MIMO DETECTION USING MARKOV CHAIN MONTE CARLO TECHNIQUES FOR NEAR-CAPACITY PERFORMANCE

Haidong Zhu, Zhenning Shi+, and Behrouz Farhang-Boroujeny**

*ECE Department, University of Utah, USA,
e-mails: haidongz@eng.utah.edu and farhang@ece.utah.edu
+Wireless Signal Processing Group, NICTA, Australia, email: zhenning.shi@nicta.com.au

ABSTRACT

In this paper, we develop a new soft-in soft-out (SISO) multiple-input multiple-output (MIMO) detection algorithm using the Markov chain Monte Carlo (MCMC) simulation techniques and study its performance when applied to a MIMO communication system. Comparison with the best MIMO detection algorithm in the current literature, the sphere decoding, show that the proposed detection algorithm can improve the gap between the present results and the capacity by as much as 2 dB.

1. INTRODUCTION

Transmission through multiple transmit and receive antennas, known as multiple-input multiple-output (MIMO) communication, has been widely studied in recent years [1, 2, 3]. MIMO communication promises an increase in the channel capacity proportional to the minimum of the number of transmit and receive antennas. Furthermore, because of presence of alternative channel paths, MIMO channels are very reliable and robust to fading effects. The challenge in realizing the very high capacity of MIMO communication systems lies in development of effective detection algorithms. Among many detection algorithms that have been proposed in the past, the sphere decoding method of Hochwald and ten Brink [4] is the one with the closest performance to the channel capacity.

In this paper, we present a novel detection algorithm based on the Markov chain Monte Carlo (MCMC) simulation techniques [5], and through simulations show that it outperforms the sphere decoding of [4] by as much as 2 dB.

2. CHANNEL MODEL

We consider a flat fading channel model whose input and output are related according to the equation

$$\mathbf{y} = \mathbf{H}\mathbf{d} + \mathbf{n} \tag{1}$$

Z. Shi is with National ICT Australia and affiliated with the Australian National University. National ICT Australia is funded through the Australian Government's *Backing Australia's Ability* initiative and in part through the Australian Research Council.

where $\mathbf{d}$ is the vector of transmit symbols, $\mathbf{n}$ is the channel additive noise vector, $\mathbf{y}$ is the received signal vector, and $\mathbf{H}$ is the channel gain matrix. The elements of $\mathbf{H}$ are the channel gains between transmit and receive antennas. Assuming that there are N_t transmit and N_r receive antennas, $\mathbf{d}$ has a length of N_t, $\mathbf{y}$ and $\mathbf{n}$ have a length of N_r and $\mathbf{H}$ is an $N_\mathrm{r} \times N_\mathrm{t}$ matrix. We assume that each element of $\mathbf{d}$ is an L-ary symbol and takes values from the the alphabet $\mathcal{A} = \{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_L\}$. We assume that $\mathbf{n}$ is an iid Gaussian sequence with the autocorrelation matrix $E[\mathbf{n}\mathbf{n}^H] = \sigma_n^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Through out this paper, we assume that the channel gain matrix $\mathbf{H}$ and the noise variance σ_n^2 are perfectly known to the receiver.

We note that the channel model (1) repeats for transmission of the successive values of $\mathbf{d}$. Hence, there should be a time index attached to all terms in (1). We avoid such a time index here for brevity.

3. ITERATIVE MIMO DETECTION

We consider an iterative MIMO detector similar to the one discussed in [4]. Fig. 1 presents the block diagram of such a detector. It consists of a soft-in soft-out (SISO) MIMO detector and a SISO channel decoder. The MIMO detector generates a set of soft output sequences for the data symbols $d_1, d_2, \cdots, d_{N_\mathrm{t}}$ (the elements of $\mathbf{d}$) based on the observed input vector $\mathbf{y}$ and the *a priori* (soft) information from the latest iteration of the channel decoder. After subtracting the *a priori* information from the output of the MIMO detector, the remaining information which is new (extrinsic) to the channel decoder is passed over for further processing. Similarly, the soft input information to the channel decoder is subtracted from its output to generate the new (extrinsic) information before being fed back to the MIMO detector.

