Detection and Decoding in Large-Scale MIMO Systems: A Non-Binary
Belief Propagation Approach

T. Lakshmi Narasimhan and A. Chockalingam
Department of ECE, Indian Institute of Science, Bangalore 560012

Abstract—In this paper, we propose a non-binary belief
propagation approach (NB-BP) for detection of $M$-ary modula-
tion symbols and decoding of $q$-ary LDPC codes in large-scale
multiuser MIMO systems. We first propose a message passing
based symbol detection algorithm which computes vector mes-
sages using a scalar Gaussian approximation of interference,
which results in a total complexity of just $O(KN\sqrt{N})$, where
$K$ is the number of uplink users and $N$ is the number of
base station (BS) antennas. We then design optimized $q$-ary
LDPC codes by matching the EXIT charts of the proposed
detector and the LDPC decoder. Simulation results show that
the proposed NB-BP detection-decoding approach using the
optimized LDPC codes achieve significantly better performance
(by about 1 dB to 7 dB at $10^{-5}$ coded BER for various system
loading factors; the number of users ranging from 16 to 128
and number of BS antennas fixed at 128) compared to using
linear detectors and off-the-shelf $q$-ary irregular LDPC codes.

I. INTRODUCTION

Wireless communication systems using multiple-input
multiple-output (MIMO) configurations with a large number
of antennas have recently attracted a lot of research attention [1, 2]. These systems can achieve high spectral and
power efficiencies. An emerging architecture for large-scale
multiuser MIMO communications is one where each base
station (BS) is equipped with a large number of antennas
and the user terminals are equipped with one antenna each.
A key requirement on the uplink (user terminal to BS link)
in such large-scale MIMO systems is to achieve reduced
detection and decoding complexities at the BS receiver to
enable practical implementation, while maintaining good
performance. When the number of BS antennas is much
larger than the number of uplink users (i.e., low system
loading factors), linear detectors like the minimum mean
square error (MMSE) detector are good in terms of both
complexity and performance [3].

Belief propagation (BP) on graphs is a promising low-
complexity high-performance approach for signal processing
in large dimensions. Decoding of turbo codes and low den-
sity parity check (LDPC) codes, and equalization/detection
[4]-[7] are popular examples of the use of BP in commu-
nications. In [6], a MIMO detection algorithm for binary
modulation, based on approximate message passing on a
factor graph is presented. We refer to this algorithm in [6]
as binary-BP (B-BP) algorithm. The total complexity of
the B-BP algorithm is very low (quadratic in the number of
dimensions) because of its Gaussian approximation of
interference. Though the performance of the B-BP algorithm
in large dimensions is very good for binary modulation
(BPSK), its performance in higher-order QAM is rather poor
(we will see this in results/discussions in Sec. III). The
BP algorithm in [7] uses a different approach. It obtains
a tree that approximates the fully-connected MIMO graph
and performs message passing on this tree. The performance
of this detection algorithm for higher-order QAM is also far
from optimal.

The non-binary BP approach achieves good performance
at low complexities for $q$-ary LDPC codes [8]. In this
paper, we propose a non-binary BP (NB-BP) approach for
both detection as well as decoding which achieves very
good complexity and performance in large-scale multiuser
MIMO systems. Our new contributions in this paper can be
summarized as follows:

• First, we propose a NB-BP based detection algo-

rithm for $M$-ary modulation, where (i) the messages
passed between nodes are constructed as vector
messages, and (ii) the interference is approximated
as a scalar Gaussian random variable. While the
scalar approximation contributes to achieving very
low complexity (lower than MMSE complexity),
the vector nature of the messages contribute to
achieving close to optimal performance in large
dimensions.

• Next, through the EXIT curve matching technique,
we obtain $q$-ary LDPC codes that are optimized for
the proposed NB-BP detector and the LDPC de-
coder. These optimized irregular $q$-ary LDPC codes
with NB-BP detection outperform off-the-shelf ir-
regular $q$-ary LDPC codes with MMSE detection,
by 1 dB to 7 dB at $10^{-5}$ coded BER for various
system loading factors; the number of users is varied
from 16 to 128 and the number of BS antennas is
fixed at 128.

