Improved Large-MIMO Detection Based on Damped Belief Propagation

Pritam Som, Tanumay Datta, A. Chockalingam, and B. Sundar Rajan

Department of ECE, Indian Institute of Science, Bangalore 560012, INDIA

Abstract— **In this paper, we consider the application of belief propagation (BP) to achieve near-optimal signal detection in *large* multiple-input multiple-output (MIMO) systems at low complexities. Large-MIMO architectures based on spatial multiplexing (V-BLAST) as well as non-orthogonal space-time block codes (STBC) from cyclic division algebra (CDA) are considered. We adopt graphical models based on Markov random fields (MRF) and factor graphs (FG). In the MRF based approach, we use pairwise compatibility functions although the graphical models of MIMO systems are fully/densely connected. In the FG approach, we employ a Gaussian approximation (GA) of the multi-antenna interference, which significantly reduces the complexity while achieving very good performance for large dimensions. We show that *i*) both MRF and FG based BP approaches exhibit *large-system behavior*, where increasingly closer to optimal performance is achieved with increasing number of dimensions, and *ii*) *damping* of messages/beliefs significantly improves the bit error performance.**

Keywords − *Large-MIMO signal detection, V-BLAST, non-orthogonal STBC, near-optimal performance, belief propagation, damping.*

I. INTRODUCTION

Multiple-input multiple-output (MIMO) systems that employ large number of transmit and receive antennas can offer very high spectral efficiencies of the order of tens to hundreds of bps/Hz [1],[2]. Achieving near-optimal signal detection at low complexities in such large-dimension systems has been a challenge. In our recent works, we have shown that certain algorithms from machine learning/artificial intelligence achieve near-optimal performance in large-MIMO systems that employ tens of transmit and receive antennas using V-BLAST and non-orthogonal space-time block codes (STBC) from cyclic division algebra [3] with tens to hundreds of dimensions in space and time, at low complexities. Such algorithms include local neighborhood search based algorithms like a *likelihood ascent search* (LAS) algorithm [4],[5] and a *reactive tabu search* (RTS) algorithm [6], and algorithms based on *probabilistic data association* (PDA) [7] and *belief propagation* (BP) [8][1]. In this paper, we present extensions to our BP based large-MIMO signal detection work in [8].

In systems characterized by fully/highly connected graphical models, BP based algorithms [9] may fail to converge, and if they do converge, the estimated marginals may be far from exact [12],[13]. It may be expected that BP might perform poorly in MIMO graphs due to the high density of connections. However, several methods are known in the literature, including and *double loop methods* [14],[15] and *damping* [16],[17], which can be applied to improve things if BP does not converge (or converges too slowly). In [14], Heskes *et al* proposed the double-loop algorithm that is provably convergent for the minimization of Bethe and Kikuchi free energies

---

[1]Similar algorithms have been earlier reported in the context of multiuser detection.

[9] that represent the cost function of BP. In [18], Pretti and Pelizzola proposed a new propagation algorithm for the minimization of the cost function (Bethe free energy) for a generic lattice model with pair interactions. The algorithm is shown to be more stable than BP, as it reaches a fixed point also for highly frustrated systems and faster than the provably convergent double loop algorithms.

In [16], Pretti proposed a modified version of BP with over-relaxed BP dynamics. At each step of the algorithm, the evaluation of messages is taken to be a weighted average between the old estimate and the new estimate. The weighted average could either be applied to the messages (resulting in *message damped BP*) or to the estimate of the probability distribution/beliefs of the variables (*probability/belief damped BP*), or to both messages and beliefs (*hybrid damped BP*). It is shown, in [16], that the probability damping BP can be derived as a limit case in which the double-loop algorithm becomes a single-loop one. We, in this paper, show that damping is quite effective in achieving good performance in large-MIMO signal detection.

