Expectation Propagation for Semi-Blind Channel Estimation in Cell-Free Networks

Zilu Zhao, Dirk Slock
Communication Systems Department, EURECOM, France
zilu.zhao@eurecom.fr, dirk.slock@eurecom.fr

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. This paper examines the role of Cell-Free (CF) Massive MIMO (MaMIMO) in advancing wireless communication networks, particularly for beyond 5G and 6G networks. Building on the foundational work by Ngo et al., CF MaMIMO, with its distributed architecture, addresses the demands for high data rates, uniform quality of service (QoS), and power efficiency. A central challenge in CF networks is pilot contamination, arising from the absence of traditional cellular boundaries and an excess of user terminals (UTs) relative to pilot sequences. We introduce an Expectation Propagation (EP)-based method for Semi-Blind bilinear estimation in CF MaMIMO networks, providing a low-complexity solution by utilizing the Central Limit Theorem. This method enhances scalability and efficiency compared to existing approaches. Additionally, we propose a shift from distributed to decentralized EP, allowing for local information sharing among Access Points (APs) about user signals.**

I. Introduction

Cell-Free (CF) Massive MIMO (MaMIMO), an evolution of the traditional MaMIMO systems, represents a significant leap in the field of advanced wireless communications. First introduced by Ngo et al. in their seminal work [1], this innovative concept reimagines wireless networks as distributed architectures intricately linked with a central processing unit (CPU). As the technological world strides towards the era of beyond 5G networks, Cell-Free MaMIMO has garnered substantial interest from both the academic and industrial sectors, as highlighted by Elhoushy et al. [2].

One of the unique features of Cell-Free networks is the absence of traditional cellular boundaries, which, while beneficial in many aspects, introduces a critical challenge: the potential for a high number of user terminals (UTs) in a given area to exceed the length of the pilot sequence. This imbalance leads to what is known as pilot contamination, a phenomenon that can severely impact the performance of Cell-Free MaMIMO networks. Addressing this issue, Semi-Blind approaches have been explored as a viable solution to mitigate the effects of pilot contamination, as detailed in the work of Gholami et al. [3]. In their study, the authors conceptualize the channel as a deterministic unknown parameter and delve into the analysis of crucial aspects of the model, such as the Cramer-Rao bound and identifiability.

In the context of Bayesian inference, the Semi-Blind approach is modeled as a bilinear inference problem, where the Access Points (APs) must concurrently estimate both the channel state information (CSI) and the user signals. Message Passing algorithms, renowned for their efficacy in various inference problems by leveraging the structures of probability models, play a critical role here. Notably, one of the most potent message passing algorithms is Expectation Propagation (EP), proposed by Minka [4]. EP, considered a form of (loopy) belief propagation (BP) with approximations for simplifying the inference problem, has shown promising results. They have been shown to have the same fixed point as the Bethe Free Energy [5]. However, a comprehensive understanding of EP's convergence properties remains an area yet to be fully explored.

Based on variable level EP (VL-EP), a centralized Bayesian methodology has been explored by Gholami et al. [6], offering a novel approach in the context of Cell-Free MaMIMO networks. A more recent development is the distributed method proposed by Karataev et al. [7]. This approach distributes the computational load by enabling each Access Point (AP) to carry out part of the computation, potentially enhancing the scalability and efficiency of the network. Approximate Message Passing (AMP) algorithms have been introduced to reduce the number of messages [8]. In [9], the authors further generalized the GAMP algorithm [10] to bilinear problems.

In [11], the author proposed Decentralized GAMP (D-GAMP), which is a decentralized method for signal recovery in linear systems. This can be viewed as a combination of consensus propagation [12] and AMP.

A. Main Contributions

In this article, we develop an EP-based method to tackle the Semi-Blind estimation problem. Compared to [7], our factorization provides a low-complexity way of estimating channels from the pilot symbols and prior information. By exploiting Central Limit Theory, our factorization scheme simplifies the computation compared to [7]. However, each AP still needs to perform matrix inversion.

Inspired by [12] and [11], we can transform the distributed EP into a decentralized EP by sharing the information of the UT signals locally among the AP.

II. System Model

We consider a semi-blind signal model at the l-th access point (AP),

$$\mathbf{Y}_l = \begin{bmatrix}\mathbf{Y}_{p,l} & \mathbf{Y}_{d,l}\end{bmatrix} = \boldsymbol{H}_l \begin{bmatrix}\mathbf{X}_p & \mathbf{X}_d\end{bmatrix} + \mathbf{V}_l. \qquad (1)$$

The received signals $\mathbf{Y}_l$ is composed of received pilot signals $\mathbf{Y}_{p,l} \in \mathbb{C}^{N\times P}$ and received data signals $\mathbf{Y}_{d,l} \in \mathbb{C}^{N\times T}$. The channels between different users are considered Gaussian distributed and independent i.e. $\mathrm{vec}(\boldsymbol{H}_l) \sim \mathcal{CN}(\mathbf{0}, \Xi_l)$ where $\Xi_l \in \mathbb{C}^{NK\times NK}$ is a block diagonal matrix of K blocks $\Xi_{kl} \in \mathbb{C}^{N\times N}$. The transmitted symbols can be decomposed as pilot symbols $\mathbf{X}_p \in \mathcal{S}^{K\times P}$ and data symbols $\mathbf{X}_d \in \mathcal{S}^{K\times T}$, where $\mathcal{S}$ is the constellation set. The elements in $\mathbf{X}_d$ are i.i.d.. The noise is considered as i.i.d. Gaussian distribution, and thus, $\mathrm{vec}(\mathbf{V}_l) \sim \mathcal{CN}(0, \sigma_v^2\mathbf{I})$.