To our knowledge, non-binary BP for detection of $M$-ary
modulation and $q$-ary LDPC code optimization for large-
scale multiuser MIMO systems have not been reported so
far.

II. SYSTEM MODEL

Consider a large-scale multiuser MIMO system where
$K$ uplink users, each transmitting with a single antenna,
communicate with a BS having a large number of receive
antennas. Let $N$ denote the number of BS antennas; $N$ is
in the range of tens to hundreds. The ratio $\alpha = K/N$ is
the system loading factor. This system model is illustrated
in Fig. 1. Each user uses an LDPC code over $\text{GF}(q)$ and
$M$-QAM modulation. Each user encodes a sequence of $k\beta$
information bits to a sequence of $n$ coded symbols using
a q-ary LDPC code with parity check matrix $F$ and code rate $R = \frac{q}{K}$, where $\beta = \log_2 q$. These coded symbols are then $M$-QAM modulated and transmitted. Assume $M = q = 2^k$, where $k$ is an integer. The transmission of one LDPC code block requires $n$ channel uses. Let $H^{(t)}_c \in \mathbb{C}^{N \times K}$ denote the channel gain matrix in the $t$th channel use and $h^{(t)}_{ij}$ denote the complex channel gain from the $i$th user to the $j$th BS antenna. The channel gains $h^{(t)}_{ij}$'s are assumed to be independent Gaussian with zero mean and variance $\sigma^2_j$, such that $\sum_j \sigma^2_j = K$. Each $\sigma^2_j$ models the imbalance in the received power from user $j$ due to path loss etc., and $\sigma^2_j = 1$ corresponds to the case of perfect power control. Let $x^{(t)}_c$ denote the modulated symbol vector transmitted in the $t$th channel use, where the $j$th element of $x^{(t)}_c$ denotes the modulation symbol transmitted by the $j$th user. Assuming perfect synchronization, the received vector at the BS in the $t$th channel use, $y^{(t)}_c$, is given by

$$y^{(t)}_c = H^{(t)}_c x^{(t)}_c + w^{(t)}_c,$$

where $w^{(t)}_c$ is the noise vector whose entries are modeled as i.i.d. $CN(0, \sigma^2)$. Dropping the channel index for convenience, (1) can be written in the real domain as

$$y = Hx + w,$$

where $y \triangleq \begin{bmatrix} \Re(y_c) \\ \Im(y_c) \end{bmatrix}$, $H \triangleq \begin{bmatrix} \Re(H_c) - \Im(H_c) \\ \Im(H_c) + \Re(H_c) \end{bmatrix}$, $x \triangleq \begin{bmatrix} \Re(x_c) \\ \Im(x_c) \end{bmatrix}$, $w \triangleq \begin{bmatrix} \Re(w_c) \\ \Im(w_c) \end{bmatrix}$, and $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts, respectively. Note that for $M$-QAM modulation, the elements of $x$ come from the underlying PAM alphabet $\mathbb{A} = \{ \pm 1, \pm 3, \cdots, \pm \sqrt{M - 1} \}$. The BS observes $y$ and performs detection and decoding.

### III. Non-Binary BP for Detection

In this section, we propose a NB-BP scheme for the detection of $M$-QAM symbols suited well for large-scale MIMO systems. A key component of the scheme is the proposed construction of vector messages over a factor graph using a scalar Gaussian approximation of interference. While the scalar approximation contributes to achieving very low complexity, the vector nature of the messages contributes to achieving very good performance.

