Turbo Multiuser Detection for Coded DS-CDMA Systems: A Gibbs Sampling Approach

Vincent Buchoux and Olivier Cappé and Éric Moulines
ENST - Image and Signal Processing Laboratory (TSI dpt)
46 rue Barrault, 75634 Paris Cedex 13, France
buchoux,cappe,moulines@tsi.enst.fr

Abstract

This paper deals with joint detection and decoding techniques for coded CDMA (Code Division Multiple Access) systems. A promising approach in this context consists of combining the results of a soft output multiuser detector (MUD) with single user soft-input soft-output (SISO) decoders in an iterative fashion (so called "turbo" principle). In a first part of the paper we describe the CDMA channel under the form of a probabilistic graphical model (also known as Bayesian, or belief, network) which provides a very generic and natural way of deriving turbo algorithms. The structure of the algorithm is then obtained by direct application of general probability propagation rules rather than by using the context dependent notions of intrinsic and extrinsic information.

It turns out however that the obtained algorithm still requires soft output multiuser detection in a pseudo model where the symbols emitted by the user are a priori independent, which is not computationally feasible. The second part of the paper describes a simulation based MUD scheme which draws upon recent advances in Markov Chain Monte Carlo methods. The performance of the overall turbo multiuser decoder is compared with that of a state-of-the-art algorithm with comparable computational cost.

Keywords: *CDMA, Multiuser Detection, Turbo Decoding, Graphical Models, Markov Chain Monte Carlo, Gibbs Sampler, Rao-Blackwellization*

1. Introduction

Multiuser detection is the crux for high-spectral efficiency in digital communication systems using CDMA. As recently advocated in many contributions [7, 9], single user performance can almost be reached in a coded CDMA system by coupling joint multiuser detection and error-correcting coding in an iterative mode inspired from the turbo-decoding principle [3]. These turbo multiuser strategies however require efficient approximations to the optimal SISO MUD such as those proposed by Wang and Poor [12] and Alexander *et al* [1].

In this paper, we consider a probabilistic graphical representation of the dispersive CDMA channel. In the graphical model framework, the structure of the turbo decoder can be derived by applying generic probability propagation rules [4], [6]. We use this approach to justify the fact that the SISO joint multiuser detector must be designed so as to estimate the marginal posterior probabilities of the emitted information symbols given the observed signal frame, in a pseudo model where the symbols are a priori independent, with prior probabilities coming from the bank of single user decoders. It is well known that this task is not computationally feasible when the number of users is not very small (say, less than four), particularly in the presence of dispersive fading. We thus propose to use a simulation based SISO MUD which uses a user-by-user and symbol-by-symbol *Gibbs sampling approach*. The use of Markov Chain Monte-Carlo (MCMC) methods (like Gibbs sampling) in digital communication settings has been proposed recently by [11]. The main contributions of this paper are an evaluation of the simulation based approach in a realistic setting, in particular with a very low number of MCMC simulations, as well as an improved scheme for estimating the posterior probabilities at the SISO MUD stage based on a *Rao-Blackwellization* procedure [10].

2. Problem Statement

2.1. Signal Model

We consider a K-users direct-sequence code-division multiple access (DS-CDMA) with channel coding over a multipath channel. Assuming that the baseband received signal is fractionally sampled and/or received with spatial diversity, the discrete time signal at time index t writes

$$\mathbf{r}^t \triangleq \left[r^t_{(1)}, \cdots, r^t_{(N)}\right]^{\mathrm{T}}$$

where the superscript T denotes transposition and N is the equivalent number of polyphase components, including oversampling and multi-sensors reception. In a frequency selective multipath environment, the CDMA channel takes the form of a finite length MIMO linear filtering

$$\mathbf{r}^t = \sum_{l=0}^{L} \mathbf{H}^l\, \mathbf{s}^{t-l} + \mathbf{w}^t \tag{1}$$