In [8], we showed that BP on Markov random field (MRF) based graphical models of MIMO systems exhibits *large-system behavior*, where the bit error performance improves with increasing number of dimensions and approaches near-optimal performance for large number of dimensions. The use of pairwise compatibility functions in the MRFs was instrumental in achieving such good performance at low complexities. Our contributions in this paper are two-fold:

- First, we further improve the performance achieved in [8] through the use of belief/message damping.
- Next, we present an alternate approach based on factor graphs (FG) [9], where we approximate the multi-antenna interference as Gaussian. We show that this FG approach also exhibits large-system behavior. The Gaussian approximation of the interference significantly reduces the complexity while achieving very good performance in detecting large-dimension MIMO signals.

The rest of this paper is organized as follows. In Sec. II, we present the system model. MRF based and FG based approaches for large-MIMO detection are presented in Secs. III and IV, respectively. Conclusions are given in Sec. V.

II. SYSTEM MODEL

MIMO systems with N_t transmit and N_r receive antennas employing spatial multiplexing (V-BLAST) and non-orthogonal space-time block codes (STBC) from cyclic division algebra (CDA) [3] achieve the full rate of N_t symbols per channel use. In addition, STBCs from CDA achieve full transmit diversity as well. We consider signal detection in both V-BLAST as well as non-orthogonal STBC MIMO systems.

Consider a V-BLAST MIMO system with N_t transmit antennas and N_r receive antennas. Let $\mathbf{x} \in \mathbb{A}^{N_t}$ denote the transmitted symbol vector, where $\mathbb{A}$ is the modulation alphabet. Let $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denote the channel gain matrix whose entries are modeled as $\mathbb{CN}(0,1)$. The received signal vector $\mathbf{y} \in \mathbb{C}^{N_r}$ is given by

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

where $\mathbf{n}$ is the noise vector whose entries are modeled as i.i.d $\mathbb{CN}(0, \sigma^2 = \frac{N_t E_s}{\gamma})$, where E_s is the average energy of the transmitted symbols and γ is the average received SNR per receive antenna.

A non-orthogonal STBC from CDA is an $N_t \times N_t$ matrix whose entries are formed using linear combinations of various data symbols [3]. Each STBC is constructed using N_t^2 data symbols and are sent in N_t channel uses. The received signal matrix can be vectorized and written in an equivalent real system model of the form (1), where the number of transmit and receive dimensions are $2N_t^2$ and $2N_tN_r$, respectively, for QAM [5].

The goal is to obtain an estimate of the transmitted symbol vector $\mathbf{x}$, given $\mathbf{y}$ and the knowledge of $\mathbf{H}$. The optimal maximum aposteriori probability (MAP) detector enumerates the joint posterior distribution

$$p(\mathbf{x}|\mathbf{y}, \mathbf{H}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{H}) p(\mathbf{x}), \tag{2}$$

and marginalizes out each variable as

$$p(x_i|\mathbf{y}, \mathbf{H}) = \sum_{x_{-i}} p(\mathbf{x}|\mathbf{y}, \mathbf{H}), \tag{3}$$

where x_{-i} stands for all entries of $\mathbf{x}$ except x_i. The MAP estimate of the bit x_i, $i = 1, \cdots, S$, is then given by

$$\widehat{x}_i = \begin{array}{c} \text{arg max} \\ a \in \mathbb{A} \end{array} p(x_i = a \mid \mathbf{y}, \mathbf{H}). \tag{4}$$

We assume that the channel gain matrix is known perfectly at the receiver, but not at the transmitter.

III. DETECTION USING MRF APPROACH

BP is a technique that solves inference problems using graphical models [9]. Well known graphical models include Bayesian belief networks, factor graphs and MRFs. In this section, we present a detection scheme based BP on MRFs.

An MRF is an undirected graph whose vertices are random variables [10],[11]. Usually, the variables in an MRF are constrained by a *compatibility function*, also known as a *clique potential* in literature. An MRF is called a pairwise MRF, if the clique potentials are all functions of two variables. The clique potentials can then be denoted as $\psi_{i,j}(x_i, x_j)$, where x_i, x_j are variables connected by an edge in the MRF. Consider a pairwise MRF in which the x_i's denote underlying *hidden* variables on which the observed variables y_i's are dependent. Let the dependence between the hidden variable x_i and the explicit variable y_i be represented by a *joint* compatibility function $\phi_i(x_i, y_i)$ (also called as 'evidence' of x_i).