The soft information that is exchanged between the MIMO detector and the channel decoder are the likelihood values (L-values) of transmitted information bits or symbols. The L-values are the ratios of the symbol probabilities as commonly defined in the literature [6]. We continue our discussion with an evaluation of symbol probabilities and through that demonstrate the challenge of estimating L-values in the

MIMO detector.

Given $\mathbf{y}$ and the *a priori* information $\boldsymbol{\lambda}_2^{\mathrm{e}}$, where $\boldsymbol{\lambda}_2^{\mathrm{e}}$ denotes the set of all *a priori* (extrinsic) information (say, the probabilities or L-values) that we have about $\mathbf{d}$ from the channel decoder, we wish to calculate the conditional probability $P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})$, for $k = 1, 2, \cdots, N_{\mathrm{t}}$ and the L possible choices of α_i. To see the difficulty of this problem, we note that

$$\begin{aligned} P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}) &= \sum_{\mathbf{d}_{-k}} P\left(d_k = \alpha_i, \mathbf{d}_{-k}|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}\right) \\ &= \sum_{\mathbf{d}_{-k}} P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}, \boldsymbol{\lambda}_2^{\mathrm{e}}) P(\mathbf{d}_{-k}|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}), \quad (2) \end{aligned}$$

where $\mathbf{d}_{-k} = [d_1 \ \ldots \ d_{k-1} \ d_{k+1} \ \cdots \ d_{N_{\mathrm{t}}}]^T$, the second identity follows by applying the chain rule [7], and the summation is over all possible values of $\mathbf{d}_{-k}$. The number of combinations that $\mathbf{d}_{-k}$ takes grows exponentially with N_{t} and thus may become prohibitive as N_{t} grows.

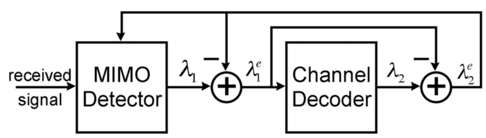

Fig. 1. Receiver structure. The MIMO detector and channel decoder are SISO blocks that exchange soft information, $\boldsymbol{\lambda}_1^{\mathrm{e}}$ and $\boldsymbol{\lambda}_2^{\mathrm{e}}$, in a turbo loop.

4. SATISTICAL ESTIMATION OF PROBABILITIES

In a recent work [8], we have shown that the prohibitive complexity of (2) can be avoided by adopting the Monte Carlo statistical methods, where accurate estimate of the summations is obtained by taking a small, but proper, samples of $\mathbf{d}_{-k}$. Gibbs samples that are obtained through a MCMC simulator of the channel model (1) are used for this purpose. The MCMC simulator begins with a random selection of $\mathbf{d}$, choosing the elements of $\mathbf{d}$ randomly according to their respective extrinsic information (probabilities). It then proceeds with changing one of the elements of $\mathbf{d}$ at a time. Say, d_k is selected and it is given one of the possible values from the alphabet $\mathcal{A}$ according to their respective probabilities. In other words, d_k is set equal to α_i according to the distribution $P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_{-k}^{\mathrm{e}})$. More particularly, to obtain the desired samples, the Gibbs sampler runs through the following routine.

- Initialize $\mathbf{d}^{(-N_{\mathrm{b}})}$
- for $n = -N_{\mathrm{b}} + 1$ to N_{s}
  draw $d_1^{(n)}$ from $P(d_1|d_2^{(n-1)}, \ldots, d_{N_{\mathrm{t}}}^{(n-1)}, \mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})$
  draw $d_2^{(n)}$ from $P(d_2|d_1^{(n)}, d_3^{(n-1)}, \ldots, d_{N_{\mathrm{t}}}^{(n-1)}, \mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})$
  $\vdots$
  draw $d_{N_{\mathrm{t}}}^{(n)}$ from $P(d_{N_{\mathrm{t}}}|d_1^{(n)}, \ldots, d_{N_{\mathrm{t}}-1}^{(n)}, \mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})$