$$= \mathbf{H}\,\mathbf{S}^t + \mathbf{w}^t \tag{2}$$

where s^t_k denotes the symbol of user k at time t, and where $\mathbf{s}^t \triangleq [s^t_1, \cdots, s^t_K]^{\mathrm{T}}$. $\mathbf{w}^t$ is an additive white and circular Gaussian noise of variance $\sigma^2\mathbf{I}_N$, and $\mathbf{H}^l$ denotes the $N \times K$ MIMO filter coefficients (in symbol period) of the L-order convolutive channel. We also denote by $\mathbf{H} \triangleq (\mathbf{H}^L, \cdots, \mathbf{H}^0)$ and $\mathbf{S}^t \triangleq \left[(\mathbf{s}^{t-L})^{\mathrm{T}}, \cdots, (\mathbf{s}^t)^{\mathrm{T}}\right]^{\mathrm{T}}$, the matrix and vector forms of the channel parameter and the channel state respectively. Each coefficient $\mathbf{H}^l$ contains the polyphase components of the lth channel coefficient of each user, *i.e.*

$$\mathbf{H}^l \triangleq \left[\mathbf{h}^l_1, \cdots, \mathbf{h}^l_K\right]$$

with $\mathbf{h}^l_k \triangleq [h^l_{(1),k}, \cdots, h^l_{(N),k}]^{\mathrm{T}}$. Thus, we may consider the received signal as a noisy mixture of the K users, as shown on figure 1. We assume that the symbols $(s^t_k)_{1\le k\le K}$ result from the modulation of q consecutive interleaved bits b^t_k in a finite and complex alphabet $\mathcal{S}$ of cardinal $Q \triangleq 2^q$. The code bits $(b^t_k)_{1\le k\le K}$ are the interleaved version of the output of K convolutional codes denoted $\mathcal{C}_k$, fed by K binary information sources $(d^t_k)_{1\le k\le K}$. In the sequel, we

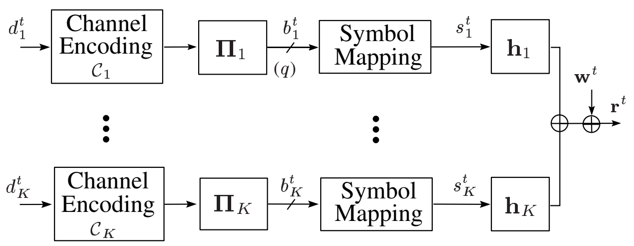


Figure 1. Equivalent Coded CDMA Channel

assume a M-symbols frame transmission leading to an observed signal denoted $\mathbf{r}^{1:M} \triangleq [(\mathbf{r}^1)^{\mathrm{T}}, \cdots, (\mathbf{r}^{M+L})^{\mathrm{T}}]^{\mathrm{T}}$ following the linear model

$$\mathbf{r}^{1:M} = \boldsymbol{\mathcal{H}}\,\mathbf{s}^{1:M} + \mathbf{w}^{1:M} \tag{3}$$

where $\mathbf{s}^{1:M} \triangleq [(\mathbf{s}^1)^{\mathrm{T}}, \cdots, (\mathbf{s}^M)^{\mathrm{T}}]^{\mathrm{T}}$ and where $\boldsymbol{\mathcal{H}}$ is a $N(M+L) \times KM$ filtering matrix composed of shifted versions of $\mathbf{H}$ on its rows. Moreover, we assume that the transmission of the symbols $\mathbf{s}^{1:M}$ corresponds to the transmission of M_d information bits.

2.2. Iterative joint detection and decoding

The optimal joint decoder-an-detector proceeds by finding the marginal maximum a posteriori estimate $\{\hat{d}^t_{k,\mathrm{MAP}}\}_{1\le t\le M_d}$ for each user, given the received signal $\mathbf{r}^{1:M}$ and the code structure $\boldsymbol{\mathcal{C}}$. Because this operation is computationally feasible only for non-interleaved systems with very small number of users, channel order and code constraint length [5], turbo schemes [8, 1, 12] appear has an appealing solution since they usually lead to near optimal performance with a manageable computation load. To described turbo algorithms it is standard to refer to the notion of "intrinsic" and "extrinsic" information although these are clearly context dependent an somewhat ill-defined in general settings. In this contribution, we adopt the point of view of Frey [4], McEliece [6] *et al* who consider the turbo principle as an instance of the probability propagation algorithm. To describe the CDMA channel, we use the representation advocated by Frey [4] ("factor graph") in which the circled symbols denote random variables, the dark dots correspond to a priori or conditional distributions, and the edges stand for conditional independence relations.

