Channel Hardening-Exploiting Message Passing (CHEMP) Receiver in Large MIMO Systems

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

***Abstract*—In this paper, we propose a multiple-input multiple-output (MIMO) receiver algorithm that exploits *channel hardening* that occurs in large MIMO channels. Channel hardening refers to the phenomenon where the off-diagonal terms of the $\mathbf{H}^H\mathbf{H}$ matrix become increasingly weaker compared to the diagonal terms as the size of the channel gain matrix H increases. Specifically, we propose a message passing detection (MPD) algorithm which works with the real-valued matched filtered received vector (whose signal term becomes $\mathbf{H}^T\mathbf{H}\mathbf{x}$, where x is the transmitted vector), and uses a Gaussian approximation on the off-diagonal terms of the $\mathbf{H}^T\mathbf{H}$ matrix. We also propose a simple estimation scheme which directly obtains an estimate of $\mathbf{H}^T\mathbf{H}$ (instead of an estimate of H), which is used as an effective channel estimate in the MPD algorithm. We refer to this receiver as the *channel hardening-exploiting message passing (CHEMP)* receiver. The proposed CHEMP receiver achieves very good performance in large-scale MIMO systems (e.g., in systems with 16 to 128 uplink users and 128 base station antennas). For the considered large MIMO settings, the complexity of the proposed MPD algorithm is almost the same as or less than that of the minimum mean square error (MMSE) detection. This is because the MPD algorithm does not need a matrix inversion. It also achieves a significantly better performance compared to MMSE and other message passing detection algorithms using MMSE estimate of H.**

Keywords *– Large-scale MIMO systems, channel hardening, message passing, detection, channel estimation.*

I. Introduction

Wireless communication systems using multiple-input multiple-output (MIMO) configurations with a large number of antennas have 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 the 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 channel estimation, 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].

Message passing on graphical models is a promising low-complexity high-performance approach for signal processing in large dimensions [4]. Decoding of turbo/LDPC codes, and equalization/detection are popular examples of the use of message passing algorithms in communications [5]. In [6], a MIMO detection algorithm based on approximate message passing on a factor graph is presented. The message passing 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.

In this this paper, we propose a promising low-complexity receiver for large-scale MIMO systems. The receiver is based on message passing. The novelty in the proposed receiver lies in the exploitation of the 'channel hardening' phenomenon that occurs in large MIMO channels [8]-[11]. Channel hardening refers to the phenomenon where the off-diagonal terms of the $\mathbf{H}^T\mathbf{H}$ matrix become increasingly weaker compared to the diagonal terms as the size of the channel gain matrix $\mathbf{H}$ increases. We exploit this for the purposes of detection and channel estimation. The proposed receiver, referred to as the *channel hardening-exploiting message passing (CHEMP)* receiver, consists of two components; a message passing detection (MPD) algorithm and an estimation scheme to obtain an estimate of $\mathbf{H}^T\mathbf{H}$. The highlights of our contributions in this paper can be summarized as follows:

- proposal of the MPD algorithm which works with the real-valued matched filtered received vector, and uses a Gaussian approximation on the off-diagonal terms of the $\mathbf{H}^T\mathbf{H}$ matrix.
- proposal of a simple estimation scheme which directly obtains an estimate of $\mathbf{H}^T\mathbf{H}$ (instead of an estimate of $\mathbf{H}$), which is used as an effective channel estimate in the MPD algorithm.
- less than the MMSE detection complexity (because matrix inversion is not needed in the MPD algorithm).
- significantly better performance compared to MMSE and other message passing detection algorithms which use an MMSE estimate of $\mathbf{H}$.

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. We consider $\alpha \leq 1$ (i.e., $K \leq N$). The system model is illustrated in Fig. 1. Let $\mathbf{H}_c \in \mathbb{C}^{N\times K}$ denote the channel gain matrix and H^c_{ij} denote the complex channel gain from the jth user to the ith BS antenna. The channel gains H^c_{ij}s are assumed

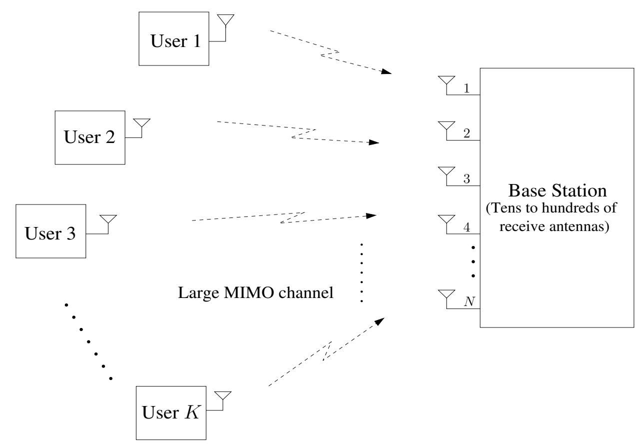

Fig. 1. Large-scale multiuser MIMO system model on the uplink.