The joint distribution of the hidden and explicit variables is given by

$$p(\mathbf{x}, \mathbf{y}) \propto \prod_{i,j} \psi_{i,j}(x_i, x_j) \prod_i \phi_i(x_i, y_i). \tag{5}$$

Since y_i's are fixed, we can drop them in the above equation and write

$$p(\mathbf{x}) \propto \prod_{i,j} \psi_{i,j}(x_i, x_j) \prod_i \phi_i(x_i). \tag{6}$$

In [8], even though the MRF of a MIMO system is fully/highly connected (where the clique size is large), we defined clique potentials as functions of only pairs of neighboring variables, i.e., the clique potentials are denoted as $\psi_{i,j}(x_i, x_j)$, where x_i and x_j are variables connected by an edge in the MRF. This is to say that the clique potentials are defined similar to those in pairwise MRFs [9] (where all cliques are of size two), although the MRF of the system itself is fully/highly connected and hence not a pairwise MRF. For the vector channel model in 1, the following compatibility functions are adopted in [8]:

$$\phi_i(x_i) = \exp\left(x_i^H \Re\{z_i\} + \ln\{p(x_i)\}\right) \tag{7}$$
$$\psi_{i,j}(x_i, x_j) = \exp\left(-x_i^H \Re\{R_{i,j}\} x_j\right), \tag{8}$$

where z_i is the ith element of the vector $\mathbf{z} \triangleq \frac{1}{\sigma^2}\mathbf{H}^H\mathbf{y}$, and R_{ij} is the element in the ith row and jth column for the matrix $\mathbf{R} \triangleq \frac{1}{\sigma^2}\mathbf{H}^H\mathbf{H}$. Message passing is carried out on this graphical model to compute the marginal probabilities of the variables. The belief at node i is

$$b_i(x_i) \propto \phi_i(x_i) \prod_{j \in \mathcal{N}(i)} m_{j,i}(x_i), \tag{9}$$

where $\mathcal{N}(i)$ denotes the neighboring nodes of node i, and the messages are defined as [9]

$$m_{j,i}(x_i) \propto \sum_{x_j} \phi_j(x_j) \psi_{j,i}(x_j, x_i) \prod_{k \in \mathcal{N}(j)\setminus i} m_{k,j}(x_j). \tag{10}$$

A. Improvement through Message/Belief Damping

In [16], Pretti proposed a modified version of BP with over-relaxed BP dynamics. At each step of the algorithm, the evaluation of messages is taken to be a weighted average between the old estimate and the new estimate. The weighted average could either be applied to the messages (resulting in *message damped BP*) or to the estimate of the probability distribution/beliefs of the variables (*probability/belief damped BP*), or to both messages and beliefs (*hybrid damped BP*).

Message Damped BP: Denoting $\widetilde{m}_{i,j}^{(t)}(x_j)$ as the updated message in iteration t obtained by message passing, the new message from node i to node j in iteration t, denoted by $m_{i,j}^{(t)}(x_j)$, is computed as a convex combination of the old message and the updated message as

$$\widetilde{m}_{i,j}^{(t)}(x_j) \propto \sum_{x_i} \phi_i(x_i) \psi_{i,j}(x_i, x_j) \prod_{k \in \mathcal{N}(i)\setminus j} m_{k,i}^{(t-1)}(x_i), \tag{11}$$

$$m_{i,j}^{(t)}(x_j) = \alpha_m m_{i,j}^{(t-1)}(x_j) + (1 - \alpha_m)\widetilde{m}_{i,j}^{(t)}(x_j), \tag{12}$$

