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

He Zhu, Beilei Gong and Zheng Wang

School of Information Science and Engineering,

Southeast University, Nanjing, China

IEEE ICCT 2023, Wuxi, China

20-22 October 2023

Contents

Introduction

◊ *System model*

Consider a real-valued uplink massive MIMO system

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

◊ *Motivation*

MMSE:

- ✓ Polynomial complexity
- ✗ Gram matrix computation
- ✗ High-dimension inversion

GAMP:

- ✓ Polynomial complexity
- ✗ High modulation order

Can we design? $\Longrightarrow$

CHEMP:

- ✓ Polynomial complexity
- ✗ Gram matrix computation
- ✗ High modulation order

LC-GAMP:

- ✓ Polynomial complexity

□ L. Xiang, "Gaussian Approximate Message Passing Detection of Orthogonal Time Frequency Space Modulation," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10999-11004, 2021.

Existing GAMP Algorithm

◊ *GAMP algorithm description*

In each iteration, the messages in both directions are updated alternately.

- The downward oriented message:

$$\mu^i_{x_k \to f_n}(x_k) \propto \frac{\mathcal{CN}\left(x_k; z^{i-1}_{x_k \to f_n}, \gamma^{i-1}_{x_k \to f_n}\right)}{\sum\limits_{x_k \in \mathbb{S}} \mathcal{CN}\left(x_k; z^{i-1}_{x_k \to f_n}, \gamma^{i-1}_{x_k \to f_n}\right)}. \quad (2)$$

- Approximated by minimizing the KL divergence

$$m^i_{x_k \to f_n} = \sum_{s \in \mathbb{S}} s \mu^i_{x_k \to f_n}(x_k = s), \quad (3)$$

$$v^i_{x_k \to f_n} = \sum_{s \in \mathbb{S}} |s|^2 \mu^i_{x_k \to f_n}(x_k = s) - |m^i_{x_k \to f_n}|^2. \quad (4)$$

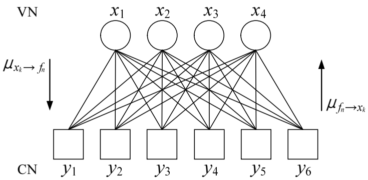

Figure: An Illustration of the FG with $N \times K = 6 \times 4$.

Existing GAMP Algorithm

◊ *GAMP algorithm description*

- The upward oriented message:

$$m^i_{f_n \to x_k} = y_n - \sum_{k' \neq k} h_{n,k'} m^i_{x_{k'} \to f_n}, \quad (5)$$

$$v^i_{f_n \to x_k} = \sigma^2_n + \sum_{k' \neq k} h^2_{n,k'} v^i_{x_{k'} \to f_n}. \quad (6)$$

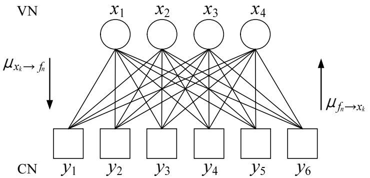

Figure: An Illustration of the FG with $N \times K = 6 \times 4$.

As the number of iterations is reached, the a *posteriori* probability $\mu_{x_k}(x_k)$ is expressed as

$$\mu_{x_k}(x_k) \propto \frac{\mathcal{CN}\left(x_k; z_{x_k}, \gamma_{x_k}\right)}{\sum_{x_k \in \mathbb{S}} \mathcal{CN}\left(x_k; z_{x_k}, \gamma_{x_k}\right)} \quad (7)$$

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

$$\hat{x}_k = \arg\min_{x_k \in \mathbb{S}} \left|\left| \sum_{s \in \mathbb{S}} s\mu_{x_k}(s) - x_k \right|\right|. \quad (8)$$

The Proposed LC-GAMP Algorithm

◊ *LC-GAMP algorithm description*

In GAMP, by rewriting (2)-(4), it can be found that

- (3) and (4) involve many multiplication and addition operations:

$$m^{i}_{x_k\to f_n} = s_1\mu^{i}_{x_k\to f_n}(s_1) + \ldots + s_{\sqrt{M}}\mu^{i}_{x_k\to f_n}(s_{\sqrt{M}}), \tag{9}$$

$$v^{i}_{x_k\to f_n} = |s_1|^2\mu^{i}_{x_k\to f_n}(s_1) + \ldots + |s_{\sqrt{M}}|^2\mu^{i}_{x_k\to f_n}(s_{\sqrt{M}}) - |m^{i}_{x_k\to f_n}|^2, \tag{10}$$

  which is rather sensitive to the increment of the modulation order.