to be independent Gaussian with zero mean and variance σ_j^2, such that $\sum_j \sigma_j^2 = K$. The σ_j^2 models the imbalance in the received power from user j due to path loss etc., and $\sigma_j^2 = 1$ corresponds to the case of perfect power control. Let $\mathbf{x}_c$ denote the modulated symbol vector, where the jth element of $\mathbf{x}_c$ denotes the modulation symbol (assume QPSK modulation) transmitted by the jth user. Assuming perfect synchronization, the received vector at the BS, $\mathbf{y}_c$, is given by

$$\mathbf{y}_c \quad = \quad \mathbf{H}_c\mathbf{x}_c + \mathbf{w}_c, \tag{1}$$

where $\mathbf{w}_c$ is the noise vector. Eqn. (1) can be written in the real domain as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2}$$

where

$$\mathbf{H} \triangleq \begin{bmatrix} \Re(\mathbf{H}_c) & -\Im(\mathbf{H}_c) \\ \Im(\mathbf{H}_c) & \Re(\mathbf{H}_c) \end{bmatrix},$$

$$\mathbf{y} \triangleq \begin{bmatrix} \Re(\mathbf{y}_c) \\ \Im(\mathbf{y}_c) \end{bmatrix}, \mathbf{x} \triangleq \begin{bmatrix} \Re(\mathbf{x}_c) \\ \Im(\mathbf{x}_c) \end{bmatrix}, \mathbf{w} \triangleq \begin{bmatrix} \Re(\mathbf{w}_c) \\ \Im(\mathbf{w}_c) \end{bmatrix},$$

$\Re(.)$ and $\Im(.)$ denote the real and imaginary parts, respectively. Note that $\mathbf{H} \in \mathbb{R}^{2N\times 2K}$, $\mathbf{y} \in \mathbb{R}^{2N}$, $\mathbf{x} \in \{\pm 1\}^{2K}$, and $\mathbf{w} \in \mathbb{R}^{2N}$. The elements of $\mathbf{w}$ are modeled as i.i.d. $\mathcal{N}(0, \sigma_n^2)$. For the real-valued system model in (2), the maximum-likelihood (ML) detection rule is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\{\pm 1\}^{2K}}{\text{argmin}} \ (\mathbf{y} - \mathbf{H}\mathbf{x})^T(\mathbf{y} - \mathbf{H}\mathbf{x}). \tag{3}$$

When the transmitted bits are equally likely, then the ML decision rule is same as the maximum a posteriori probability (MAP) decision rule, given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\{\pm 1\}^{2K}}{\text{argmax}} \ \Pr(\mathbf{x} \mid \mathbf{y}, \mathbf{H}). \tag{4}$$

The exact computation of (3) and (4) requires exponential complexity in K. Message passing algorithms can provide approximate marginalization of the joint distribution in (4) at low complexities. In Section III, we propose one such message passing algorithm, whose novelty lies in exploiting the channel hardening phenomenon that happens in large MIMO channels.

A. Channel hardening in large MIMO channels

Channel hardening refers to the phenomenon where the variance of the mutual information of the MIMO channel grows very slowly relative to its mean or even shrink as the number of antennas grows [8]. Consider a $n_r \times n_t$ MIMO channel. As n_r and n_t are increased keeping their ratio fixed, the distribution of the singular values of the MIMO channel matrix becomes less sensitive to the actual distribution of the entries of the channel matrix (as long as the entries are i.i.d.) [9]. This is a result of the Marčenko-Pastur law [10], which states that if the entries of a $n_r \times n_t$ matrix $\mathbf{H}$ are zero mean i.i.d. with variance $1/n_r$, then the empirical distribution of eigenvalues of $\mathbf{H}^H\mathbf{H}$ converges almost surely, as $n_r, n_t \to \infty$ with $n_t/n_r = \alpha$, to a density function [11]

$$p_\alpha(x) = \Big(1 - \frac{1}{\alpha}\Big)^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\alpha x}, \tag{5}$$

where $(x)^+ = \max(x, 0)$, $a = (1-\sqrt{\alpha})^2$, and $b = (1+\sqrt{\alpha})^2$. An effect of the Marčenko-Pastur law is that very tall or very wide matrices[1] are very well conditioned. The law also implies that the channel "hardens", i.e., the eigenvalue histogram of a single realization converges to the average asymptotic eigenvalue distribution.

