A Low-Complexity Gaussian Approximate Message Passing Detection Algorithm For Massive MIMO With High Order Modulation

He Zhu, Beilei Gong, Zheng Wang
School of Information Science and Engineering
Southeast University, Nanjing, China
Email: he_zhu@seu.edu.cn, beilei_gong@seu.edu.cn, wznuaa@gmail.com

***Abstract*—In this paper, the low-complexity Gaussian approximate message passing (LC-GAMP) detection algorithm is proposed for massive MIMO with high order modulation. By using the partial set instead of the full set of the constellation symbols, the iterations in GAMP can be significantly reduced with negligible performance loss. Meanwhile, the proposed LC-GAMP algorithm is further improved via deep learning (DL). After tuning the update step size by introducing the trainable hyper-parameters, considerable detection performance improvement can be achieved. The simulation results show that the proposed LC-GAMP algorithm not only achieves a better trade-off between performance and computational complexity than the original GAMP algorithm, but also has a competitive performance compared to other MP based detection schemes.**

***Index Terms*—Low complexity, massive MIMO detection, message passing (MP), deep learning (DL).**

I. Introduction

Massive MIMO has become a crucial technology for the beyond fifth-generation (B5G) and sixth-generation (6G) wireless communication systems due to its high spectral efficiency, excellent energy efficiency and strong link reliability [1]. However, the dramatic increase of the system dimension also imposes an unbearable pressure on the uplink signal detection of massive MIMO [2]. To this end, as a promising technique in signal processing, in recent years the message passing (MP) based on factor graphs (FGs) has gained much attention in the research of massive MIMO detection [3].

Specifically, the authors of [4] proposed the channel hardening-exploiting message passing (CHEMP) receiver, but it costs a high complexity due to the computation of the Gram matrix. On the other hand, the Gaussian approximation based message passing (AMP-G) detection was proposed in [5], which avoids the computation of the Gram matrix. Scaled-and-added Gaussian message passing (SA-GMP) was presented in order to speed up the convergence of GMP algorithm [6]. In [7], Gaussian approximate message passing (GAMP) algorithm was given, which is attractive due to its improved detection performance. However, the complexity of GAMP still increases rapidly with the number of users and the modulation order. In addition, the DL aided detection scheme was proposed in [9] and obtained an improvement in detection performance on approximate message passing (AMP) algorithm.

In this paper, with respect to GAMP for massive MIMO detection, the LC-GAMP detection algorithm is proposed for the further complexity reduction, which considers only a partial set of constellation symbols during each iteration. Then, the LC-GAMP algorithm is improved in terms of the detection performance by tuning the update step size in LC-GAMP via the trainable hyper-parameters. Simulation results show that the proposed LC-GAMP algorithm achieves significantly complexity reduction with better detection performance compared to GAMP and the other MP based detection schemes.

II. GAMP Algorithm for Massive MIMO Detection

Consider an uplink massive MIMO system with N receive antennas at the base station, serving K single-antenna users simultaneously at the transmitter with $N \gg K$. Then the relationship between the receive and transmit signals is

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

Here, $\tilde{\mathbf{y}} \in \mathbb{C}^N$ denotes the received signal vector, $\tilde{\mathbf{H}} \in \mathbb{C}^{N\times K}$ is the channel gain matrix with each element following independent Gaussian distribution with zero mean and unit variance, $\tilde{\mathbf{x}}$ is the K-dimensional transmit signal vector, and each element is a M-quadrature amplitude modulation (M-QAM) symbol, $\tilde{\mathbf{n}} \in \mathbb{C}^N$ denotes additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2. Here, for notational simplicity, we convert the complex system in (1) into an equivalent but real one, namely