We model the system as a factor graph with $N$ observation nodes and $K$ variable nodes (see Fig. 2), and perform an approximate marginalization of the symbol probabilities over this graph. The $i$th element in the received vector $y$ can be written as

$$y_i = h_{ij} x_j + z_i, \quad i = 1, \cdots, 2N, \quad j = 1, \cdots, 2K,$$

where $z_i \triangleq \sum_{l=1, l \neq j}^{2K} h_{il} x_l + w_i$ is the interference-plus-noise term, $x_j$ is the $j$th element in $x$, $h_{ij}$ is the $(i,j)$th element in $H$, and $w_i$ is the $i$th element in $w$. As in [6], we approximate the scalar term $z_i$ as Gaussian r.v. with mean $(\mu_{ij})$ and variance $(\sigma^2_{ij})$, given by

$$\mu_{ij} = \sum_{l=1, l \neq j}^{2K} h_{il} E(x_l), \quad \sigma^2_{ij} = \sum_{l=1, l \neq j}^{2K} h_{il}^2 \text{Var}(x_l) + \sigma^2.$$

**Construction of vector messages:** Let $a_{ij}$ denote the message passed from the $i$th observation node to the $j$th variable node, and $v_{ji}$ denote the message passed from the $j$th variable node to $i$th observation node (see Fig. 2). $a_{ij}$ and $v_{ji}$ are vectors of size $\sqrt{M} \times 1$, and they are constructed to be functions of the approximate likelihood and posterior probabilities. Using the Gaussian approximation made above, the likelihood and the posterior probabilities can be approximated as

$$\Pr(y_i|H,x_j = s) \approx \frac{1}{\sigma_{ij} \sqrt{2\pi}} \exp \left( \frac{-(y_i - \mu_{ij} - h_{ij}s)^2}{2\sigma^2_{ij}} \right),$$

where $s \in \mathbb{A}$, and

$$\Pr(x_j = s|y, H) \approx \prod_{i=1}^{2N} \Pr(y_i|H, x_j = s) \approx \prod_{i=1}^{2N} \frac{\exp \left( \frac{-(y_i - \mu_{ij} - h_{ij}s)^2}{2\sigma^2_{ij}} \right)}{\sigma_{ij}},$$

respectively. With these approximations, the messages are defined as

$$a_{ij}(s) = \frac{1}{\sigma_{ij} \sqrt{2\pi}} \exp \left( \frac{-(y_i - \mu_{ij} - h_{ij}s)^2}{2\sigma^2_{ij}} \right),$$

$$v_{ji}(s) = \sum_{i=1, i \neq j}^{2N} a_{ij}(s),$$

where $a_{ij}(s)$ and $v_{ji}(s)$ are the elements of $a_{ij}$ and $v_{ji}$, respectively, corresponding to the symbol $s$. The mean and variance at the $i$th observation node are computed as

---

Remark: Although a similar scalar approximation of interference is used in [6], here the proposed formulation of messages as vectors achieves significantly improved $M$-QAM detection performance compared to the scalar messages based BP in [6].
Step 2) After a given number of iterations, the final symbol probabilities are made on the bit probability values computed using a damping factor

$$\mu_{ij} = \sum_{l=1, l \neq j}^{2K} h_{il}s^T v_{li},$$

(9)

$$\sigma_{ij}^2 = \sum_{l=1, l \neq j}^{2K} h_{il}^2 \left( \left( s \circ s \right)^T v_{li} - (s^T v_{li})^2 \right) + \sigma_e^2.$$  (10)

where $s$ is the vector of all elements in $\mathbb{A}$ (e.g., for $M = 16$, $s = [-3 - 1 + 1 + 3]^T$), and $\circ$ denotes the Hadamard product of vectors. Message passing:

Step 1) Initialize the posterior probability values $v_{ji}(s)$’s as $1/\sqrt{M}$.

Step 2) Compute $a_{ij}$ messages using (9), (10), and (7).