Channel hardening can bring in several advantages in large dimensional signal processing. For example, linear detection in large systems will require inversion of large matrices. Inversion of large random matrices can be done fast using series expansion techniques [12],[13]. Because of channel hardening, approximate matrix inversions using series expansion and deterministic approximations from limiting distribution become effective in large dimensions.

An interesting aspect in channel hardening is that as the size of $\mathbf{H}$ increases, the off-diagonal terms of the $\mathbf{H}^H\mathbf{H}$ matrix become increasingly weaker compared to the diagonal terms, i.e., $\frac{\mathbf{H}^H\mathbf{H}}{n_r} \to \mathbf{I}_{n_t}$ for $n_r, n_t \to \infty$ with $n_t/n_r = \alpha$. This phenomenon is pictorially illustrated in Fig. 2, where we have plotted $\mathbf{H}^T\mathbf{H}$ for the real-valued channel model in (2) for 8×8, 32×32, 64×64, and 128×128 channels. In proposing the new receiver algorithm in the next section, we will work with approximations to the off-diagonal terms of the $\mathbf{H}^T\mathbf{H}$ matrix and estimates of $\mathbf{H}^T\mathbf{H}$, which are found to achieve very good performance in large dimensions at low complexities.

III. The Proposed CHEMP Receiver

In this section, we present the proposed CHEMP receiver. The proposed CHEMP receiver has two main components: 1) a message passing based detection (MPD) algorithm, and 2) a scheme to estimate $\mathbf{H}^T\mathbf{H}$. The proposed MPD algorithm works with the real-valued matched filtered received vector (whose signal term becomes $\mathbf{H}^T\mathbf{H}\mathbf{x}$), and uses a Gaussian approximation on the off-diagonal terms of the $\mathbf{H}^T\mathbf{H}$ matrix. Before we describe the proposed MPD algorithm, we state the

[1] In practice, the channel matrix in a multiuser system with tens of single-antenna users and hundreds of BS antennas will become a very tall matrix on the uplink, and a very wide matrix on the downlink.

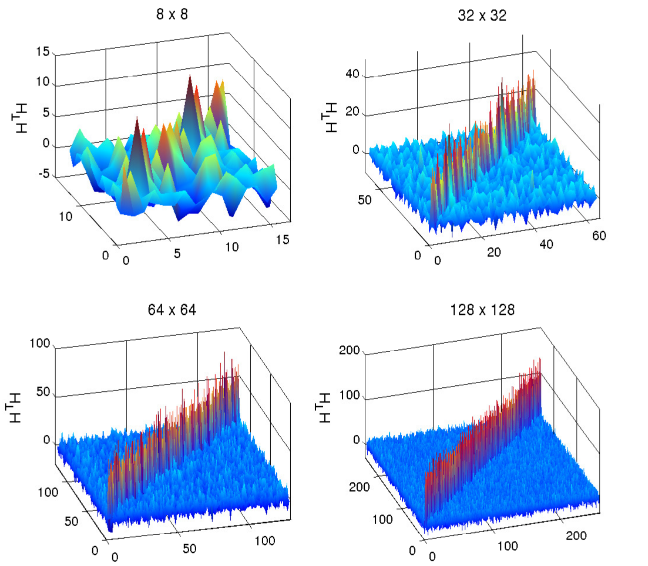

Fig. 2. Magnitude plots of $\mathbf{H}^T\mathbf{H}$ for 8×8, 32×32, 64×64, and 128×128 MIMO channels.

following lemma which will be used in the development and analysis of the detection algorithm.

Lemma 1. *Let X_i and Y_i be Gaussian random variables with zero mean and variance σ_x^2 and σ_y^2, respectively. Let $Z_i \triangleq X_iY_i$ and $Z \triangleq \frac{1}{n}\sum_{i=1}^{n} Z_i$.*

- *When X_i and Y_i are independent, $\mathbb{E}Z_i = 0$ and $\mathbb{E}Z_i^2 = \sigma_x^2\sigma_y^2$. Then by central limit theorem, for large n, $Z \sim \mathcal{N}(0, \frac{\sigma_x^2\sigma_y^2}{n})$. When X_i and Y_i are i.i.d., $Z \sim \mathcal{N}(0, \frac{\sigma_x^4}{n})$.*
- *When $X_i = Y_i$, Z is a χ^2 random variable of degree n. $\mathbb{E}Z = \sigma_x^2$ and Var$(Z) = \frac{2\sigma_x^4}{n}$.* □

A. Proposed MPD algorithm

Consider the real-valued system model in (2):