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

with $\mathbf{H} = [\mathcal{R}\{\tilde{\mathbf{H}}\}\ -\mathcal{I}\{\tilde{\mathbf{H}}\};\ \mathcal{I}\{\tilde{\mathbf{H}}\}\ \mathcal{R}\{\tilde{\mathbf{H}}\}] \in \mathbb{R}^{2N\times 2K}$, $\mathbf{y} = [\mathcal{R}\{\tilde{\mathbf{y}}\};\ \mathcal{I}\{\tilde{\mathbf{y}}\}] \in \mathbb{R}^{2N}$, $\mathbf{x} = [\mathcal{R}\{\tilde{\mathbf{x}}\};\ \mathcal{I}\{\tilde{\mathbf{x}}\}] \in \mathbb{S}^{2K}$, and $\mathbf{n} = [\mathcal{R}\{\tilde{\mathbf{n}}\};\ \mathcal{I}\{\tilde{\mathbf{n}}\}] \in \mathbb{R}^{2N}$, where $\mathbb{S} = \{-\sqrt{M}+1, ..., -1, 1, ..., \sqrt{M}-1\}$ denotes the M-QAM.

Theoretically, the GAMP algorithm is based on a pairwise FG consisting of the variable nodes (VNs) and check nodes (CNs) [7]. Specifically, in GAMP algorithm, given the initial setup $m^0_{f_n\to x_k} = 0$ and $v^0_{f_n\to x_k} \to +\infty$, the downward

Corresponding author: Zheng Wang (e-mail: wznuaa@gmail.com)

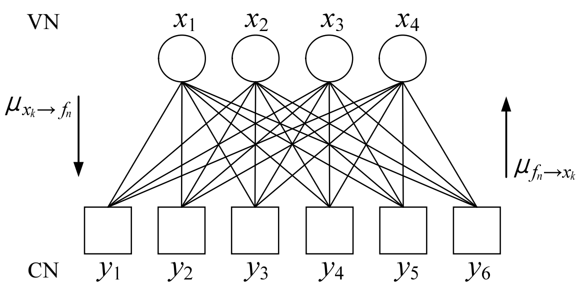

Fig. 1. An Illustration of the FG with $N \times K = 6 \times 4$.

oriented message $\mu_{x_k \to f_n}$ transmitted from VN x_k to CN f_n in the i-th $(i = 1, 2, ..., t)$ iteration can be formulated as

$$\mu^i_{x_k \to f_n}(x_k) \propto \frac{\mathcal{CN}\left(x_k; z^{i-1}_{x_k \to f_n}, \gamma^{i-1}_{x_k \to f_n}\right)}{\sum\limits_{x_k \in \mathbb{S}} \mathcal{CN}\left(x_k; z^{i-1}_{x_k \to f_n}, \gamma^{i-1}_{x_k \to f_n}\right)} \tag{3}$$

with

$$\gamma^{i-1}_{x_k \to f_n} = \left(\sum_{n' \neq n} \frac{h^2_{n',k}}{v^{i-1}_{f_{n'} \to x_k}} \right)^{-1}, \tag{4}$$

$$z^{i-1}_{x_k \to f_n} = \gamma^{i-1}_{x_k \to f_n} \sum_{n' \neq n} \frac{h_{n',k} m^{i-1}_{f_{n'} \to x_k}}{v^{i-1}_{f_{n'} \to x_k}}. \tag{5}$$

Next, the message $\mu^i_{x_k \to f_n}(x_k)$ in (3) is approximated by a Gaussian distribution by minimizing the KL divergence $D_{KL}(\mu^i_{x_k \to f_n}(x_k) || \mathcal{CN}(x_k; m^i_{x_k \to f_n}, v^i_{x_k \to f_n}))$ for arbitrary prior probability [8]. This gives the expressions

$$m^i_{x_k \to f_n} = \sum_{s \in \mathbb{S}} s \mu^i_{x_k \to f_n}(x_k = s), \tag{6}$$

$$v^i_{x_k \to f_n} = \sum_{s \in \mathbb{S}} |s|^2 \mu^i_{x_k \to f_n}(x_k = s) - |m^i_{x_k \to f_n}|^2. \tag{7}$$