where $\alpha_m \in [0,1)$ is referred as the *message damping factor*.
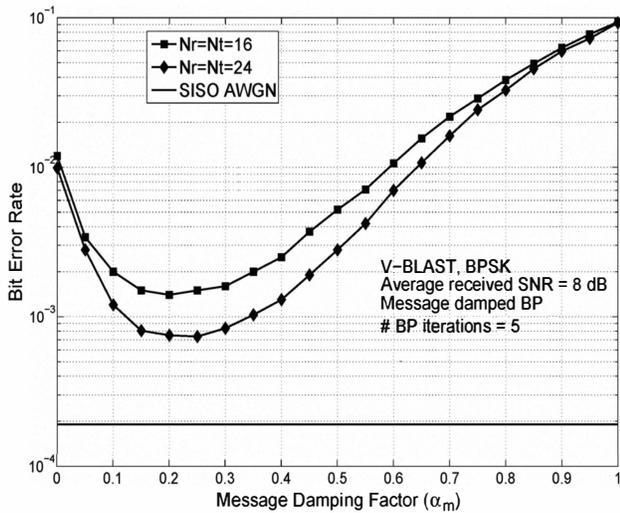
Fig. 1. BER performance of the MRF based message damped BP algorithm as a function of message damping factor, α_m, in V-BLAST with $N_t = N_r = 16$ at 8 dB SNR and BPSK. # BP iterations = 5.

Fig. 2. BER performance of the MRF based message damped BP algorithm as a function of SNR in V-BLAST for different $N_t = N_r = 16$ and BPSK. # BP iterations = 5, $\alpha_m = 0.2$.

Belief Damping: Instead of damping the messages in each iteration, the beliefs of the variables can be computed in each iteration as a weighted average, as

$$\widetilde{\mathsf{b}}_i^{(t)}(x_i) \quad \propto \quad \phi_i(x_i) \prod_{j \in \mathcal{N}(i)} m_{j,i}^{(t)}(x_i), \qquad (13)$$

$$\mathsf{b}_i^{(t)}(x_i) \quad = \quad \alpha_b \, \mathsf{b}_i^{(t-1)}(x_i) + (1 - \alpha_b) \, \widetilde{\mathsf{b}}_i^{(t)}(x_i), \quad (14)$$

where $\alpha_b \in [0, 1)$ is referred to as the *belief damping factor*.

Hybrid Damping: As a more general damping strategy, we can update both the messages as well as the beliefs according to (12) and (14), respectively, in each iteration. Different combinations of (α_m, α_b) values specializes to different strategies; for e.g., $(\alpha_m = \alpha_b = 0)$ corresponds to undamped BP, $(\alpha_m \neq 0, \alpha_b = 0)$ corresponds to message damped BP, $(\alpha_m = 0, \alpha_b \neq 0)$ corresponds to belief damped BP, and $(\alpha_m \neq 0, \alpha_b \neq 0)$ corresponds to hybrid damped BP.

The above damping operations do not increase the order of complexity of the algorithm, which, for $N_t = N_r$, is $O(N_t^2)$ per-symbol complexity for V-BLAST and $O(N_t^4)$ per-symbol complexity for non-orthogonal STBC [8].