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{6}$$

Performing matched filter operation on $\mathbf{y}$, we have

$$\mathbf{H}^T\mathbf{y} = \mathbf{H}^T\mathbf{H}\mathbf{x} + \mathbf{H}^T\mathbf{w}. \tag{7}$$

From (7) we write the following:

$$\mathbf{z} = \mathbf{J}\mathbf{x} + \mathbf{v}, \tag{8}$$

where

$$\mathbf{z} \triangleq \frac{\mathbf{H}^T\mathbf{y}}{N}, \quad \mathbf{J} \triangleq \frac{\mathbf{H}^T\mathbf{H}}{N}, \quad \mathbf{v} \triangleq \frac{\mathbf{H}^T\mathbf{w}}{N}. \tag{9}$$

The ith element of $\mathbf{z}$ can be written as

$$z_i = J_{ii}x_i + \underbrace{\sum_{j=1, j\neq i}^{2K} J_{ij}x_j + v_i}_{\triangleq g_i}, \tag{10}$$

where J_{ij} is the element in the ith row and the jth column of $\mathbf{J}$, x_i is the ith element of $\mathbf{x}$, and $v_i = \sum_{j=1}^{2N} \frac{H_{ji}w_j}{N}$ is the ith element of $\mathbf{v}$. Note that the variable g_i defined in (10) denotes the interference-plus-noise term, which involves the off-diagonal elements of $\frac{\mathbf{H}^T\mathbf{H}}{N}$ (i.e., J_{ij}, $i \neq j$). We approximate the g_i term to have a Gaussian distribution with mean μ_i and variance σ_i^2, i.e., the distribution of g_i is approximated as $\mathcal{N}(\mu_i, \sigma_i^2)$. By central limit theorem, this approximation is accurate for large K, N. The mean and variance in this approximation are given by

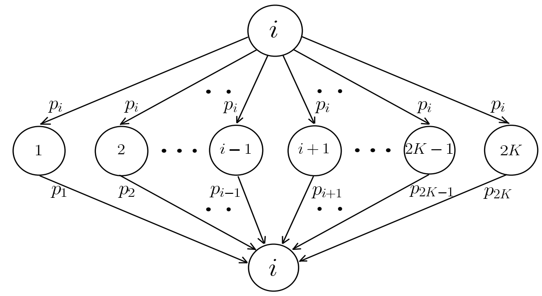

Fig. 3. Message passing in the proposed MPD algorithm.

$$\begin{aligned} \mu_i &= \mathbb{E}(g_i) = \sum_{j=1, j\neq i}^{2K} J_{ij}\mathbb{E}(x_j) \\ \sigma_i^2 &= \text{Var}(g_i) = \sum_{j=1, j\neq i}^{2K} J_{ij}^2\text{Var}(x_j) + \sigma_v^2. \end{aligned} \tag{11}$$

Denoting the probability of the symbol x_j as p_j, we have

$$\mathbb{E}(x_j) = (2p_j - 1), \quad \text{Var}(x_j) = 4p_j(1-p_j). \tag{12}$$

Also, note that by Lemma 1, $\sigma_v^2 = \frac{\sigma_n^2}{2N}$. Because of the above Gaussian approximation, the a posteriori probabilities of the transmitted symbols can be written as

$$\Pr(x_i|z_i) \propto \exp\Big(\frac{-1}{2\sigma_i^2}(z_i - J_{ii}x_i - \mu_i)^2\Big). \tag{13}$$

From (13), the log-likelihood ratio (LLR) of x_i, denoted by L_i, can be written as

$$L_i = \ln\frac{\Pr(z_i|x_i=+1)}{\Pr(z_i|x_i=-1)} = \frac{2J_{ii}}{\sigma_i^2}(z_i - \mu_i). \tag{14}$$

From (14), the probability of symbol x_i, can be written as

$$p_i = \frac{e^{L_i}}{1+e^{L_i}}. \tag{15}$$

Message passing: The system is modeled as a fully-connected graph, where the data symbols in $\mathbf{x}$ represent the nodes. There are $2K$ nodes in the graph. The message sent from the ith node to every other node is the probability p_i, computed from (15). Likewise, node i will receive similar messages from every other node; i.e., node i will receive message p_j from node j, $\forall j \neq i$. Figure 3 illustrates the above message passing schedule. Note that the computation of the messages p_i in (15) requires the computation of (11) and (14). The algorithm is initialized with $p_i = 0.5$, $\forall i$, and message passing is carried out for a certain number of iterations, after which the algorithm stops. The values of p_is at the end are taken as the soft values of x_is.