- Many symbols in $\mathbb{S}$ correspond to very small values $\mu^{i}_{x_k\to f_n}$:

$$\mu^{i}_{x_k\to f_n}(x_k = s_j) \propto \frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\sum_{s_j\in\mathbb{S}}\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}. \tag{11}$$

  The size of $\mu^{i}_{x_k\to f_n}(x_k = s_j)$ heavily depends on $d^{i}_{x_k\to f_n}(s_j) = |s_j - z^{i-1}_{x_k\to f_n}|$.

The Proposed LC-GAMP Algorithm

◊ *LC-GAMP algorithm description*

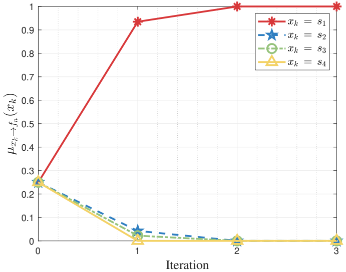

Figure: Illustration of $\mu^{i}_{x_k \to f_n}(x_k)$ for $x_k = s_1, \dots, s_4$ respectively in a 128×16 massive MIMO system using 16-QAM.

$$\mu^{i}_{x_k \to f_n}(x_k = s_1) \gg \mu^{i}_{x_k \to f_n}(x_k = s_2) > \dots \tag{12}$$

By calculating $d^{i}_{x_k \to f_n}(s_j) = |s_j - z^{i-1}_{x_k \to f_n}|$ in each iteration, we can easily find out the trivial symbols in $\mathbb{S}$ that can be ignored.

The Proposed LC-GAMP Algorithm

◊ *LC-GAMP algorithm description*

Ignore these trivial symbols during the iterations, i.e., substitute $\mathbb{S}$ with $\mathbb{A}$:

$$\hat{\mu}^{i}_{x_k \to f_n}(x_k = s_j) \propto \frac{\exp\left(\frac{-d^{i\,2}_{x_k \to f_n}(s_j)}{2\gamma^{i-1}_{x_k \to f_n}}\right)}{\sum_{s_j \in \mathbb{A}} \exp\left(\frac{-d^{i\,2}_{x_k \to f_n}(s_j)}{2\gamma^{i-1}_{x_k \to f_n}}\right)}. \tag{13}$$

Similarly,

$$\hat{m}^{i}_{x_k \to f_n} = s_1 \hat{\mu}^{i}_{x_k \to f_n}(s_1) + .. + s_{|\mathbb{A}|} \hat{\mu}^{i}_{x_k \to f_n}(s_{|\mathbb{A}|}), \tag{14}$$

$$\hat{v}^{i}_{x_k \to f_n} = |s_1|^2 \hat{\mu}^{i}_{x_k \to f_n}(s_1) + \ldots + |s_{|\mathbb{A}|}|^2 \hat{\mu}^{i}_{x_k \to f_n}(s_{|\mathbb{A}|}) - |\hat{m}^{i}_{x_k \to f_n}|^2. \tag{15}$$

Significant reduction in complexity can be achieved by the proposed approximation scheme.

The Proposed LC-GAMP Algorithm

◊ *Approximate Set Size Selection Criteria*

The size of $\mathbb{A}$ has a great impact on the approximation accuracy and implementation efficiency. Based on $\mu^i_{x_k\to f_n}$ and $\hat{\mu}^i_{x_k\to f_n}$, define the approximation error e

$$e = \sum_{s_j\in\mathbb{S}} \frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} - \sum_{s_j\in\mathbb{A}} \frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} = \sum_{s_j\notin\mathbb{A}, s_j\in\mathbb{S}} \frac{\exp\left(\frac{-d^{i\,2}_{x_k\to f_n}(s_j)}{2\gamma^{i-1}_{x_k\to f_n}}\right)}{\rho^i_{n,k}(\mathbb{S})} \tag{16}$$

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

By letting $|\mathbb{A}| = 2N$, we can arrive at the following results.

The Proposed LC-GAMP Algorithm

◊ *Approximate Set Size Selection Criteria*

Lemma 1: The approximation error e in (16) decays exponentially with the increment of N as

$$e < \alpha^{-4N^2} \tag{17}$$

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

- The approximation error is negligible even very few symbols are considered in $\mathbb{A}$, since it **decays rapidly** in an **exponential** way.
- $\boldsymbol{N = 1}$ is recommended for practical signal detection, which means only the first two constellation symbols are employed during each iteration.