B. Simulation Results

In Figs. 1 to 3, we present the simulated BER performance of the algorithm focusing on the effect of damping on the performance. The number of BP iterations is set to 5 in all these figures. Figure 1 shows the the variation of BER as a function of the message damping factor, α_m, for V-BLAST with $N_t = N_r = 16, 24$ and SNR = 8 dB. Note that $\alpha_m = 0$ corresponds to the case of undamped BP. It can be observed from Fig. 1 that, depending on the choice of the value of α_m, message damping can significantly improve the BER performance of the BP algorithm. There is an optimum value of α_m at which the BER improvement over no damping case is maximum. For the chosen set of system parameters in Fig. 1, the optimum value of α_m is observed to be about 0.2. For
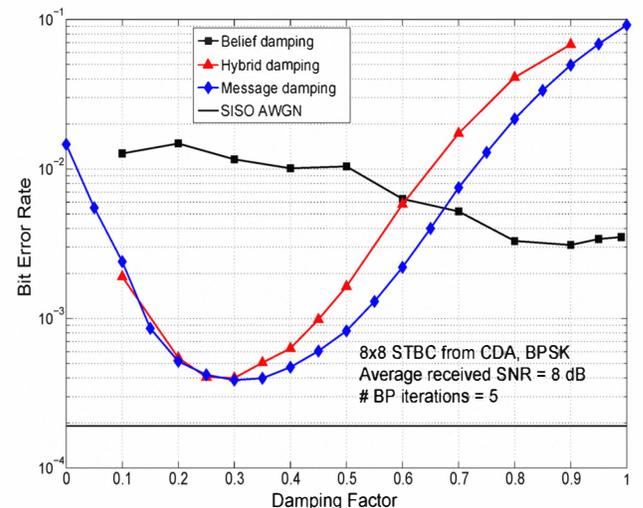
Fig. 3. Effect of message, belief, and hybrid damping on the BER performance of 8×8 STBC from CDA with $t = e^{\mathbf{j}}$, $\delta = e^{\sqrt{5}\mathbf{j}}$ at 8 dB SNR. BPSK, $N_t = N_r$, # BP iterations = 5, $\alpha_m = \alpha_b$ for hybrid damping.

this optimum value of $\alpha_m = 0.2$, it is observed that about an order of BER improvement is achieved with message damping compared to that without damping. From Fig. 1, it can further be seen that the performance improves for increasing $N_t = N_r$ (i.e., performance of the $N_t = N_r = 24$ system is better that of the $N_t = N_r = 16$ system). This illustrates the large system behavior of the algorithm, where the performance moves more towards SISO AWGN performance when $N_t = N_r$ is increased from 16 to 24. The large-system behavior of the algorithm is further illustrated in Fig. 2, where we plot the BER performance of V-BLAST as a function of SNR for different $N_t = N_r = 4, 8, 16, 24$ and 32 for $\alpha_m = 0.2$.

In Fig. 3, we present a comparison of the BER performance achieved using message damping, belief damping and hybrid damping based BP detection of 8×8 non-orthogonal STBC from CDA with $t = e^{\mathbf{j}}$, $\delta = e^{\sqrt{5}\mathbf{j}}$ [3] at 8 dB SNR. For message damping and belief damping, α_m and α_b are varied in the range 0 to 1. For hybrid damping, we set $\alpha_m = \alpha_b$ and

varied it in the range 0 to 1. From Fig. 3, it can be seen that i) with damping, there is an optimum value of the damping factor at which the BER performance is the best (e.g., for message damping, the optimum damping factor is about 0.3 in Fig. 3), ii) message damping performs better than belief damping for small values of the damping factor, whereas belief damping performs better at high values of the damping factor; however, over the entire range of the damping factor, the best performance of message damping is significantly better than the best performance of belief damping, and iii) for the chosen condition of $\alpha_m = \alpha_b$, hybrid damping performance is similar to that of message damping; however, if the damping factors α_m and α_b are jointly optimized in hybrid damping, then the best performance with hybrid damping is expected to be better than the best performances of both message damping and belief damping.

IV. DETECTION USING FACTOR GRAPH APPROACH

In this section, we present a detection scheme based on BP on factor graphs of MIMO systems. Consider the linear vector channel model in (1). Each entry of the observation vector $\mathbf{y}$ will be treated as a function node (observation node), and each transmitted symbol will be a variable node. In each iteration of the algorithm, messages are passed back and forth the variable nodes and the observation nodes. The received signal at the ith receive antenna can be written as

$$
\begin{aligned}
y_i &= \sum_{j=1}^{N_t} h_{ij} x_j + n_i \\
&= h_{ik} x_k + \underbrace{\sum_{j=1, j\neq k}^{N_t} h_{ij} x_j + n_i}_{Interference}, \quad (15)
\end{aligned}
$$

where h_{ij} denotes the channel coefficient from jth transmit to ith receive antenna. We approximate the interference from other streams to be Gaussian. That is,