2.3. Graphical Model of the CDMA Channel

We represent the coded frame-based CDMA system described in section 2.1 by the graphical model given in figure 2, where $p(\mathbf{r}^{1:M}\,|\,\mathbf{s}^{1:M})$ denotes the conditional distribution of the observations $\mathbf{r}^{1:M}$ given the symbols $\mathbf{s}^{1:M}$ and $\mathcal{C}_k$ denotes a prior code structure on the symbols $s^{1:T}_k$ (see figure 3). From eq. (3), we may write

$$p(\mathbf{r}^{1:M}\,|\,\mathbf{s}^{1:M}) = \frac{1}{(\pi\sigma^2)^{(M+L)}} \exp\left(\frac{-\left\|\mathbf{r}^{1:M} - \boldsymbol{\mathcal{H}}\,\mathbf{s}^{1:M}\right\|^2}{\sigma^2}\right) \tag{4}$$

Making the convolutive structure of eq. (1) appear in the above equation would be useless since, even in the simplest case (*i.e.* when there is no inter-symbol interference), the required state space has dimension Q^K, and, for a L-order convolutive channel, the space dimension increases to Q^{KL} (see eq. (2)). Applying the probability propagation rules [4] to the graph of figure 2, one obtains an iterative scheme which consists of "messages"[1] passing be-

[1] Although these are quite often referred to a APP (for a posteriori probabilities), these messages are not always interpretable as probabilities : in a graph without cycle, they either correspond to an unnormalized probability distribution or to a likelihood defined up to an unknown normalizing constant. In a graph with cycles, such as the one we are considering, the messages have no direct interpretation in term of probabilities.

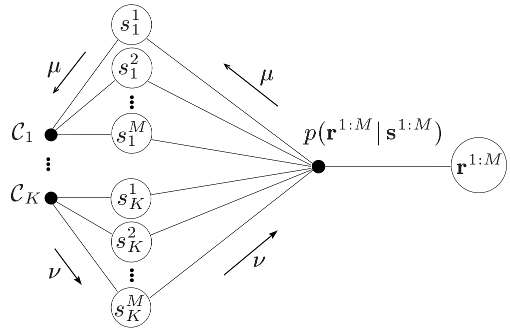


Figure 2. Graphical Bayesian Model of the CDMA system

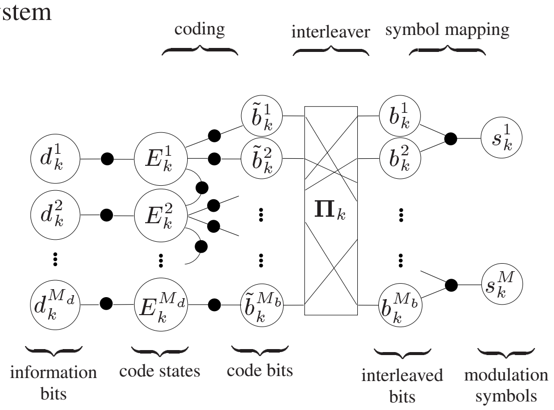


Figure 3. Graphical representation of the code structure $\mathcal{C}_k$

tween the multiple access channel and the K code structures $\boldsymbol{\mathcal{C}} \triangleq (\mathcal{C}_k)_{1 \le k \le K}$. The messages are either propagated from the multiple access channel to the symbols s_k^t (μ messages on figure 2) or in the reverse direction (μ on figure 2). The probability propagation rules for messages of type μ and ν write

$$\left\{ \begin{array}{lll} (i) & \mu_k^t(s_k^t) = & \displaystyle\sum_{(t',k') \neq (t,k)} p(\mathbf{r}^{1:M} | \, \mathbf{s}^{1:M}) \prod_{(t',k') \neq (t \neq k)} \nu_{k'}^{t'}(s_{k'}^{t'}) \\ (ii) & \nu_k^t(s_k^t) = & \displaystyle\sum_{t' \neq t} \mathcal{C}_k(\mathbf{s}_k^{1:M}) \prod_{t' \neq t} \mu_k^{t'}(s_k^{t'}) \end{array} \right.$$