Meanwhile, the mean and variance of the upward oriented message $\mu^i_{f_n \to x_k}(x_k)$ transmitted from CN f_n to VN x_k at the i-th iteration are expressed as

$$m^i_{f_n \to x_k} = y_n - \sum_{k' \neq k} h_{n,k'} m^i_{x_{k'} \to f_n}, \tag{8}$$

$$v^i_{f_n \to x_k} = \sigma^2_n + \sum_{k' \neq k} h^2_{n,k'} v^i_{x_{k'} \to f_n}. \tag{9}$$

To sum up, in each iteration, the messages in both directions are exchanged between CNs and VNs, and are updated alternately until the number of iterations t is reached. Subsequently, the a *posteriori* probability $\mu_{x_k}(x_k)$ is expressed as

$$\mu_{x_k}(x_k) \propto \frac{\mathcal{CN}(x_k; z_{x_k}, \gamma_{x_k})}{\sum\limits_{x_k \in \mathbb{S}} \mathcal{CN}(x_k; z_{x_k}, \gamma_{x_k})} \tag{10}$$

with

$$\gamma_{x_k} = \left(\sum_n \frac{h^2_{n,k}}{v^t_{f_n \to x_k}} \right)^{-1} \text{ and } z_{x_k} = \gamma_{x_k} \sum_n \frac{h_{n,k} m^t_{f_n \to x_k}}{v^t_{f_n \to x_k}}. \tag{11}$$

Algorithm 1: GAMP Detection for Massive MIMO

Input: $\mathbf{y} \in \mathbb{R}^{2N}$, $\mathbf{H} \in \mathbb{R}^{2N \times 2K}$
Output: $\hat{\mathbf{x}} = [\hat{x}_1, ..., \hat{x}_K]^T \in \mathbb{S}^{2K}$
Initialize: $m^0_{f_n \to x_k} = 0$ and $v^0_{f_n \to x_k} \to +\infty$
for $i = 1, ..., t$ **do**
  compute $\gamma^{i-1}_{x_k \to f_n}$ and $z^{i-1}_{x_k \to f_n}$ by (4)-(5)
  compute $\mu^i_{x_k \to f_n}$, $m^i_{x_k \to f_n}$, $v^i_{x_k \to f_n}$ by (3), (6), (7)
  compute $m^i_{f_n \to x_k}$ and $v^i_{f_n \to x_k}$ by (8)-(9)
end
compute γ_{x_k}, z_{x_k} and $\mu_{x_k}(x_k)$ by (10)-(12)
compute $\hat{x}_k$ by (12)

Finally, based on $\mu_{x_k}(x_k)$ in (10), the target transmitted signal $\mathbf{x}$ can be estimated by

$$\hat{x}_k = \arg \min_{x_k \in \mathbb{S}} || \sum_{s \in \mathbb{S}} s \mu_{x_k}(s) - x_k ||. \tag{12}$$

III. The Proposed LC-GAMP Detection Algorithm

A. Algorithm Description

Typically, under M-QAM, $m^i_{x_k \to f_n}$ and $v^i_{x_k \to f_n}$ in (6) and (7) can be rewritten as

$$m^i_{x_k \to f_n} = s_1 \mu^i_{x_k \to f_n}(s_1) + ... + s_{\sqrt{M}} \mu^i_{x_k \to f_n}(s_{\sqrt{M}}), \tag{13}$$

$$v^i_{x_k \to f_n} = |s_1|^2 \mu^i_{x_k \to f_n}(s_1) + ... + |s_{\sqrt{M}}|^2 \mu^i_{x_k \to f_n}(s_{\sqrt{M}}) - |m^i_{x_k \to f_n}|^2. \tag{14}$$

Clearly, it can be found that the computations in (13) and (14) involve many multiplication and addition operations, which is rather sensitive to the increment of the modulation order. Meanwhile, we also noticed that most of the terms in (13) and (14) are quite small during the iterations. This is due to the fact that many constellation symbols in $\mathbb{S}$ correspond to very small values $\mu^i_{x_k \to f_n}(x_k)$.