Step 3) Compute $v_{ji}$ messages using (8). This forms one iteration of the algorithm. Repeat Steps 2) and 3) for a certain number of iterations. Damping on $v_{ji}$ messages can be done in the $m$th iteration using a damping factor $\delta$

$$v_{ji}^{(m)} = (1 - \delta)v_{ji}^{(m-1)} + \delta v_{ji}^{(m-1)}, \quad \delta \in [0, 1].$$

(11)

After a given number of iterations, the final symbol probabilities are computed as

$$P_{x_j}(s) \triangleq \Pr(x_j = s) \propto \prod_{i=1}^{2N} a_{ij}(s).$$

(12)

A listing of the proposed NB-BP algorithm is listed in Algorithm 1. The $P_{x_j}(s)$ values $\forall s, j$ are fed as soft inputs to the $q$-ary LDPC decoder. In uncoded systems, hard bit decisions are made on the bit probability values computed as

$$\Pr(b_j^p = 1) = \sum_{s \in \mathbb{A} : p \text{th bit in } s = 1} P_{x_j}(s),$$

(13)

where $b_j^p$ is the $p$th bit in the $j$th user’s symbol.

Complexity: From (8), (9) and (10), the total complexity of the NB-BP scheme proposed above is $O(KN\sqrt{M})$. This is because the summations in (9), (10) can be computed by summing over all node indices once and subtracting from it the term to be excluded at each node. Note that the $O(KN\sqrt{M})$ complexity of the proposed scheme is much less compared to the MMSE detector complexity of $O(KN^2)$. This is because the MMSE detector needs a matrix inversion, whereas the proposed NB-BP scheme does not need a matrix inversion. More interestingly, as discussed next, even with this less than MMSE complexity, the proposed NB-BP scheme performs increasingly closer to optimal performance in large-scale MIMO systems.

BER Performance: Figure 3 shows the uncoded BER performance of the proposed NB-BP detector for 16-QAM in multiuser MIMO with $N = 32, 64, 128, 256, \alpha = 1$, and $\sigma_e = 1$. The performance of the B-BP detector in [6], MMSE detector, MF detector, and unfaded SISO AWGN performance are also plotted for comparison. For using the B-BP scheme in [6] for $M$-QAM detection, each $M$-QAM symbol is written in the form of linear combination of the constituent $q$ bits and the equivalent system model is written as $y = H(1_K \otimes m)x_b + w$, where $m = [2^0 2^1 \ldots 2^{q-1}]$, $x_b \in \{\pm 1\}^{K\beta}$ is the vector of information bits, and the B-BP algorithm is run on the equivalent bit-level system model with the resulting complexity being the same as that of NB-BP. In the simulations, the number of BP iterations used is 40 and the damping factor used is $\delta = 0.2$. From Fig. 3, we observe that the NB-BP detector performs considerably better than the MMSE and MF detectors. In large dimensions (e.g., $N = 256$), the NB-BP detector performance gets very close to SISO-AWGN performance. Also, the NB-BP scheme significantly outperforms the B-BP scheme (e.g., for $N = 256$, NB-BP performs better than B-BP by about 8 dB at $10^{-3}$ BER). This is because, with $M$-QAM, the assumption that the elements of $x_b$ in B-BP are independent is not true, and this results in a degraded performance in B-BP when applied to $M$-QAM.

---

```
Algorithm 1: The proposed NB-BP detection algorithm.

Input: $y, H, \sigma_e^2$

Initialize: $v_{ji}^{(0)}(s) \leftarrow 1/\sqrt{M}, \forall i, j, s$

for $m = 1 \rightarrow \text{Number of iterations}$

for $i = 1 \rightarrow 2N$

\[ \mu_i \leftarrow \sum_{l=1}^{2K} h_{il}s^T v_{li}^{(m-1)} \]

\[ \sigma_{ij}^2 \leftarrow \sum_{l=1}^{2K} h_{il}^2 \left( s \circ s \right)^T v_{li}^{(m-1)} - (s^T v_{li}^{(m-1)})^2 + \sigma_e^2 \]

\[ \forall s \in \mathbb{A} \]

\[ v_{ji}^{(m)}(s) \leftarrow \frac{1}{\sigma_i} \exp \left( \frac{-|y_i - a_{ij}(s)|^2}{2\sigma_{ij}^2} \right) \]

end

for $j = 1 \rightarrow 2K$

\[ \forall s \in \mathbb{A} \]

\[ v_{ji}^{(m)}(s) \leftarrow \frac{1}{Z} \prod_{l=1}^{2N} a_{ij}(s) \]

end

end

Output: $P_{x_j}(s) \leftarrow \frac{1}{Z} \prod_{l=1}^{2N} a_{ij}(s)$, $Z$ is normalizing constant.
```
Next, in Figs. 4(a) and 4(b), we show performance and complexity comparison of NB-BP with MMSE, ZF, and MF detectors for varying loading factors at 17 dB SNR, $N = 128$ and 16-QAM. As mentioned earlier, we can see that the NB-BP scheme achieves better performance than ZF and MMSE detectors at lesser complexity than these detectors across various loading factors, $\alpha$.