The Gibbs sampler is run for $N_{\mathrm{b}} + N_{\mathrm{s}}$ iterations. From this the first N_{b} iterations are used to allow Markov chain to converge to its steady state so that proper samples of $\mathbf{d}$ could be collected afterwards. During the next N_{s} iterations, i.e., for $n = 1$ to N_{s}, the Gibbs sampler generates N_{s} samples of each symbol d_k when $\mathbf{d}_{-k}$, $\mathbf{y}$ and $\boldsymbol{\lambda}_2^{\mathrm{e}}$ are given, for $k = 1, 2, \cdots, N_{\mathrm{t}}$. These samples are used to obtain an estimate of $P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})$ using the approximation

$$P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}) \approx \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\mathrm{e}}) \quad (3)$$

where $\mathbf{d}_{-k}^{(n)} = [d_1^{(n)} \ \cdots \ d_{k-1}^{(n)} \ d_{k+1}^{(n-1)} \ \cdots \ d_{N_{\mathrm{t}}}^{(n-1)}]^T$ and $\mathbf{d}_{-k}^{(n)}$, in (3) and other equations that follow, is a shorthand notation for $\mathbf{d}_{-k} = \mathbf{d}_{-k}^{(n)}$.

The approximation (3) is obtained through direct application of Monte Carlo integration [5] to (2). In this paper, we propose the following approximation:

$$P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}) \approx \frac{\sum_{n=1}^{N_{\mathrm{s}}} P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\mathrm{e}}) P(\mathbf{d}_{-k}^{(n)}|\mathbf{y}, \boldsymbol{\lambda}_2^{e})}{\sum_{n=1}^{N_{\mathrm{s}}} P(\mathbf{d}_{-k}^{(n)}|\mathbf{y}, \boldsymbol{\lambda}_2^{e})}. \quad (4)$$

This approximation is obtained from the theory of importance sampling [5], and assuming that $\mathbf{d}_{-k}^{(n)}$ are chosen from a set of equally distributed but important samples of $\mathbf{d}_{-k}$. Such important samples are obtained by running Gibbs sampler and deleting repetitions of samples $\mathbf{d}_{-k}$. Computer simulations presented in Section 6 show that (4) has much better performance than (3) .

5. COMPUTATION OF EXTRINSIC L-VALUES

We assume that the alphabet size L is a power of 2. This means each data symbol d_k carries $J = \log_2 L$ bits of coded information, where J bis an integer. Let $b_j(d_k)$ denote the jth bit of d_k, and $U_j^+(\mathcal{A})$ and $U_j^-(\mathcal{A})$ the subsets of the alphabet $\mathcal{A}$ in which the jth bit of each element is $+1$ and -1, respectively. Accordingly, the L-value of $b_j(d_k)$ at the MIMO detector output is obtained as

$$\lambda_1(b_j(d_k)) = \ln \frac{\sum_{U_j^+(\mathcal{A})} P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})}{\sum_{U_j^-(\mathcal{A})} P(d_k = \alpha_i|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}})} \quad (5)$$

Substituting (4) in (5), we obtain (6), shown at the top of the next page. Using the Bayes rule, we get

$$\begin{aligned} P(\mathbf{d}_{-k}^{(n)}|\mathbf{y}, \boldsymbol{\lambda}_2^{\mathrm{e}}) &= \frac{p(\mathbf{d}_{-k}^{(n)}, \mathbf{y}|\boldsymbol{\lambda}_2^{\mathrm{e}})}{p(\mathbf{y}|\boldsymbol{\lambda}_2^{\mathrm{e}})} \\ &= \frac{p(\mathbf{y}|\mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\mathrm{e}}) P^{\mathrm{e}}(\mathbf{d}_{-k}^{(n)})}{p(\mathbf{y}|\boldsymbol{\lambda}_2^{\mathrm{e}})} \quad (7) \end{aligned}$$