To make this point more specific, we can rewrite the downward oriented message $\mu_{x_k \to f_n}$ in (3) as

$$\mu^i_{x_k \to f_n}(x_k = s_j) \propto \frac{\exp\left(\frac{-d^{i\,2}_{x_k \to f_n}(s_j)}{2\gamma^{i-1}_{x_k \to f_n}} \right)}{\sum\limits_{s_j \in \mathbb{S}} \exp\left(\frac{-d^{i\,2}_{x_k \to f_n}(s_j)}{2\gamma^{i-1}_{x_k \to f_n}} \right)} \tag{15}$$

with

$$d^i_{x_k \to f_n}(s_j) = |s_j - z^{i-1}_{x_k \to f_n}|. \tag{16}$$

Obviously, the size of $\mu^i_{x_k \to f_n}(x_k = s_j)$ heavily depends on the distance between s_j and $z^{i-1}_{x_k \to f_n}$, i.e., $d^i_{x_k \to f_n}(s_j)$. A larger value $d^i_{x_k \to f_n}(s_j)$ corresponds to a smaller $\mu^i_{x_k \to f_n}(x_k = s_j)$. For a better understanding, in Fig. 2, we give the result of tracking one set of $\mu^i_{x_k \to f_n}(x_k)$ by simulations. Specifically, the following observations can be found:

$$\mu^i_{x_k \to f_n}(x_k = s_1) \gg \mu^i_{x_k \to f_n}(x_k = s_2) > ..., \tag{17}$$

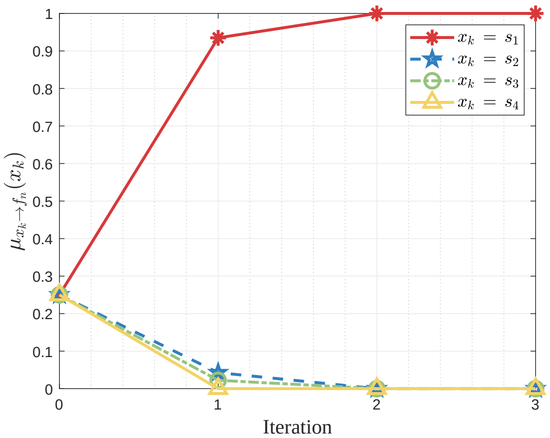

Fig. 2. Illustration of $\mu^i_{x_k\to f_n}(x_k)$ for $x_k = s_1, ..., s_4$ respectively in a 128×16 massive MIMO system using 16-QAM.

where this trend becomes more obvious as the iteration proceeds. Given it, by ignoring these trivial symbols directly during the iterations, significant complexity reduction can be achieved but only with negligible performance loss. Therefore, assuming that the approximation set of symbols we employ is $\mathbb{A}$, then (15) can be approximated by

$$\hat{\mu}^i_{x_k\to f_n}(x_k = s_j) \propto \frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\sum\limits_{s_j\in\mathbb{A}}\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}. \tag{18}$$

Similarly, we can also ignore the small terms related to these symbols in (13) and (14), so as to

$$\hat{m}^i_{x_k\to f_n} = s_1\hat{\mu}^i_{x_k\to f_n}(s_1) + .. + s_{|\mathbb{A}|}\hat{\mu}^i_{x_k\to f_n}(s_{|\mathbb{A}|}), \tag{19}$$

$$\hat{v}^i_{x_k\to f_n} = |s_1|^2\hat{\mu}^i_{x_k\to f_n}(s_1) + ... + |s_{|\mathbb{A}|}|^2\hat{\mu}^i_{x_k\to f_n}(s_{|\mathbb{A}|}) - |\hat{m}^i_{x_k\to f_n}|^2. \tag{20}$$

To summarize, we outline the proposed LC-GAMP detection algorithm in Algorithm 2.

B. Approximate Set Size Selection Criteria

Obviously, according to (15) and (18), the size of $\mathbb{A}$ has a great impact on the approximation accuracy and implementation efficiency. Hence, how to choose the appropriate size of $\mathbb{A}$, i.e., $|\mathbb{A}|$, is a worthy issue to consider in the proposed LC-GAMP detection.