IV. OPTIMIZED LDPC CODE DESIGN FOR NB-BP DETECTOR-DECODER

In this section, we compute the EXIT chart of the NB-BP detector combined with the $q$-ary LDPC decoder, and obtain the optimal degree profile distribution of the $q$-ary LDPC code. The $q$-ary LDPC codewords are decoded by a message passing algorithm over a bipartite graph made of $n$ variable nodes and $k$ check nodes. A detailed description of non-binary LDPC decoding algorithm can be found in [8],[12]. As in [11], we formulate an integrated graph consisting of three sets of nodes, namely, variable nodes set, observation nodes set, and check nodes set. There are $nN$ observation nodes corresponding to the received vectors, $nK$ variable nodes corresponding to the transmitted coded symbol vectors, and $K(n-k)$ check nodes corresponding to the check equations of the LDPC code. Combining the NB-BP detector proposed in the previous section and the non-binary LDPC decoder, a joint message passing scheme is formulated for joint detection-decoding.

**EXIT analysis:** We use the EXIT chart analysis for analyzing the behavior of joint detector-decoder. If $I_A$ is the average mutual information between the coded symbols and input a priori information, and $I_E$ is the average mutual information between the coded symbols and the extrinsic output, then the EXIT function is $f(I_A) = I_E$. To obtain the EXIT characteristics of the joint detector-decoder, we first obtain the EXIT curves of the NB-BP detector and combine it with that of the LDPC decoder.

Let $I_{E,nbbp}$ and $I_{A,nbbp}$ denote the $I_E$ and $I_A$, respectively, for the NB-BP detector. Then the EXIT function is $I_{E,nbbp} = f(\gamma, K/N, I_{A,nbbp})$, where $\gamma$ is the average received SNR. Since an analytical evaluation of this function is difficult, we compute it through Monte Carlo simulations [10]. If $d_v$ and $d_c$ denote the variable node and check node degrees, respectively, of the LDPC code, then the EXIT function [12], [10] of the LDPC variable nodes set is given by $I_{E,V} = J \left( \sqrt{(d_v - 1)(J^{-1}(I_{A,V}))^2 + c_\gamma} \right)$, and the EXIT function of the LDPC check nodes set is given by $I_{E,C} = 1 - J \left( J^{-1}(1 - I_{A,C}) \sqrt{(d_c - 1)} \right)$, where $c$ is a constant dependent on $q$ and $J(.)$ is as defined in [10].

We formulate the EXIT function of the combination of the NB-BP detector and the variable nodes set of the $q$-ary LDPC decoder as

$$I_{E,jV}(\gamma, d_v, I_{A,jV}, I_{E,nbbp}) = J \left( \sqrt{(d_v - 1)(J^{-1}(I_{A,jV}))^2 + J^{-1}(I_{E,nbbp})^2} \right),$$

where $I_{E,jV}$ and $I_{A,jV}$ are the $I_E$ and $I_A$, respectively, for the combined variable nodes set. We match this EXIT curve with that of the check nodes set, such that the EXIT curve of the check nodes set lie below the EXIT curve of the combined variable nodes set.