Algorithm 1 Proposed MPD algorithm

Require: $\mathbf{z}$, $\mathbf{J}$, σ_v^2, Δ

Initialize: $p_i^0 \leftarrow 0.5,\ i = 1, \cdots, 2K$
for $t = 1$ **to** *number_of_iterations* **do**
  for $i = 1$ **to** $2K$ **do**
    $\mu_i \leftarrow \sum_{j=1, j\neq i}^{2K} J_{ij}(2p_j^{t-1} - 1)$
    $\sigma_i^2 \leftarrow \sum_{j=1, j\neq i}^{2K} 4J_{ij}^2 p_j^{t-1}(1 - p_j^{t-1}) + \sigma_v^2$
    $L_i^t \leftarrow \frac{2J_{ii}}{\sigma_i^2}(z_i - \mu_i)$
    $\tilde{p}_i^t \leftarrow \frac{e^{L_i}}{1+e^{L_i}}$
  end for
  $\mathbf{p}^t \leftarrow (1 - \Delta)\tilde{\mathbf{p}}^t + \Delta \mathbf{p}^{t-1}$
end for

These soft values can be directly fed to the channel decoder in coded systems. In uncoded systems, a hard estimate of symbol x_i can be obtained as

$$\hat{x}_i = \begin{cases} +1 & \text{if } p_i \geq 0.5 \\ -1 & \text{otherwise.} \end{cases} \quad (16)$$

B. Improving convergence rate

At the end of the tth iteration of the detection algorithm described above, we obtain the probability of the ith user's information bit, p_i^t. The rate of convergence of this sequence $\{p_i^0, p_i^1, p_i^2, \cdots, p_i^t, \cdots\}$ can be improved by certain techniques including damping. Damping of messages passed in message passing algorithms is a scheme known to improve the rate of convergence of iterative algorithms. At the tth iteration, the message is damped by obtaining a convex combination of the message computed at the tth iteration and the message at the $(t-1)$th iteration, with a damping factor $\Delta \in [0, 1)$. Thus, if $\tilde{p}_i^t$ is the computed probability at the tth iteration, the message at the end of tth iteration is

$$p_i^t = (1 - \Delta)\tilde{p}_i^t + \Delta p_i^{t-1}. \quad (17)$$

A listing of the proposed MPD algorithm with damping is given in **Algorithm 1**.

C. Complexity comparison between MPD and MMSE

The computational complexity of the MPD algorithm is as follows. The complexity (in number of real operations) required to compute (11) and (15) is of order $O(K^2)$. The complexities of computing $\mathbf{z}$ and $\mathbf{J}$ are of orders $O(NK)$ and $O(NK^2)$, respectively. So, the total complexity of the proposed MPD is $O(NK^2)$, which is attractive for large-scale MIMO systems.

In Table I, we present an interesting comparison between the complexities of MPD and MMSE detection for $N = 128, 256$, and K varied from 16 to 256. Since we have used 20 iterations for MPD in all the BER simulations, we have taken the number of iterations to be 20 for the calculation of the MPD complexity. From Table I, the following interesting observations can be made: 1) for large N (e.g., $N = 256$), MPD complexity is less than MMSE complexity. This is because MPD needs only matrix multiplication and not matrix inversion, whereas MMSE detection needs both matrix multiplication and inversion. 2) for $N = 128$, the MPD complexity for $K = 64, 96, 128$ is less than the MMSE complexity. For $K = 16, 32$, the MPD complexity is almost the same as (marginally higher than) MMSE complexity, because the number of iterations ($= 20$) is comparable with K ($= 16, 32$). Also, MPD performs better than MMSE detection, and achieves close to optimal detection performance for large K, N, and different system loading factors. We will see this performance advantage of MPD next.

K	Complexity in no. of real operations $\times 10^6$			
	$N = 128$		$N = 256$	
	MMSE	MPD (prop.)	MMSE	MPD (prop.)
16	0.1775	0.1798	0.3331	0.2964
32	0.7482	0.7496	1.3216	1.1908
64	3.5936	3.2001	5.7890	4.7738
96	9.5846	7.2088	14.4507	10.7489
128	19.7700	12.8146	28.3552	19.1160
256	-	-	157.3737	76.5051

TABLE I
COMPARISON BETWEEN THE COMPLEXITIES (IN NUMBER OF REAL OPERATIONS) OF THE PROPOSED MPD AND THE MMSE DETECTION FOR DIFFERENT VALUES OF K, N. NUMBER OF ITERATIONS FOR MPD = 20.