$$
y_i = h_{ik} x_k + \underbrace{\sum_{j=1, j\neq k}^{N_t} h_{ij} x_j + n_i}_{\triangleq z_{ik}}, \quad (16)
$$

where the interference plus noise term, z_{ik}, is modeled as $\mathbb{CN}(\mu_{z_{ik}}, \sigma_{z_{ik}}^2)$ with

$$
\mu_{z_{ik}} = \sum_{j=1, j\neq k}^{N_t} h_{ij} \mathbb{E}(x_j), \quad (17)
$$

and

$$
\sigma_{z_{ik}}^2 = \sum_{j=1, j\neq k}^{N_t} |h_{ij}|^2 \operatorname{Var}(x_j) + \sigma^2. \quad (18)
$$

Now, for BPSK signaling, the log-likelihood ratio (LLR) of the symbol $x_k \in \{+1, -1\}$ at observation node i, denoted by Λ_i^k, can be written as

$$
\begin{aligned}
\Lambda_i^k &= \log \frac{p(y_i|\mathbf{H}, x_k = 1)}{p(y_i|\mathbf{H}, x_k = -1)} \\
&= \frac{4}{\sigma_{z_{ik}}^2} \Re\left(h_{ik}^*(y_i - \mu_{z_{ik}})\right). \quad (19)
\end{aligned}
$$
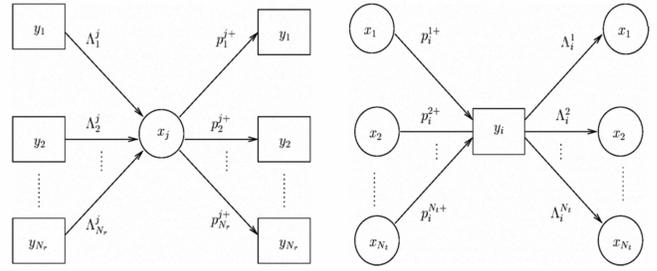
Fig. 4. Message passing between variable nodes and observation nodes.

The LLR values computed at the observation nodes are passed to the variable nodes (Fig. 4). Using these LLRs, the variable nodes compute the probabilities

$$
p_i^{k+} \triangleq p_i(x_k = +1|\mathbf{y}) = \frac{\exp(\sum_{l\neq i} \Lambda_l^k)}{1 + \exp(\sum_{l\neq i} \Lambda_l^k)}, \quad (20)
$$

and pass them back to the observation nodes (Fig. 4). This message passing is carried out for a certain number of iterations. Messages can be damped as described in Sec. III and then passed. At the end, x_k is detected as

$$
\widehat{x}_k = \operatorname{sgn}\left(\sum_{i=1}^{N_r} \Lambda_i^k\right). \quad (21)
$$

A. Computational Complexity

The computation complexity of the algorithm involves i) LLR calculations at the observation nodes as per (19), which has $O(N_t N_r)$ complexity, and ii) calculation of probabilities at variable nodes as per (20), which also requires $O(N_t N_r)$ complexity. Hence, the overall complexity of the algorithm is $O(N_t N_r)$ for detecting N_t number of transmitted symbols. So the per-symbol complexity is $O(N_t)$ for $N_t = Nr$. It is noted that this complexity is one order less than that of the MRF based approach in the previous section. This is because the computation of $\mathbf{H}^T\mathbf{H}$, which is needed in the MRF approach, is not needed in the FG approach due to the Gaussian approximation of the interference. The Gaussian approximation thus reduces the complexity significantly, while achieving very good performance as well in large dimensions (we'll see this in the following subsection on performance results). For non-orthogonal STBC, the complexity of FG based approach is $O(N_t^2)$ per-symbol.