**Optimized LDPC code design procedure:** To approach the capacity of the channel, the EXIT curves of the check nodes set and the variable nodes set should be matched. This matching is done by obtaining an appropriate degree distribution of the variable nodes and the check nodes, thereby designing irregular LDPC codes for a specific channel and receiver. The design methodology we adopt is described in detail in [13]. We use the method described in [9] to optimize the non-zero entries of the parity check matrix. By this method, the combination of the non-zero entries of a row of the parity check matrix that maximize the average entropy of the syndrome vector is chosen to be the entries of the row of our parity check matrix, $F$. We obtained optimized non-binary irregular LDPC codes using the design procedure described above, and the obtained codes for different system settings and loading factors are given in Table I.
Coded BER performance: We evaluated the coded BER performance of the proposed non-binary LDPC codes and compared with those of other LDPC codes, namely, (i) random non-binary ‘regular’ LDPC code, (ii) non-binary irregular LDPC code in [14], and (iii) optimized ‘binary’ irregular LDPC code in [13]. Figure 5 shows the simulated coded BER performance of the proposed rate-1/2 non-binary LDPC code with \( n = 1000 \) coded symbols using NB-BP detection and decoding in a system with \( N = 64 \), \( \alpha = 1 \), and 16-QAM. It can be seen that the proposed code significantly outperforms other codes; e.g., by about 1.2 to 3 dB at \( 10^{-5} \) coded BER. The better performance of the proposed code is because of the matching of EXIT charts of the combined NB-BP detector and non-binary LDPC decoder. Also, the proposed code’s performance is just about 2.3 dB away from capacity. Figure 6 shows the average SNR required to achieve a coded BER of \( 10^{-5} \) by the proposed rate-1/2 non-binary LDPC codes with NB-BP detection as a function of loading factor for \( N = 128 \), \( n = 1000 \) coded symbols, and 16-QAM. It can be seen that this performance is better than the performance achieved by the non-binary LDPC code in [14] with MMSE detection and NB-BP detection, by 1 dB to 7 dB for various system loading factors.

V. CONCLUSIONS

We proposed a promising non-binary BP algorithm for \( M \)-QAM signal detection in large-scale MIMO systems. An interesting feature of the proposed algorithm from an implementation point of view is that it achieves close to optimum performance in large-scale MIMO systems with less than MMSE complexity. It also enabled the design of good LDPC codes matched to large MIMO channels.

TABLE I. DEGREE PROFILES OF OPTIMIZED RATE-1/2 16-ARY LDPC CODES FOR DIFFERENT LARGE MULTIUSER MIMO CONFIGURATIONS. \( p_v \), \( p_c \): FRACTION OF VARIABLE NODES WITH DEGREE \( d_v \) AND CHECK NODES WITH DEGREE \( d_c \).  

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( N = 128, \alpha = 1 )</th>
<th>( N = 128, \alpha = 0.6 )</th>
<th>( N = 128, \alpha = 0.25 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (d_v, p_v) )</td>
<td>( (2, 0.4768), (6, 0.0104), (8, 0.3174), (10, 0.0024), (20, 0.0113) )</td>
<td>( (2, 0.6236), (8, 0.168),(16, 0.1833), (20, 0.0222) )</td>
<td>( (2, 0.3557), (3, 0.8015), (8, 0.0067), (12, 0.0358) )</td>
</tr>
<tr>
<td>( (d_c, p_c) )</td>
<td>( (6, 0.55206), (10, 0.1973), (18, 0.1517), (32, 0.1304) )</td>
<td>( (8, 0.5649), (16, 0.1755), (18, 0.2596) )</td>
<td>( (15, 0.7287), (8, 0.1793), (10, 0.0922) )</td>
</tr>
</tbody>
</table>

Fig. 5. Performance comparison of the proposed irregular non-binary LDPC codes with other LDPC codes for \( n = 1000 \) coded symbols, 16-QAM and \( N = K = 64 \), at a spectral efficiency of 128 bits/s/Hz.

Fig. 6. Performance comparison of the proposed irregular non-binary LDPC codes with other LDPC codes for \( n = 1000 \) coded symbols and 16-QAM. \( N = 128 \) for different loading factors.

REFERENCES