Both (i) and (ii) have the usual interpretation

- if $\nu_k^t(s_k^t)$ ($1 \le k \le K$ and $1 \le t \le M$), are independent priors on symbols as illustrated in the pseudo model given in figure 4(a), then the messages μ_k^t are given by

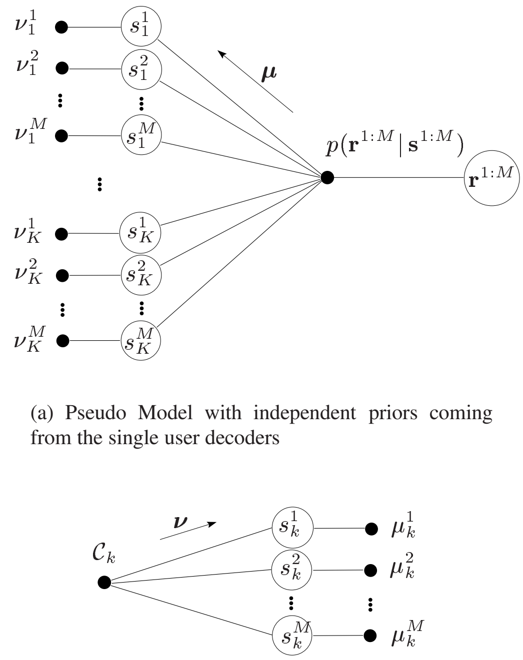
$$\mu_k^t(s_k^t) \quad \propto \quad P(s_k^t | \boldsymbol{\nu}, \mathbf{r}^{1:M}) / \nu_k^t(s_k^t) \tag{5}$$

- if $\mu_k^t(s_k^t)$ (for $1 \le t \le M$) are interpreted as the likelihood of pseudo observations conditionally independent given $\mathbf{s}_k^{1:M}$, then the messages ν_k^t are given by

$$\nu_k^t(s_k^t) \quad \propto \quad P(s_k^t | \boldsymbol{\mu}, \mathcal{C}_k) / \mu_k^t(s_k^t) \tag{6}$$

Because figure 4(b) corresponds to the structure of a Hidden Markov Model (see also figure 3 which shows the detailed graphical representation of the code structure $\mathcal{C}_k$), eq. (6) can be implemented efficiently for a given code $\mathcal{C}_k$ using the *forward-backward algorithm* [2].

Note that the messages $\mu_k^t(s_k^t)$, at the first iteration, are computed from equi-distributed initial values of the priors $\nu_k^t(s_k^t)$. The main problem consists nevertheless in estimating $\mu_k^t(s_k^t)$, or equivalently, the posterior probability $P(s_k^t | \boldsymbol{\nu}, \mathbf{r}^{1:M})$ in model 4(a).



(a) Pseudo Model with independent priors coming from the single user decoders

(b) Pseudo Model with independent likelihood terms coming from the MUD

Figure 4. Message propagation in the CDMA graphical model

3. A Sampling-Based Approach

Given independent priors $\nu_k^t(s_k^t)$ on symbols s_k^t, we propose to estimate the posterior distribution $P(s_k^t | \boldsymbol{\nu}, \mathbf{r}^{1:M})$ through a particular Markov Chain Monte-Carlo (MCMC) technique, namely the *Gibbs sampler*. MCMC [10] is a class of powerful techniques which make it possible to estimate posterior quantities in high-dimensional models. MCMC is a stochastic approximation procedure, comparable in many respects to conventional Monte Carlo except for the fact that successive simulations are Markov dependent.