$$\lambda_1(b_j(d_k)) = \ln \frac{\sum_{U_j^+(\mathcal{A})} \sum_{n=1}^{N_s} P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\rm e}) P(\mathbf{d}_{-k}^{(n)}|\mathbf{y}, \boldsymbol{\lambda}_2^{\rm e})}{\sum_{U_j^-(\mathcal{A})} \sum_{n=1}^{N_s} P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\rm e}) P(\mathbf{d}_{-k}^{(n)}|\mathbf{y}, \boldsymbol{\lambda}_2^{\rm e})} \tag{6}$$

$$\lambda_1(b_j(d_k)) = \ln \frac{\sum_{U_j^+(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} p(\mathbf{y}|\mathbf{d}_{k,i}^{(n)}) P^{\rm e}(\mathbf{d}_{-k}^{(n)}) P_{-j}^{\rm e}(d_k = \alpha_i)}{\sum_{U_j^-(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} p(\mathbf{y}|\mathbf{d}_{k,i}^{(n)}) P^{\rm e}(\mathbf{d}_{-k}^{(n)}) P_{-j}^{\rm e}(d_k = \alpha_i)} \cdot \frac{P^{\rm e}(b_j(d_k) = +1)}{P^{\rm e}(b_j(d_k) = -1)} \tag{10}$$

$$\lambda_1^{\rm e}(b_j(d_k)) = \ln \frac{\sum_{U_j^+(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} p(\mathbf{y}|\mathbf{d}^{(n)}) P^{\rm e}(\mathbf{d}_{-k}^{(n)}) P_{-j}^{\rm e}(d_k = \alpha_i)}{\sum_{U_j^-(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} p(\mathbf{y}|\mathbf{d}^{(n)}) P^{\rm e}(\mathbf{d}_{-k}^{(n)}) P_{-j}^{\rm e}(d_k = \alpha_i)}. \tag{11}$$

where $P^{\rm e}(\mathbf{d}_{-k}^{(n)})$ is a shorthand notation for $P(\mathbf{d}_{-k}^{(n)}|\boldsymbol{\lambda}_2^{\rm e})$, i.e., the probability of $\mathbf{d}_{-k} = \mathbf{d}_{-k}^{(n)}$ given the available extrinsic information. Similarly, we obtain

$$P(d_k = \alpha_i|\mathbf{y}, \mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\rm e}) = \frac{p(\mathbf{y}|\mathbf{d}_{-k}^{(n)}, d_k = \alpha_i) P^{\rm e}(d_k = \alpha_i)}{p(\mathbf{y}|\mathbf{d}_{-k}^{(n)}, \boldsymbol{\lambda}_2^{\rm e})}. \tag{8}$$

Note that we use $P(\cdot)$ and $p(\cdot)$ to denote probability and probability density functions, respectively. Morerover, because of the interleaving effect, the extrinsic bit information from the channel decoder may be assumed to be independent of each others. Hence,

$$P^{\rm e}(d_k = \alpha_i) = \prod_{l=1}^{J} P^{\rm e}(b_l(d_k) = b_l(\alpha_i)). \tag{9}$$

Substituting (7), (8) and (9) in (6), we obtain (10) at the top of this page, where $\mathbf{d}_{k,i}^{(n)} = [d_1^{(n)} \cdots d_{k-1}^{(n)}\ \alpha_i\ d_{k+1}^{(n-1)} \cdots d_{N_{\rm t}}^{(n-1)}]^T$ and $P_{-j}^{\rm e}(d_k = \alpha_i) = \prod_{l=1, l\neq j}^{L} P^{\rm e}(b_l(d_k) = b_l(\alpha_i))$. Note that we have removed $\boldsymbol{\lambda}_2^{\rm e}$ from the first terms under summations in (10), because when $\mathbf{d}$ is fully specified the extrinsic information become irrelevant. On the other hand, recalling that $\lambda_1(b_j(d_k)) = \lambda_1^{\rm e}(b_j(d_k)) + \lambda_2^{\rm e}(b_j(d_k))$ and $\lambda_2^{\rm e}(b_j(d_k)) = \ln \frac{P^{\rm e}(b_j(d_k)=+1)}{P^{\rm e}(b_j(d_k)=-1)}$, we obtain, from (10), (11), also shown at the top of this page.