To clearly depict the approximation accuracy, based on the right-hand sides (RHS) of (15) and (18), we define the approximation error as

$$e = \sum_{s_j\in\mathbb{S}}\frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} - \sum_{s_j\in\mathbb{A}}\frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} = \sum_{s_j\notin\mathbb{A}, s_j\in\mathbb{S}}\frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} \tag{21}$$

Algorithm 2: LC-GAMP Detection for Massive MIMO

Input: $\mathbf{y}\in\mathbb{R}^{2N}$, $\mathbf{H}\in\mathbb{R}^{2N\times 2K}$
Output: $\hat{\mathbf{x}} = [\hat{x}_1, ..., \hat{x}_K]^T \in \mathbb{S}^{2K}$
Initialize: $m^0_{f_n\to x_k} = 0$ and $v^0_{f_n\to x_k} \to +\infty$
for $i = 1, ..., t$ **do**
compute $\gamma^{i-1}_{x_k\to f_n}$ and $z^{i-1}_{x_k\to f_n}$ by (4)-(5)
compute $d^i_{x_k\to f_n}$ by (16)
select the symbols to form $\mathbb{A}$ by $d^i_{x_k\to f_n}$
compute $\hat{\mu}^i_{x_k\to f_n}$, $\hat{m}^i_{x_k\to f_n}$, $\hat{v}^i_{x_k\to f_n}$ by (18)-(21)
compute $m^i_{f_n\to x_k}$ and $v^i_{f_n\to x_k}$ by (8)-(9)
end
compute γ_{x_k}, z_{x_k} and μ_{x_k} by (10)-(12)
compute $\hat{x}_k$ by (12)

with $\rho^i_{n,k}(\mathbb{S}) = \sum\limits_{s_j\in\mathbb{S}}\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)$. Then, by letting $|\mathbb{A}| = 2N$, we can arrive at the following results.

Lemma 1: The approximation error e in (21) decays exponentially with the increment of N as

$$e < \alpha^{-4N^2} \tag{22}$$

with $\alpha = \exp(1/(2\gamma^{i-1}_{x_k\to f_n}))$.

Proof. Assuming that s_m corresponds to the minimum value of $d^i_{x_k\to f_n}(s_j)$ for $1\le j\le\sqrt{M}$. Then, by letting $d = |s_m - z^{i-1}_{x_k\to f_n}|$, we have

$$d^i_{x_k\to f_n}(s_j) = \begin{cases} j-1+d, & \text{if } j \text{ is even} \\ j-d, & \text{if } j \text{ is odd}\end{cases}. \tag{23}$$

Then, (21) can be rewritten as

$$\begin{aligned} e &= \sum_{j=N+1}^{\sqrt{M}/2}\frac{\alpha^{-(2(j-1)+d)^2} + \alpha^{-(2j-d)^2}}{\rho^i_{n,k}(\mathbb{S})} \\ &< \sum_{j=N+1}^{\sqrt{M}/2}\frac{\alpha^{-(2j-2)^2}(\alpha^{-d^2} + \alpha^{-(2-d)^2})}{\rho^i_{n,k}(\mathbb{S})} \\ &< (1+\alpha^{-4}+\alpha^{-16}+...)\cdot\frac{\alpha^{-(2N)^2}(\alpha^{-d^2}+\alpha^{-(2-d)^2})}{\rho^i_{n,k}(\mathbb{S})} \\ &\approx \frac{\alpha^{-(2N)^2}(\alpha^{-d^2} + \alpha^{-(2-d)^2})}{\rho^i_{n,k}(\mathbb{S})}. \end{aligned} \tag{24}$$

Finally, due to $\rho^i_{n,k}(\mathbb{S}) > (\alpha^{-(2(j-1)+d)^2} + \alpha^{-(2j-d)^2})|_{j=1} = \alpha^{-d^2} + \alpha^{-(2-d)^2}$, we can easily arrive at the result in (22), which completes the proof. □

From (22), we can conclude that the approximation error e decays rapidly in an exponential way, which is negligible even very few symbols are considered in $\mathbb{A}$. Therefore, we suggest to use $N = 1$, i.e., $|\mathbb{A}| = 2$ for the practical signal detection, which means only the first two constellation symbols are employed during each iteration.