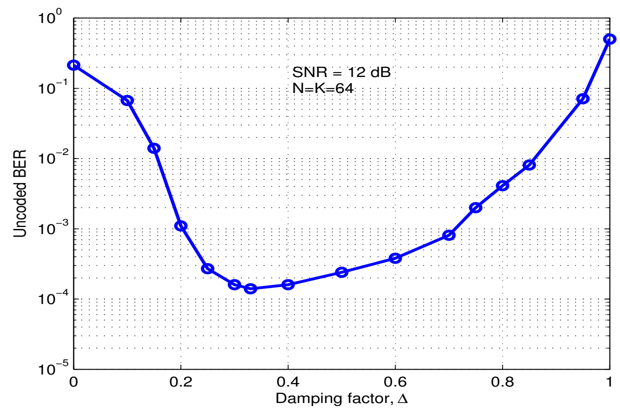

Fig. 4. Uncoded BER performance of the proposed MPD algorithm as a function of damping factor Δ. $N = K = 64$, SNR=12 dB.

D. BER performance of MPD

In this subsection, we present the uncoded BER performance of MPD obtained through simulations for different system parameter settings. We will now assume perfect knowledge $\mathbf{H}$. We will relax this assumption later. First, in Fig. 4, we plot the uncoded BER of MPD at an average SNR of 12 dB for $N = K = 64$ for various values of the damping factor Δ. The number of message passing iterations used is 20. From this figure, we observe that a damping factor of $\Delta = 0.33$ is optimal. This value of Δ is found to give good performance for other values of system parameters as well. So we have used this value of Δ in all the simulations.

In Fig. 5, we plot the uncoded BER of MPD and MMSE detector for different values of N ($= 4, 8, 16, 32, 64, 128$) for a system loading factor of $\alpha = 1$. Since optimal detection performance for such large-dimension systems is hard

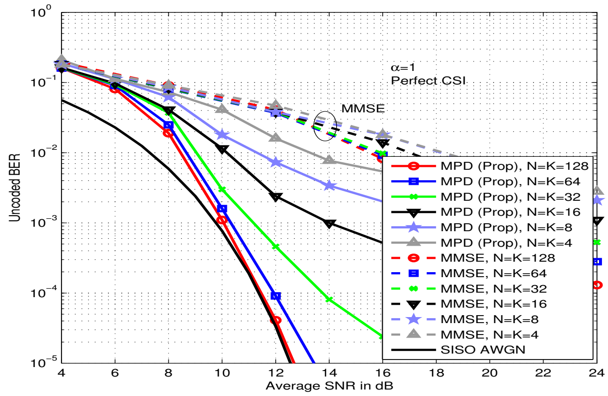

Fig. 5. Uncoded BER performance of the MPD algorithm and the MMSE detector for $N = K = 4, 8, 16, 32, 64, 128$.

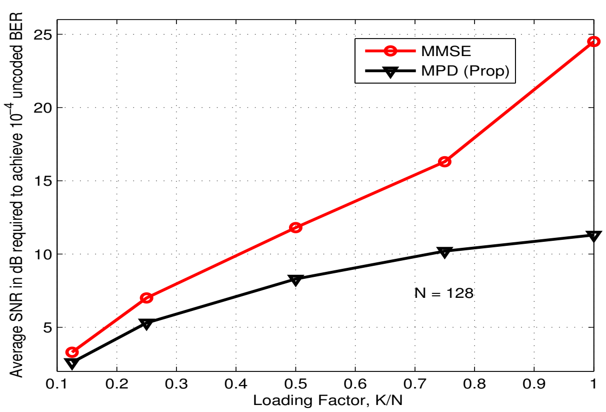

Fig. 6. Comparison between the average SNR required to achieve an uncoded BER of 10^{-4} in MMSE detection and MPD at different loading factors with $N = 128$.

to obtain, we have plotted single-input single-output (SISO) additive white Gaussian noise (AWGN) channel performance as a lower bound on the optimum detection performance. From Fig. 5, it is observed that the performance of MPD improves for increasing N, K, and moves closer to the SISO-AWGN performance for large N, K; for e.g., the MPD performance for $N = K = 64, 128$ gets very close to SISO-AWGN performance. It can be also seen that MPD is better than MMSE detector.

Interestingly, the MPD performance for different loading factors is better than MMSE detection performance. This is shown in Fig. 6, where the average SNRs required to achieve an uncoded BER of 10^{-4} in MPD and MMSE detection are plotted. It can be observed from Fig. 6 that the MPD outperforms the MMSE detection by about 1.2 dB at a loading factor of $\alpha = 0.125$. This performance advantage of MPD over MMSE detection increases for increasing values of α. For example, the performance advantage of MPD over MMSE detection is 6.5 dB and 12.5 dB for $\alpha = 0.75$ and $\alpha = 1$, respectively. This performance advantage of MPD becomes very attractive given that MPD complexity is almost same or less than the MMSE detection complexity (as discussed in Section III-C).