When the channel gain matrix $\mathbf{H}$ is known and the noise vector $\mathbf{n}$ is Gaussian and satisfies $E[\mathbf{n}\mathbf{n}^H] = \sigma_n^2 \mathbf{I}$, $p(\mathbf{y}|\mathbf{d}_{k,i}^{(n)}) = \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-||\mathbf{y}-\mathbf{H}\mathbf{d}_{k,i}^{(n)}||^2/2\sigma_n^2}$.

As in turbo decoders [6], to avoid numerical instability, a log-domain implementation of (11) may be used. To this end, we define $\eta_{k,i,j}^{(n)} = \ln\left(P^{\rm e}(\mathbf{d}_{-k}^{(n)}) P_{-j}^{\rm e}(d_k = \alpha_i)\right) - \frac{||\mathbf{y}-\mathbf{H}\mathbf{d}^{(n)}||^2}{2\sigma_n^2}$ and note that (11) can be rearranged as

$$\lambda_1^{\rm e}(b_j(d_k)) = \ln \sum_{U_j^+(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} e^{\eta_{k,i,j}^{(n)}} - \ln \sum_{U_j^-(\mathcal{A})} \sum_{n=1}^{N_{\rm s}} e^{\eta_{k,i,j}^{(n)}} \tag{12}$$

which can performed in a computationally efficient manner by using the identity $\ln(e^x + e^y) = \max(x, y) + \ln(1 + e^{|x-y|})$ and following the standard methods [6].

Alternatively, we can adopt a max-log-MAP type approach [6] and use the following approximation:

$$\lambda_1^{\rm e}(b_j(d_k)) \approx \max_{U_j^+(\mathcal{A}),n} \eta_{k,i,j}^{(n)} - \max_{U_j^-(\mathcal{A}),n} \eta_{k,i,j}^{(n)}. \tag{13}$$

Since this approximation incur very little loss in performance, but results in considerably less complex algorithm, the simulation results presented in the next section are based on (13).

6. COMPUTER SIMULATIONS

We have studied the performance of both (3) and (4) through computer simulations. The general conclusion that could be derived from these simulations is that (4) outperforms (3) by a significant gap. We have also compared the proposed MCMC detection algorithm (implemented using (4)) with the best available detection algorithms [4, 9] and found that the proposed algorithm outperforms these methods. Moreover, the MCMC algorithm is directly applicable to the important case where $N_{\rm t} > N_{\rm r}$. The detection methods of [4] and [9], on the other hand, are only appropriate in the case where $N_{\rm t} \leq N_{\rm r}$. Extensions of these method to the case $N_{\rm t} > N_{\rm r}$ although possible, is not recommended. This has been clearly noted in [4] and the use of space-time codes of [10] has been recommended. In this paper, because of the limited space, we limit ourselves to two sets of simulation results (i) a comparison of (3) and (4), and (ii) a comparison of the proposed MIMO detector with the sphere decoding of [4].

To compare the relative behavior of (3) and (4), we evaluated the extrinsic L-values of the information bits at the output of the MIMO detector for $N_{\rm t} = 4$ and $N_{\rm r} = 2$. Quadrature phase shift keying (QPSK) symbols are used. The channel code is a rate 1/2 convolutional code with the generator polynomials $1+D+D^2$ and $1+D^2$. Estimator (3) is based on Gibbs samples from a single Markov chain with $N_{\rm b} = 20$ and $N_{\rm s}$ varied from 1 to 160. The estimator (4), on the other hand, is based on a set of 3 Markov chains run in parallel. Samples the three chains are used in (4). For each case the exact L-values are calculated and compared with the estimated values. The difference between the exact and approximate L-values are squared and ensemble averaged over a large number of the channel realizations to obtain an estimate of the mean square error (MSE) of the estimates. The SNR at the receiver input is set equal to 8 dB. The mutual information between the bit L-values and the transmitted information is set equal to 0.2. The results, shown in Fig. 2, clearly show the superior performance of the estimator (4).