B. Simulation Results

We evaluated the BER performance of the Gaussian approximation (GA) based BP algorithm through simulations. Figures 5 and 6 show uncoded BER performance for V-BLAST and non-orthogonal STBCs, respectively, for 4-QAM. For V-BLAST, the number of antennas considered are $N_t = N_r = 8, 16, 24, 32, 64$; V-BLAST with $N_t = 64$ and 4-QAM has 128 real dimensions). For non-orthogonal STBC MIMO, 4×4, 8×8, and 16×16 information lossless STBCs from CDA (with $\delta = t = 1$ [3]) are considered; 16×16 STBC from CDA with 4-QAM has 512 real dimensions[2]. The number of

---

[2]Since the simulation of MAP for such large number of dimensions is prohibitively complex, we have plotted the SISO AWGN performance as a lower bound for comparison.
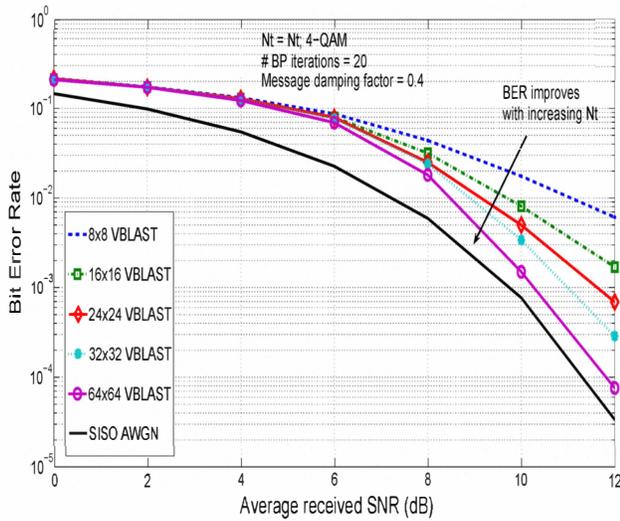
Fig. 5. BER performance of the FG based BP algorithm with Gaussian approximation of interference in large V-BLAST MIMO systems. $N_t = N_r = 8, 16, 24, 32, 64$, 4-QAM, # BP iterations = 20, $\alpha_m = 0.4$.
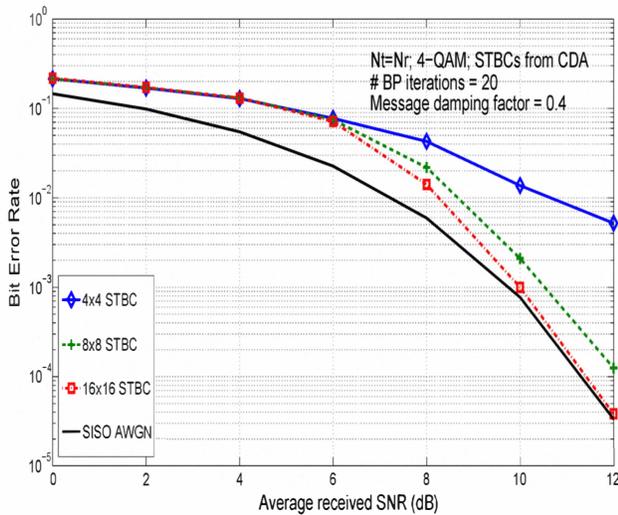
Fig. 6. BER performance of the FG based BP algorithm with Gaussian approximation of interference in large non-orthogonal STBC MIMO systems. $4 \times 4, 8 \times 8, 16 \times 16$ STBCs from CDA with $\delta = t = 1$. $N_t = N_r$. 4-QAM, # BP iterations = 20, $\alpha_m = 0.4$.

BP iterations used is 20. In our simulations, we found that a message damping factor of 0.4 is optimum for the system parameters considered. Accordingly, we have used 0.4 as the message damping factor in Figs. 5 and 6. From Figs. 5 and 6, we observe that the GA based algorithm exhibits large system behavior in both large V-BLAST as well as large non-orthogonal STBC MIMO systems. We have also evaluated the coded BER performance of the algorithm for 64×64 V-BLAST with 4-QAM and rate-1/2 turbo code (64 bps/Hz spectral efficiency), and found that the algorithm performs close to the theoretical minimum SNR for a 64×64 MIMO channel, to within about 4.5 dB.