3.1. The Gibbs sampling approach

For Bayesian probabilistic inference, all quantities of interest may be written as $g_{\mathbf{Y}} \triangleq \mathbb{E}\{g(x_1, \cdots, x_T)|\mathbf{Y}\}$ where $g(.)$ is a function of some random variables $x_{1:T}$ and $\mathbf{Y}$ are the observations. The Gibbs sampler approximates $g_{\mathbf{Y}}$ by constructing an homogeneous Markov chain $\{x_{1:T}^{(n)}\}$ according to the following mechanism

- `for` $t = 1, 2, \cdots, T$
  `Sample` $x_t^{(n)}$ `from` $p(x_t|x_{1:t-1}^{(n)}, x_{t+1:T}^{(n-1)}, \mathbf{Y})$
- `end (repeated for` $n = 1 \ldots$`)`

The resulting sequence $\{x_{1:T}^{(n)}\}_{n>0}$ is a sample path of a Markov chain which has $p(x_{1:T}|\mathbf{Y})$ as its stationary distribution. Under some regularity conditions, $g_{\mathbf{Y}}$ may be estimated using the ergodic theorem by

$$\widehat{g}_{\mathbf{Y}} = \frac{1}{n_s} \sum_{n=0}^{n_s} g(x_{1:T}^{(n)}) \xrightarrow[n_s \to \infty]{} g_{\mathbf{Y}} \quad \text{(with prob. 1)} \quad (7)$$

3.2. Application to Joint Multiuser Detection

Let $\mathbf{s}_{-k}^{-t}$ denote the sequence $\mathbf{s}^{1:M}$ in which the symbol s_k^t has been suppressed. The Gibbs sampling update at simulation index $n+1$ (simulation of $\mathbf{s}^{(n+1)}$ given $\mathbf{s}^{(n)}$) consists of

- `for` $t = 1, 2, \cdots, M$
  - `for` $k = 1, 2, \cdots, K$
  `Sample` $s_k^{t\,(n+1)}$ `from` $P(s_k^t|\mathbf{s}_{-k}^{-t\,(n+1)}, \boldsymbol{\nu}, \mathbf{r}^{1:M})$
  - `end (for k)`
- `end (for t)`

where $\mathbf{s}_{-k}^{-t\,(n+1)}$ denotes the incompletely updated sequence at simulation index $n+1$ and time index t, which contains both newly simulated symbols (those with time indexes smaller t) and symbols taken from the previously simulated sequence $\mathbf{s}^{(n)}$ (for the time indexes greater than t) :

$$\mathbf{s}_{-k}^{-t\,(n+1)} \quad \triangleq \quad (\, \mathbf{s}_{1:t-1}^{(n+1)} \,,\, \mathbf{s}_{1:k-1,t}^{(n+1)} \,,\, \mathbf{s}_{k+1:K,t}^{(n)} \,,\, \mathbf{s}_{t+1:M}^{(n)} \,)$$

The posterior distribution $P(s_k^t|\mathbf{s}_{-k}^{-t\,(n+1)}, \boldsymbol{\nu}, \mathbf{r}^{1:M})$ is given by

$$\begin{aligned} & P(s_k^t|\mathbf{s}_{-k}^{-t\,(n+1)}, \boldsymbol{\nu}, \mathbf{r}^{1:M}) \\ & \quad = \frac{1}{C}\, p(\mathbf{r}^{1:M}|s_k^t, \mathbf{s}_{-k}^{-t\,(n+1)}) \underbrace{p(s_k^t|\mathcal{C})}_{\nu_k^t(s_k^t)} \end{aligned} \quad (8)$$

where C is a normalizing constant obtained by summation over the possible values of s_k^t of the right-hand side of the expression. Expression (8) is not as computationally demanding as expected, since the possible values of s_k^t are sharing common factors in $p(\mathbf{r}^{1:M}|s_k^t, \mathbf{s}_{-k}^{-t\,(n+1)})$ that vanish due to the normalizing constant C. Without any loss of information, one can replace $p(\mathbf{r}^{1:M}|s_k^t, \mathbf{s}_{-k}^{-t\,(n+1)})$ by $p(\mathbf{r}^{t:t+L}|s_k^t, \mathbf{s}_{-k}^{-t\,(n+1)})$, which is computationally much less demanding and is of complexity $\mathcal{O}(n_s.K.N)$ per turbo iteration per user and per symbol.