Fig. 3 compares the BER results of the MCMC detector with the sphere decoding of [4]. The results are for a MIMO channel with $N_\mathrm{t} = 8$ and $N_\mathrm{r} = 8$ and constellation sizes of QPSK, 16-QAM and 64-QAM. The channel code is the turbo code used in [4]. As seen, the MCMC is able to reduce the gap between the BER curves and the capacity limit by about 2 dB for the cases of 16-QAM and 64-QAM. For the case of QPSK, the MCMC detector performs similar to the sphere decoding which in this case searches over all possible values of symbols, i.e., it is an exact MAP detector [4].

7. CONCLUSION

We presented a new MCMC detection algorithm which was found to perform much superior to a recent work [8]. This algorithm also was compared with the sphere decoding of [4] and its superior performance was demonstrated through simulations. We did not discuss the complexity issues, because of the limited space. However, at this point we wish to comment that the MCMC detector that we proposed in this paper has significantly lower complexity than sphere decoding. Moreover, the ability to run parallel Gibbs samplers make this method very appropriate for VLSI implementation where parallel processing can be very useful in implementing system with very fast throughputs.

8. REFERENCES

[1] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1 (2), August 1996, pp. 41 – 59.

[2] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment," *Wirelss Commun. Mag.*, vol. 6, pp. 311-335, March 1998.

[3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, November – December 1999.

[4] B.M. Hochwald, and S. ten Brink, "Achieving near-capacity on a multiple-anatenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, March 2003.

[5] C.P. Robert and G. Casella, *Monte Carlo Statistical Methods.* Springer-Verlag, New York, 1999.

[6] P. Robertson, E. Villebrun, and P Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," ICC'95, 18-22 June 1995, vol. 2, pp. 1009 - 1013.

[7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* 3rd Ed., McGraw-Hill.

[8] Z. Shi, H. Zhu, and B. Farhang-Boroujeny, "Markov chain Monte Carlo Techniques in Iterative Detectors: A Novel Approach Based on Monte Carlo Integration," Accepted for presentation in Globecom 2004.

[9] H.E. Gamal and A.R. Hammons, Jr., "A new approach to layered space-time coding and signal processing," *IEEE Trans Info. Theory,* pp. 2321 - 2334, Sept. 2001.

[10] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory,* vol. 48, pp. 1804-1824, July 2002.

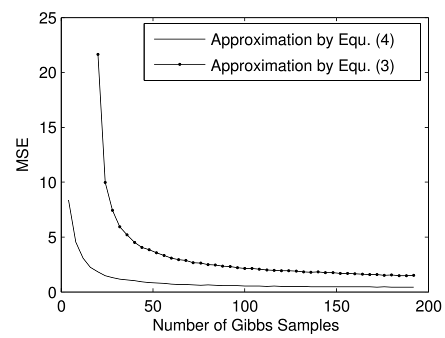

Fig. 2. Comparison of (3) and (4). Presented are the MSE of the L-value estimates as a function of the number of Gibbs samples.

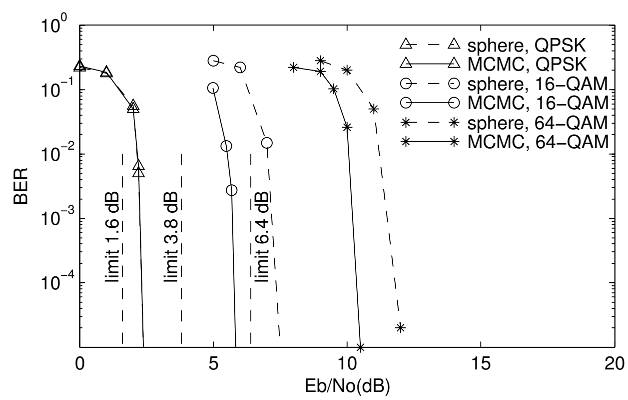

Fig. 3. BER results of the MCMC detector and sphere decoding.