E. Channel estimation for MPD

A key issue in large-scale MIMO systems is the estimation of channel gains. In conventional approaches, the NK channel gains in the channel matrix are estimated and used for the detection of transmitted symbols. Note that in our transformed system model (8), the influence of the channel on vector $\mathbf{z}$ is through $\mathbf{H}^T\mathbf{H}$, rather than through $\mathbf{H}$ as such. We propose to exploit this observation on the structure of the system model (8). Specifically, we propose to directly obtain an estimate of $\mathbf{H}^T\mathbf{H}$ and use it in the MPD algorithm, rather than obtaining an estimate of $\mathbf{H}$ as done in conventional approaches. We note that this approach is simple and novel, and it works very well in the MPD algorithm (as we will see in the performance results). We present the scheme to obtain an estimate of the $\mathbf{H}^T\mathbf{H}$ matrix next.

Estimating the $\mathbf{H}^T\mathbf{H}$ matrix:

Note that we have defined $\mathbf{J} = \mathbf{H}^T\mathbf{H}$. We are interested in obtaining $\hat{\mathbf{J}}$, an estimate of $\mathbf{J}$. We assume that the channel is slowly fading, where the channel matrix $\mathbf{H}$ remains constant over one frame duration (which is taken to be equal to the coherence time of the channel). The length of one frame is L_f channel uses. Each frame consists of a pilot part and a data part. The pilot part consists of K channel uses, and the data part consists of $L_f - K$ channel uses.

Let $\mathbf{X}_p = P\mathbf{I}_K$ denote the pilot matrix, where in the ith channel use, $1 \leq i \leq K$, user i transmits a pilot tone with amplitude P and the other users remain silent. The received pilot matrix at the BS is then given by

$$\mathbf{Y}_p \;=\; \mathbf{H}\mathbf{X}_p + \mathbf{W}_p \;=\; P\mathbf{H} + \mathbf{W}_p, \tag{18}$$

where $P = \sqrt{KE_s}$, E_s is the average symbol energy, and $\mathbf{W}_p$ is the noise matrix. Using Lemma 1, we obtain an estimate of the matrix $\mathbf{J}$ as

$$\hat{\mathbf{J}} \;=\; \frac{\mathbf{Y}_p^T\mathbf{Y}_p}{NP^2} - \frac{\sigma_v^2}{P^2}\mathbf{I}_K. \tag{19}$$

An estimate of the vector $\mathbf{z}$ is obtained as

$$\hat{\mathbf{z}} \;=\; \frac{\mathbf{Y}_p^T\mathbf{y}}{NP}. \tag{20}$$

The estimates $\hat{\mathbf{J}}$ and $\hat{\mathbf{z}}$ are used as inputs to the MPD algorithm in place of $\mathbf{J}$ and $\mathbf{z}$.

Note on complexity: A key advantage of the above estimation scheme is its low complexity. The computation of $\hat{\mathbf{J}}$ and $\hat{\mathbf{z}}$ in (19) and (20) requires only matrix and vector multiplications. Note that even when perfect knowledge of $\mathbf{H}$ or an estimate of $\mathbf{H}$ is available, similar computations are needed to compute $\mathbf{J}$ and $\mathbf{z}$. Further note that the additional complexity needed to obtain an estimate of $\mathbf{H}$ in the conventional approach is avoided in our approach.

F. BER performance of the CHEMP receiver

As stated before, we refer to the combination of proposed MPD algorithm and the channel estimation scheme proposed

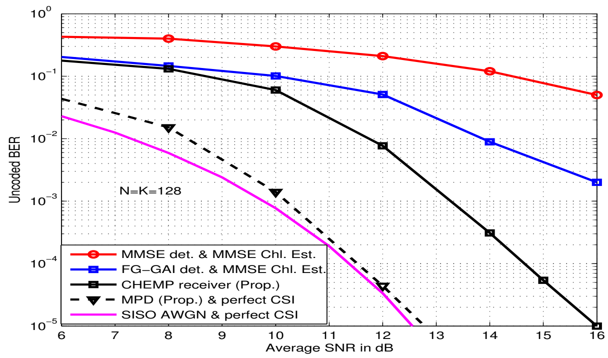

Fig. 7. Comparison of the BER performance of the proposed CHEMP receiver with those of 1) MMSE detector with MMSE channel estimate, and 2) FG-GAI detector in [6] with MMSE channel estimate, for $N = K = 128$.