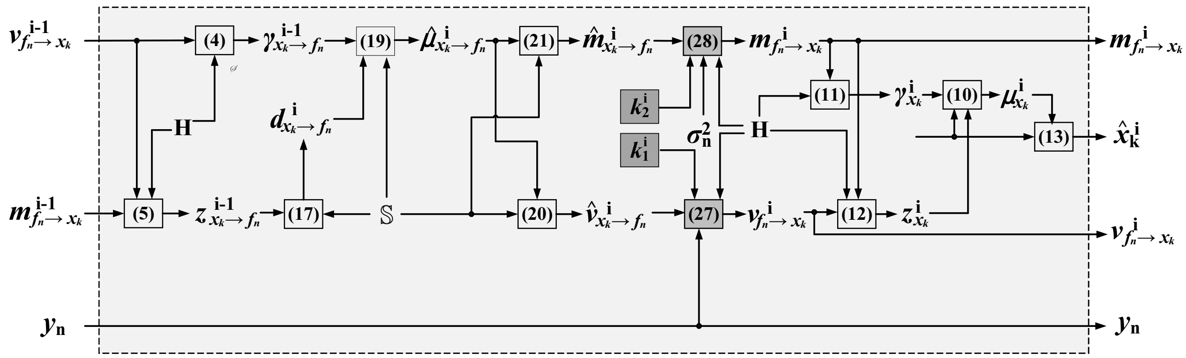

Fig. 3. Illustration of the i-layer of LC-GAMP with DL.

TABLE I
COMPLEXITY COMPARISONS IN $N \times K$ MASSIVE MIMO SYSTEMS UNDER M-QAM PER ITERATION

	multiplication	addition
RI	$(4K+2)K$	$(4K+2)K$
SA-GMP	$36NK$	$40NK-4N-4K$
CHEMP	$(2\sqrt{M}+1)(2K-1)2N$	$((2\sqrt{M}+1)(2K-1)-2)2N$
GAMP	$(24\sqrt{M}+24)NK$	$(20\sqrt{M}+28)NK-4N-4K$
LC-GAMP	$(24A+24)NK$	$(16A+4\sqrt{M}+28)NK-4N-4K$
Reduced number	$24(\sqrt{M}-A)NK$	$16(\sqrt{M}-A)NK$

IV. DEEP LEARNING AIDED LC-GAMP DETECTION

In this section, the proposed LC-GAMP detection algorithm is now improved with DL. First of all, it can be deep unfolded into a t-layer forward propagation network [9]. Then, to improve the detection performance, we introduce the trainable hyper-parameters into LC-GAMP.

In particular, according to (12), we can find that the estimation accuracy of $\hat{x}_k$ depends on the reliability of $\mu_{x_k}(x_k)$, which is estimated by the Gaussian function in (10) with the mean z_{x_k} and variance γ_{x_k}. Given it, we can introduce the hyper-parameters (k_1^i, k_2^i) to provide the appropriate step sizes for the update of (8) and (9), so as to improve the reliability of the equivalent mean and variance for $\mu_{x_k}(x_k)$. In this way, (8) and (9) can be modified as

$$\widetilde{m}^i_{f_n\to x_k} = y_n - k_1^i \sum_{k'\neq k} h_{n,k'} m^i_{x_{k'}\to f_n}, \tag{25}$$

$$\tilde{v}^i_{f_n\to x_k} = \sigma_n^2 + k_2^i \sum_{k'\neq k} h^2_{n,k'} v^i_{x_{k'}\to f_n}. \tag{26}$$

For a better understanding, the structure of the LC-GAMP with DL is illustrated in Fig. 3, which is a revised version of LC-GAMP by adding the learnable scalar variables (k_1^i, k_2^i). In this way, we are intended to obtain the optimal (k_1^i, k_2^i) by DL. Meanwhile, (k_1^i, k_2^i) are trained by minimizing the following mean squared error (MSE) loss function:

$$l\left(\mathbf{x}_l; \hat{\mathbf{x}}\right) = \frac{1}{t}\sum_{i=1}^{t} ||\mathbf{x}_l - \hat{\mathbf{x}}^i||^2. \tag{27}$$

Here $\hat{\mathbf{x}}^i$ is the estimation of the training data $\mathbf{x}_l$ at the i-th layer, and the outputs of all layers are taken into account.