Finally, the estimation of the posterior distributions is performed using a *Rao-Blackwell* procedure [10] : Instead of using $g(s) = 1$ if $s_k^t = s$ and (zero otherwise) to estimate $P(s_k^t = s|\boldsymbol{\nu}, \mathbf{r}^{1:M})$ from (7) as suggested by [11], we use $g(s) = P(s_k^t = s|\mathbf{s}_{-k}^{-t}, \boldsymbol{\nu}, \mathbf{r}^{1:M})$. This term is readily available since it is already computed during the Gibbs update as detailed in (8). For independent (ie. Monte Carlo) simulations, this choice would lead to a systematic decrease of the variance of the approximation error by virtue of the Rao-Blakwell theorem. For MCMC methods, this is not true anymore (because of the Markov dependence) but a similar gain is generally observed empirically [10]. For communications applications with a small number n_s of MCMC simulations, this procedure also avoids the risk of estimating posterior probabilities that are exactly 0 or 1 (in [11] this was achieved through the use of ad hoc thresholds).

Since we are considering cases where the number of MCMC simulations n_s is kept small, the initialization of the Gibbs sampler does play a significant role and we use a MMSE linear receiver inspired from [12] for the very first turbo iteration. For subsequent turbo iterations, the initial guess $\mathbf{s}^{(0)}$ is chosen as the marginal maximum a posteriori decision (given the current estimate of the symbol probabilities).

4. Simulations Results and Conclusion

Simulations results are illustrated on figure 5 for a $K = 8$-users coded DS-CDMA system using orthogonal spreading (OVSF) codes of spreading factor 8 and QPSK modulation. We use for each user a rate-$1/2$ convolutional code with octal generator polynomials $(23, 35)$, and pseudo-random interleavers of size 256. The CDMA channel is defined from a standardized ETRVA channel with 6-Rayleigh faded paths and a raised-cosine spectral emission pulse with roll-off 0.2, leading to a global channel order $L = 3$. An oversampling rate of $2/T_c$ and a single sensor antenna were used at the receiver, thus leading to $N = 16$. During the simulations, the channel was kept constant and is known to the receiver.

In order to illustrate the robustness of the proposed method to the near-far effect, the signal-to-noise ratio (SNR) of the first 4 users is kept constant and equal to 0dB. The Gibbs receiver is illustrated on figures 5(a) and 5(b) for 4 iterations of the turbo mode and $n_s = 50$ Gibbs simulations. The performance of the single user system on a synchronous AWGN channel is drawn in bold line, and the linear SISO MMSE receiver proposed in [12] are also presented on figures 5(c) and 5(d) for comparison purpose. The number of MCMC simulations ($n_s = 50$) was chosen intentionnally so that the Gibbs receiver presents a comparable complexity to the linear SISO MMSE for identical system caracteristics.

Figures 5(a) show that the Gibbs receiver is very efficient for the weak users in presence of a strong near-far effect, leading to near single-user performance in 2 iterations only, whereas the linear SISO MMSE receiver shows comparatively poor performance after 2 iterations. When users have equal 0dB power, single-user performance can nearly be reached after 3 iterations for the Gibbs which converges more rapidly than the MMSE receiver for which 4 iterations are needed.

In conclusion, the performance and the low cost of the Gibbs receiver show that MCMCs can be efficiently applied to turbo multiuser detection, even for heavy-loaded systems and severe transmission conditions. Proper initialization of the symbols and Rao-Blackwellization of the quantities of interest allows us to perform only few MCMC simulations and to still get a very robust behaviour.

