Engineering at Princeton University. He received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977 and was with the University of Illinois from 1977 until 1990, when he returned to Princeton as a faculty member. His research interests are in the area of statistical signal processing, with applications in wireless communications and related fields. Among his publications in these areas are the recent books *Wireless Communication Systems: Advanced Techniques for Signal Reception* (Prentice-Hall, 2004) and *Wireless Networks: Multiuser Detection in Cross-Layer Design* (Kluwer, 2004). He is a member of the National Academy of Engineering and is a Fellow of the IEEE, the Institute of Mathematical Statistics, and other organizations. He is a past president of the IEEE Information Theory Society and has received several recent honors, including the IEEE 2001 Graduate Teaching Award and the 2001 Joint Paper Award of the 2001 IEEE Communications and Information Theory Societies. He is currently on sabbatical leave from Princeton, working under a fellowship from the John Simon Guggenheim Memorial Foundation.

## References


[1] P.D. Alexander, A.J. Grant, and M.C. Reed, "Iterative detection in code-division multiple-access with error-control coding," *Eur. Trans. Telecomm.,* vol. 9, no. 5, pp. 419–425, 1998.

[2] P.D. Alexander, M.C. Reed, J.A. Asenstorfer, and C.B. Schlegel, "Iterative multiuser interference reduction: Turbo CDMA," *IEEE Trans. Commun.,* vol. 47, no. 7, pp. 1008–1014, 1999.

[3] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory,* vol. 48, no. 4, pp. 1772–1773, July 2002.

[4] F. Brannstrom, T.M. Aulin, and L.K. Rasmussen, "Iterative detectors for trellis-code multiple-access," *IEEE Trans. Commun.,* vol. 50, no. 9, pp. 1478–1485, 2002.

[5] H. Dai and H.V. Poor, "Turbo multiuser detection for coded DMT VDSL systems," *IEEE J. Select. Areas Commun.,* vol. 20, no. 2, pp. 351–362, 2002.

[6] H. Dai and H.V. Poor, "Iterative space-time processing for multiuser detection in multipath CDMA channels," *IEEE Trans. Signal Processing,* vol. 50, no. 9, pp. 2116–2127, Sept. 2002.

[7] H. Dai and H.V. Poor, "Advanced signal processing for powerline communications," *IEEE Commun. Mag.,* vol. 41, no. 5, pp. 100–107, May 2003.

[8] A.D. Damnjanovic and B.R. Vojcic, "Iterative multiuser detection," *J. Commun. Networks,* vol. 3, no. 3, pp. 211–218, 2001.

[9] C. Douillard, et al., "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecommun.,* vol. 6, no. 5, pp. 507–511, 1995.

[10] E. Fishler and H.V. Poor, "Low-complexity multiuser detectors for time hopping impulse radio systems," *IEEE Trans. Signal Process.,* to be published.

[11] T.R. Giallorenzi and S.G. Wilson, "Multiuser ML sequence estimator for convolutional coded asynchronous CDMA," *IEEE Trans. Commun.,* vol. 44, no. 8, pp. 997–1008, 1996.

[12] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. Turbo Codes and Related Topics,* Brest, France, 1997, pp. 1–11.

[13] C. Heegard and S.B. Wicker, *Turbo Coding.* Boston, MA: Kluwer, 1995.

[14] M.L. Honig, G. Woodward, and P.D. Alexander, "Adaptive multiuser parallel-decision feedback with iterative decoding," in *Proc. 2000 IEEE Int. Symp. Inform. Theory,* Sorrento, Italy, June 25–30, 2000, p. 335.

[15] S.K. Jayaweera and H.V. Poor, "Low complexity receiver structures for space-time coded multiuser systems," *Eur. J. Applied Signal Processing,* vol. 2002, no. 3, pp. 275–288, 2002.

[16] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.,* vol. 19, no. 8, pp. 1450–1460, 2001.

[17] B. Lu and X. Wang, "Iterative receivers for multiuser space-time coding systems," *IEEE J. Select. Areas Commun.,* vol. 18, pp. 2322–2335, 2000.

[18] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.,* vol. 46, no. 7, pp. 870–880, 1998.

[19] R. Müller and J.B. Huber, "Iterated soft decision interference cancellation for CDMA," in *Proc. 9th Tyrrhenian Int. Workshop Digital Commun.,* Italy, 1997.

[20] A. Naguib and N. Seshadri, "Combined interference cancellation and ML decoding of space-time block codes," in *Proc. 1998 Global Telecommunications Conf.,* Sydney, Australia, Nov. 1998.

[21] H.V. Poor, *An Introduction to Signal Detection and Estimation,* 2nd ed. New York: Springer-Verlag, 1994.

[22] H.V. Poor, "Turbo multiuser detection: A primer," *J. Commun. Networks,* vol. 3, no. 3, pp. 196–201, 2001.

[23] H.V. Poor, "Dynamic programming in digital communications: Viterbi decoding to turbo multiuser detection," *J. Optim. Theory Appl.,* vol. 115, no. 3, pp. 629–657, 2002.

[24] H.V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory,* vol. 43, pp. 858–871, 1997.

[25] J.G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill: 2000.

[26] D. Reynolds and X. Wang, "Low-complexity turbo-equalization for diversity channels," *Signal Processing,* vol. 81, pp. 989–995, 2001.

[27] D. Reynolds and X. Wang, "Turbo multiuser detection with unknown interferers," *IEEE Trans. Commun.,* vol. 50, no. 4, pp. 616–622, 2002.

[28] A. Roumy, I. Fijalkow, and D. Pirez, "Low cost turbo receiver for coded asynchronous DS-CDMA in frequency-selective channels" *J. Commun. Networks,* vol. 3, pp. 202–210, 2001.

[29] M. Sellathura and S. Haykin, "TURBO-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Processing,* vol. 50, no. 10, pp. 2538–2546, Oct. 2002.

[30] R. Koetter, A.C. Singer, and M. Tüchler, "Turbo equalization," *IEEE Signal Processing Mag.,* vol. 21, no. 1, pp. 67–80, Jan. 2004.

[31] S. Verdú, *Multiuser Detection.* Cambridge, UK: Cambridge Univ. Press: 1998.

[32] X. Wang and H.V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.,* vol. 47, no. 7, pp. 1046–1061, 1999.

[33] X. Wang and H.V. Poor, "Turbo multiuser detection for turbo-coded CDMA," in *Proc. 1999 IEEE Wireless Communications and Networking Conf.,* New Orleans, LA, Oct. 1999.

[34] X. Wang and H.V. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception.* Upper Saddle River, NJ: Prentice-Hall, 2004.

[35] Z.N. Wu and J.M. Cioffi, "Low-complexity iterative decoding with decision-aided equalization for magnetic recording channels," *IEEE J. Select. Areas Commun.,* vol. 19, no. 4, pp. 699–708, 2001.

[36] Z.G. Yang and X. Wang, "Blind turbo multiuser detection for long-code multipath CDMA," *IEEE Trans. Commun.,* vol. 50, no. 1, pp. 112–125, 2002.