Clearly, the only difference between LC-GAMP and LC-GAMP with DL lies in the computation of $m^i_{f_n\to x_k}$ and $v^i_{f_n\to x_k}$, where the introduced hyper-parameters (k_1^i, k_2^i) do not deteriorate the complexity. Therefore, the proposed LC-GAMP with DL maintains the same complexity order as LC-GAMP but realizes an improved detection performance.

V. SIMULATION RESULTS

A. Complexity Comparisons

We now perform the complexity comparisons in terms of the numbers of multiplication and addition operations per iteration, which are shown in Table I and Table II.

In particular, the difference of the complexities between GAMP and LC-GAMP comes from the computations of $\mu^i_{x_k\to f_n}(x_k)$, $m^i_{x_k\to f_n}$ and $v^i_{x_k\to f_n}$. From Table I, in LC-GAMP, the reduced ratio in terms of multiplication and addition operations are $\frac{\sqrt{M}-A}{\sqrt{M}+1}$ and $\frac{(16\sqrt{M}-A)NK}{(20\sqrt{M}+28)NK-4N-4K}$ compared to GAMP, respectively.

Moreover, the detailed numerical comparisons between GAMP and LC-GAMP algorithms under 64-QAM and $|\mathbb{A}| = 2$ are shown in Table II. Clearly, compared to the original GAMP algorithm, the proposed LC-GAMP algorithm can achieve about 67% complexity reduction per iteration, which means the proposed LC-GAMP algorithm is a better choice for massive MIMO detection with high order modulation.

B. Performance Comparisons

To fully illustrate the superiority of the proposed LC-GAMP and LC-GAMP with DL algorithms, the detection performance comparisons are presented in terms of the bit error rate (BER) in massive MIMO systems, where the iteration numbers of all

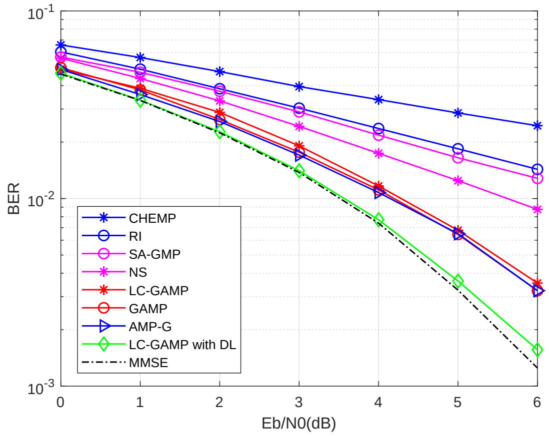

Fig. 4. BER performance comparison in a 128×16 massive MIMO system using 64-QAM.

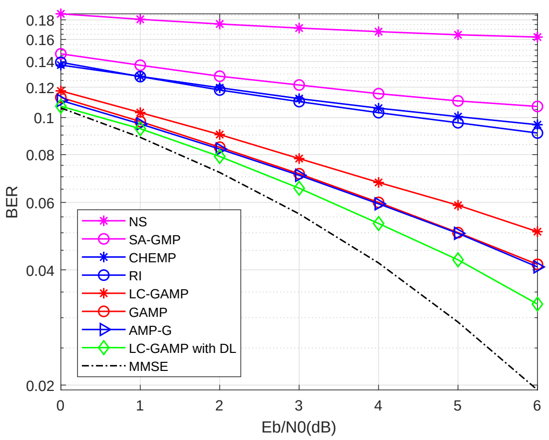

Fig. 5. BER performance comparison in a 128×32 massive MIMO system using 64-QAM.

iterative detection schemes are set as 3 with the approximation set size $|\mathbb{A}| = 2$ for the proposed LC-GAMP and LC-GAMP with DL.