Deep learning aided LC-GAMP detection

The estimation accuracy depends on the reliability of $\mu^{i}_{x_k \to f_n}(x_k)$, which is estimated by Gaussian function with the mean $z^{i-1}_{x_k \to f_n}$ and variance $\gamma^{i-1}_{x_k \to f_n}$.

$$\gamma^{i-1}_{x_k \to f_n} = \left(\sum_{n' \neq n} \frac{h^2_{n',k}}{v^{i-1}_{f_{n'} \to x_k}} \right)^{-1}, \tag{18}$$

$$z^{i-1}_{x_k \to f_n} = \gamma^{i-1}_{x_k \to f_n} \sum_{n' \neq n} \frac{h_{n',k} m^{i-1}_{f_{n'} \to x_k}}{v^{i-1}_{f_{n'} \to x_k}}. \tag{19}$$

Introduce the **learnable hyper-parameters** to provide the appropriate step sizes for the update of $m^{i}_{f_n \to x_k}$ and $v^{i}_{f_n \to x_k}$.

$$\widetilde{m}^{i}_{f_n \to x_k} = y_n - \boldsymbol{k}^{i}_{1} \sum_{k' \neq k} h_{n,k'} m^{i}_{x_{k'} \to f_n}, \tag{20}$$

$$\tilde{v}^{i}_{f_n \to x_k} = \sigma^2_n + \boldsymbol{k}^{i}_{2} \sum_{k' \neq k} h^2_{n,k'} v^{i}_{x_{k'} \to f_n}. \tag{21}$$

Deep learning aided LC-GAMP detection

For a better understanding, the structure of the LC-GAMP with DL is illustrated

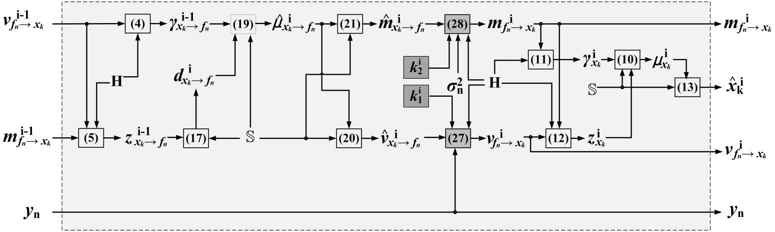

Figure: Illustration of the i-layer of LC-GAMP with DL.

Meanwhile, (k_1^i, k_2^i) are trained by minimizing the following mean squared error (MSE) loss function:

$$l\left(\mathbf{x}_l;\hat{\mathbf{x}}\right) = \frac{1}{t}\sum_{i=1}^{t} ||\mathbf{x}_l - \hat{\mathbf{x}}^i||^2. \qquad (22)$$

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

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

Simulation Results

◊ *Complexity Comparisons*

Table: 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	$(24\lvert\mathbb{A}\rvert+24)NK$	$(16\lvert\mathbb{A}\rvert+4\sqrt{M}+28)NK-4N-4K$
Reduced number	$24(\sqrt{M}-\lvert\mathbb{A}\rvert)NK$	$16(\sqrt{M}-\lvert\mathbb{A}\rvert)NK$

The reduced ratio of multiplication and addition operations are $\frac{\sqrt{M}-|\mathbb{A}|}{\sqrt{M}+1}$ and $\frac{(16\sqrt{M}-|\mathbb{A}|)NK}{(20\sqrt{M}+28)NK-4N-4K}$ compared to GAMP, respectively.

Table: Numerical Comparisons under 64-QAM and $|\mathbb{A}|=2$ per Iteration

	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%**

Simulation Results

◊ *Performance Comparisons*

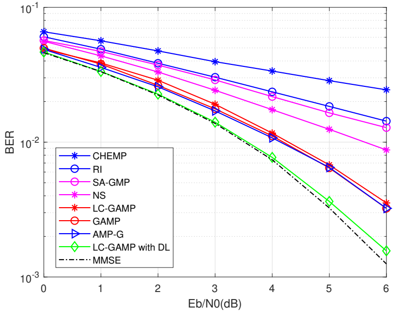

Figure: In a 128 × 16 MIMO system using 64-QAM.

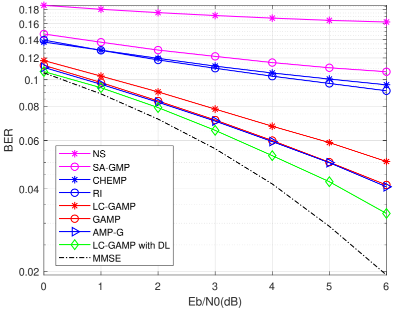

Figure: In a 128 × 32 MIMO system using 64-QAM.

- The detection performance of LC-GAMP is comparable to that of GAMP and AMP-G, and is much better than other MP-based schemes, NS and RI.
- The assistance of deep learning leads to improved detection performance.