References

[1] P. D. Alexander, M. C. Reed, J. A. Asenstorfer, and C. B. Schlegel. Iterative Multiuser Interference Reduction: Turbo CDMA. *IEEE Transactions on Communications*, 47(7):1008–1014, July 1999.

[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. on Inform. Theory*, 20(3):284–287, March 1974.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima. Near shannon limit error-correcting coding and decoding: turbocodes. In *Proceedings of ICC'93*, pages 1064–1070, Geneva, Switzerland, May 1993.

[4] B. Frey. *Graphical Models for Machine Learning and Digital Communication*. MIT Press, Cambridge, MA, 1998.

[5] T. Giallorenzi and S. Wilson. Multiuser ML Sequence Estimator for Convolutional Coded Asynchronous DS-CDMA Systems. *IEEE Transactions on Communications*, 44((8)):997–1008, August 1996.

[6] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng. Turbo Decoding as Instance of Pearl's "Belief Propagation" Algorithm. *IEEE Journal on Selected Areas in Communications*, 16(2):140–152, February 1998.

[7] M. Moher. Turbo-Based Multiuser Detection. In *Proc. IEEE Int. Symp. on Information Theory*, page 195, Ulm, Germany, June 1997.

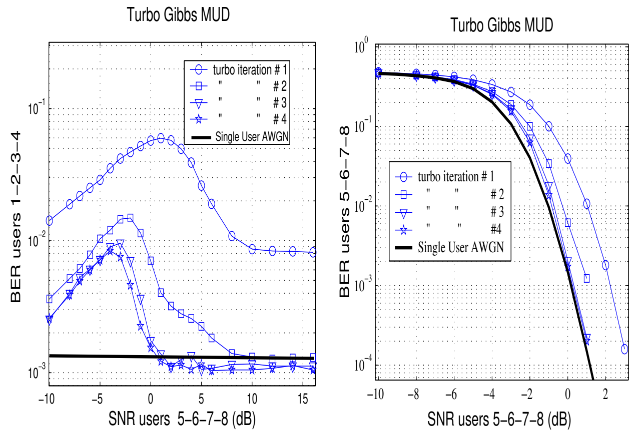
[8] M. Moher. An Iterative Multiuser Decoder for Near-Capacity Communications. *IEEE Transactions on Communications*, 46:870–880, July 1998.

[9] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer. Iterative Multiuser Detection for DS-CDMA with FEC. In *The International Symposium on Turbo Codes and Related Topics*, pages 162–165, Brest, France, September 1997.

[10] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 1999.

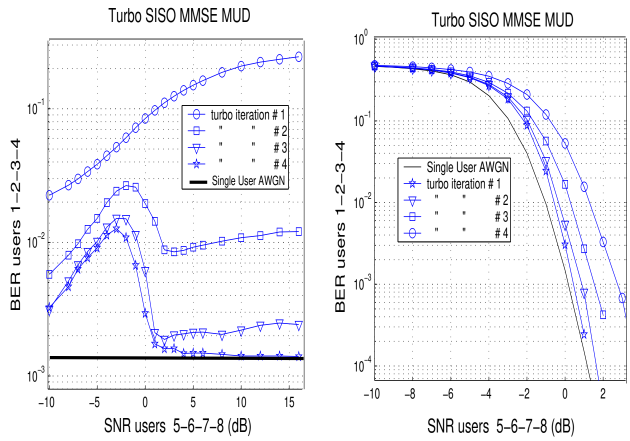
[11] X. Wang and R. Cheng. Adaptive Bayesian Multiuser Detection for Synchronous CDMA with Gaussian and Impulsive Noise. *IEEE Transactions on Signal Processing*, 47(7):2013, July 2000.

[12] X. Wang and H. V. Poor. Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA. *IEEE Transactions on Communications*, 46(7):1046–1061, July 1999.



(a) constant users

(b) varying users



(c) constant users

(d) varying users

Figure 5. Performance of the turbo Gibbs MUD and the SISO MMSE MUD in a 8-users dispersive CDMA system. SNRs of users $1 - 4$ is fixed to 0dB.