TABLE II
NUMERICAL COMPARISONS UNDER 64-QAM AND $|\mathbb{A}| = 2$

	128×16		128×32	
	multiplication	addition	multiplication	addition
GAMP	442368	384448	884736	769408
LC-GAMP	147456	187840	294912	376192
Reduced number	294912	196608	589824	393216
Reduced ratio	66.7%	51.1%	66.7%	51.1%

Fig. 4 shows the BER performance in a 128×16 massive MIMO system using 64-QAM. As can be seen clearly, the detection performance of LC-GAMP is comparable to that of GAMP and AMP-G, and is much better than other MP-based schemes (CHEMP, SA-GMP), NS [10] and RI [10]. The same conclusion can be drawn from Fig. 5 where a 128×32 massive MIMO is applied, which is consistent with Lemma 1 in the previous section. Meanwhile, it can be found that the detection performance of LC-GAMP with DL is approaching to that of MMSE in the 128×16 system but with a growing performance gap in the 128×32 systems. This is because the system antenna ratio N/K has decreased while more iterations are required to ensure the performance of the MP-based schemes, and LC-GAMP with DL may outperform MMSE in respect of the detection performance if the iteration number is large enough. However, from Table I we can find that the complexity of the LC-GAMP algorithm is only $\mathcal{O}(ANKt)$, which is much lower than that of MMSE with $\mathcal{O}(N^3)$.

VI. CONCLUSION

In this paper, by using the strategy of constellation set approximation, the LC-GAMP detection algorithm for massive MIMO with high order modulation was proposed. By demonstration, significant complexity reduction can be achieved by the proposed scheme but only with negligible performance loss. Moreover, by introducing the trainable hyper-parameters to tune the update step size, the detection performance of the LC-GAMP algorithm was further improved with deep learning, which leads to a competitive detection trade-off between performance and complexity.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China under Grants No. 62371124, in part by ZTE Corporation under Research Program 2023ZTE01-04.

REFERENCES

[1] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro and C. Studer, "Large-Scale MIMO Detection for 3GPP LTE: Algorithms and FPGA Implementations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 916-929.

[2] A. -A. Lu, X. Gao, Y. R. Zheng and C. Xiao, "Low Complexity Polynomial Expansion Detector With Deterministic Equivalents of the Moments of Channel Gram Matrix for Massive MIMO Uplink," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 586-600, 2016.

[3] J. Hu and T. M. Duman, "Graph-Based Detector for BLAST Architecture," *2007 IEEE International Conference on Communications*, Glasgow, UK, 2007, pp. 1018-1023.

[4] T. L. Narasimhan and A. Chockalingam, "Channel Hardening-Exploiting Message Passing (CHEMP) Receiver in Large-Scale MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 847-860, 2014.

[5] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang and Q. Guo, "Low-Complexity Iterative Detection for Large-Scale Multiuser MIMO-OFDM Systems Using Approximate Message Passing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 902-915, 2014.

[6] L. Liu, C. Yuen, Y. L. Guan, Y. Li and C. Huang, "Gaussian Message Passing for Overloaded Massive MIMO-NOMA," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 210-226, 2019.

[7] L. Xiang, Y. Liu, L. -L. Yang and L. Hanzo, "Gaussian Approximate Message Passing Detection of Orthogonal Time Frequency Space Modulation," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10999-11004, 2021.

[8] Z. Zhang, Y. Dong, K. Long, X. Wang and X. Dai, "Decentralized Baseband Processing With Gaussian Message Passing Detection for Uplink Massive MU-MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 2152-2157, 2022.

[9] M. Miyoshi *et al.*, "Parameter-Learned AMP for MIMO Signal Detection," *2022 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, Seoul, Korea, Republic of, 2022, pp. 99-103.

[10] M. A. Albreem et al., "Low Complexity Linear Detectors for Massive MIMO: A Comparative Study," *IEEE Access*, vol. 9, pp. 45740-45753, 2021.