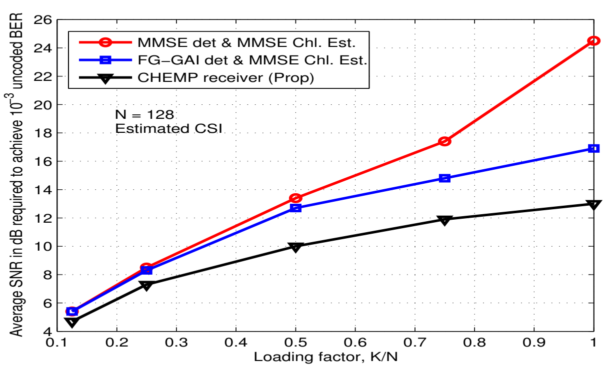

Fig. 8. Comparison between the average SNR required to achieve an uncoded BER of 10^{-3} in 1) proposed CHEMP receiver, 2) MMSE detector with MMSE channel estimate, and 3) FG-GAI detector in [6] with MMSE channel estimate, at different loading factors with $N = 128$.

in the previous subsection as the CHEMP receiver. Here, we present the uncoded BER performance of the CHEMP receiver. The number of iterations used in the MPD algorithm is 20. We compare the performance of the CHEMP receiver with two other receivers, namely, 1) MMSE detector with MMSE channel estimate, and 2) FG-GAI (factor graph with Gaussian approximation of interference) detector in [6] with MMSE channel estimate. We note that the FG-GAI detector in [6] is also a message passing algorithm which used a Gaussian approximation of interference. But this approximation was done on the original system model in (2), whereas in the proposed MPD algorithm, the Gaussian approximation is done on the matched filtered system model in (8) and the proposed channel estimation scheme is not applicable for the FG-GAI detection algorithm in [6].

In Fig. 7, we present an uncoded BER performance comparison between 1) proposed CHEMP receiver, 2) MMSE detector with MMSE channel estimate, and 3) FG-GAI detector in [6] with MMSE channel estimate. It can be seen that the performance of the proposed CHEMP receiver is significantly better than those of the MMSE and FG-GAI detectors with MMSE estimate of the channel. This shows that the proposed approach in CHEMP receiver is simple and effective in terms of both complexity and performance.

In Fig. 8 we illustrate a comparison between the the average SNR required to achieve an uncoded BER of 10^{-3} in 1) proposed CHEMP receiver, 2) MMSE detector with MMSE channel estimate, and 3) FG-GAI detector in [6] with MMSE channel estimate, at different loading factors with $N = 128$. From this figure, we observe that the CHEMP receiver outperforms the other two receivers. For example, the CHEMP receiver outperforms the MMSE detector with MMSE channel estimate by about 0.6 dB to 6 dB for loading factors in the range of 0.125 to 0.75. Likewise, the performance advantage of the CHEMP receiver over FG-GAI detector with MMSE channel estimate is about 0.6 dB to 4 dB for loading factors in the range of 0.125 to 1.

IV. Conclusions

We proposed a promising message passing based receiver (referred to as the 'CHEMP receiver') for low complexity detection and channel estimation in large-scale MIMO systems. Further, 1) an analysis of the convergence and convergence rate of the CHEMP receiver, 2) an analytical reasoning as to why the CHEMP receiver performs better with the proposed channel estimation scheme, and 3) LDPC code design matched for the large MIMO channel and the CHEMP receiver, are available in [14].

References

[1] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.,* vol. 26, no. 3, pp. 473-485, Apr. 2008.

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?" *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 2, pp. 160-171, Feb. 2013.

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

[5] R. J. McEliece, D. J. C. MacKay, and J-F. Cheng, "Turbo decoding as an instance of Pearls "belief propagation" algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140-152, Feb. 1998.

[6] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical Models," *IEEE J. Sel. Topics Signal Proc.*, vol. 5, no. 8, pp. 1497-1511, Dec. 2011

[7] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *IEEE Trans. Inform. Theory*, vol. 57. no. 8, pp. 4973-4982, Aug. 2011.

[8] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, pp. 1893-1909, Sep. 2004.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[10] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR Sbornik*, pp. 457-483, 1967.

[11] A. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*, Foundations and Trends in Communications and Information Theory, Now Publishers, Inc., 2004.

[12] S. Moshavi, E. Kanterakis, and D. Schilling, "Multistage linear receivers for DS-CDMA systems," *Intl. Jl. of Wireless Inform. Netw.*, vol. 3, no. 1, pp. 1-17, Jan. 1996.

[13] M. Wu, B. Yin, A. Vosoughi, C. Studer, J. R. Cavallaro, C. Dick, "Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink," *Proc. IEEE ISCAS'2013,* May 2013.

[14] T. Lakshmi Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," arXiv:1310.3062 [cs.IT] 11 Oct 2013.