V. CONCLUSIONS

We presented BP algorithms, based on Markov random field (MRF) and factor graph (FG) representations of MIMO systems, for signal detection in large-MIMO systems. The al-

gorithms exhibited large-system behavior, which makes BP a natural choice for detection in large-dimension systems including large-MIMO systems. Belief/message damping was shown to significantly improve the bit error performance. The MRF approach achieved reduced complexity through the use of pairwise compatibility functions. The FG approach employed Gaussian approximation (GA) of the multi-antenna interference. The GA is found to be very effective in reducing the complexity (by an order compared to that of the MRF approach), while achieving near-optimal performance as well when the number of dimensions is large. The illustrated feasibility of BP based algorithms for large-MIMO signal detection is significant, given that practical 12×12 V-BLAST MIMO systems operating at 50 bps/Hz have been already demonstrated [19], and 16×16, 24×24 and 32×32 MIMO systems can be potentially considered in wireless standards like IEEE 802.11 VHT and IEEE 802.16/LTE-A.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, November 1999.
[2] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.
[3] B. A. Sethuraman, B. Sundar Rajan, and V. Shashidhar, "Full-diversity high-rate space-time block codes from division algebras," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2596-2616, October 2003.
[4] K. Vishnu Vardhan, Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE JSAC Spl. Iss. on Multiuser Detection for Adv. Commun. Systems and Networks*, vol. 26, no. 3, pp. 473-485, April 2008.
[5] Saif K. Mohammed, Ahmed Zaki, A. Chockalingam, and B. Sundar Rajan, "High-rate space-time coded large-MIMO systems: Low-complexity detection and channel estimation," to appear in *IEEE Jl. Sel. Topics in Signal Processing (JSTSP): Spl. Iss. on Managing Complexity in Multiuser MIMO Systems*, December 2009. Online arXiv:0809.2446v3 [cs.IT] 16 Sept 2009.
[6] N. Srinidhi, Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search," *Proc. IEEE ISIT'2009*, Seoul, June-July 2009.
[7] Saif K. Mohammed, A. Chockalingam, and B. Sundar Rajan, "Low-complexity near-MAP decoding of large non-orthogonal STBCs using PDA," *Proc. IEEE ISIT'2009*, Seoul, June-July 2009.
[8] Madhekar Suneel, Pritam Som, A. Chockalingam, and B. Sundar Rajan, "Belief propagation based decoding of large non-orthogonal STBCs," *Proc. IEEE ISIT'2009*, Seoul, June-July 2009.
[9] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," *MERL Tech. Rep. TR-2001-22*, January 2002.
[10] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, California, 1988.
[11] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication*, Cambridge: MIT Press, 1998.
[12] J. M. Mooij, *Understanding and Improving Belief Propagation*, Ph.D Thesis, Radboud University Nijmegen, May 2008.
[13] J. M. Mooij and H. J. Kappen, "Sufficient conditions for convergence of the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4422-4437, December 2007.
[14] T. Heskes, K. Albers, and B. Kappen, "Approximate inference and constrained optimization," *Proc. Uncertainty in AI*, August 2003.
[15] A. L. Yuille, "A Double-Loop Algorithm to minimize Bethe and Kikuchi Free Energies", *Neural Computation*, 2002
[16] M. Pretti, "A message passing algorithm with damping," *Jl. Stat. Mech.: Theory and Practice*, November 2005.
[17] T. Heskes, "On the uniqueness of loopy belief propagation fixed points," *Neural Comput.*, vol. 16, no. 11, pp. 2379-2413, Nov. 2004.
[18] M. Pretti and A. Pelizzola, "Stable propagation algorithm for the minimization of the Bethe free energy," *Jl. Phys. A: Math. Gen.*, Nov. 2003.
[19] H. Taoka and K. Higuchi, "Field experiment on 5-Gbit/s ultra-high-speed packet transmission using MIMO multiplexing in broadband packet radio access," *NTT DoCoMo Tech. Journ.*, vol. 9, no. 2, pp. 25-31, September 2007.