The direct MMSE estimate of $\boldsymbol{H}_l$ can be obtained by $p(\boldsymbol{H}_l|\mathbf{Y}_{p,l})$ which is equivalent to a Gaussian linear model

$$\text{vec}(\mathbf{Y}_{p,l}) = (\mathbf{X}_p^T \otimes \mathbf{I}_N) \cdot \text{vec}(\boldsymbol{H}_l) + \text{vec}(\mathbf{V}_l). \quad (2)$$

If there is pilot contamination, the direct MMSE approach contains a matrix inversion of size $NK \times NK$.
To simplify the computation, we obtain the factorization scheme from the following probabilistic model

$$\begin{aligned} &p(\{\mathbf{Y}_{p,l}\}, \{\mathbf{Y}_{d,l}\}, \{\boldsymbol{H}_l\}, \mathbf{X}_d, \{\mathbf{V}_l\}) \\ =&\prod_{k,t} p(x_{d,kt}) \prod_l \prod_{t_1=1}^{T} p(\boldsymbol{y}_{d,lt_1}|\boldsymbol{H}_l, \boldsymbol{x}_{d,:t_1}) \\ &\cdot \prod_{t_2=1}^{P} p(\boldsymbol{y}_{p,lt_2}|\boldsymbol{H}_l) \prod_k p(\mathbf{h}_{lk}). \end{aligned} \quad (3)$$

By exploiting the structure of the factorization scheme, we come up with an EP algorithm.

III. Expectation Propagation

Expectation Propagation is a method to approximate the factors by distributions of the desired form [13]. The projection of a given distribution p into a target family F is [10]

$$\text{proj}(p) = \arg\min_{q \in F} KLD(p\|q), \quad (4)$$

where $KLD(p\|q) = \int p(x) \ln \frac{p(x)}{q(x)} dx$ is the Kullback–Leibler divergence. With a given factorization, the update algorithm in EP can be interpreted as message passing: [10]

$$\begin{aligned} \mu_{x_i;\Psi}(x_i) &= \prod_{\Phi \neq \Psi} \mu_{\Phi;x_i}(x_i); \\ \mu_{\Psi;x_i}(x_i) &= \frac{\text{proj}(b_\Psi(x_i))}{\mu_{x_i;\Psi}(x_i)}, \end{aligned} \quad (5)$$

where x_i represents one variable node, Ψ, Φ denote factor nodes and $b_\Psi(x_i)$ is the belief of x_i at node Ψ:

$$b_\Psi(x_i) = \mu_{x_i;\Psi}(x_i) \int \Psi(\boldsymbol{x}) \prod_{j \neq i} \mu_{x_j;\Psi}(x_j) d\boldsymbol{x}_{\bar{i}}. \quad (6)$$

In (6), we use $\boldsymbol{x}_{\bar{i}}$ to denote all elements in $\boldsymbol{x}$ except the i-th one. Although it is not necessary to normalize the messages, the messages appearing in this paper are all normalized to 1. For simplicity, we use similar notations in [7] and define the factors

$$\begin{aligned} &\Psi_{1,kt} = p_{x_{d,kt}};\ \Psi_{2,lt} = p_{\boldsymbol{y}_{d,lt|\boldsymbol{H}_l, x_{d,:t}}} \\ &\Psi_{3,lt} = p_{\boldsymbol{y}_{p,lt|\boldsymbol{H}_l}},\ \Psi_{4,lk} = p_{\mathbf{h}_{lk}}. \end{aligned} \quad (7)$$

Furthermore, we define the projection family of belief of $x_{d,kt}$ to be the categorical distribution [7]. All the beliefs of $\mathbf{h}_{lk}$ are projected to Gaussian.

IV. Update Rules

A. Messages from $\Psi_{1,kt}$

The factor $\Psi_{1,kt}$ contains only one variable. The message from $\Psi_{1,kt}$ to $x_{d,kt}$ can be computed directly since no projection is needed

$$\mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt}) = p(x_{d,kt}). \quad (8)$$

B. Belief of $x_{d,kt}$ at $\Psi_{2,lt}$

The factor $\Psi_{2,lt}$ is connected to $\forall k, \mathbf{h}_{lk}, x_{d,kt}$.
According to the EP rule, the message to $x_{d,kt}$ is

$$\mu_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) = \frac{\text{proj}[b_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt})]}{\mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt})}, \quad (9)$$

where the belief is defined as an approximated posterior for $x_{d,kt}$:

$$\begin{aligned} &b_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \propto \sum_{\boldsymbol{x}_{d,\bar{k}t}} \int p(\boldsymbol{y}_{d,lt}| \sum_i x_{d,it}\mathbf{h}_{li}) \\ &\cdot \prod_i \mu_{\mathbf{h}_{li};\Psi_{2,lt}}(\mathbf{h}_{li}) \mu_{x_{d,it};\Psi_{2,lt}}(x_{d,it}) d\boldsymbol{H}_l \\ &= \mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt}) \sum_{\boldsymbol{x}_{d,\bar{k}t}} \int p(\boldsymbol{y}_{d,lt}|x_{d,kt}\mathbf{h}_{lk} + \sum_{i \neq k} x_{d,it}\mathbf{h}_{li}) \\ &\cdot \mu_{\mathbf{h}_{lk};\Psi_{2,lt}}(\mathbf{h}_{lk}) \\ &\cdot \prod_{i \neq k} \mu_{\mathbf{h}_{li};\Psi_{2,lt}}(\mathbf{h}_{li}) \mu_{x_{d,it};\Psi_{2,lt}}(x_{d,it}) d\boldsymbol{H}_l. \end{aligned} \quad (10)$$

We use the notation $\boldsymbol{x}_{d,\bar{k}t}$ to denote all the elements in $\boldsymbol{x}_{d,:t}$ except the k-th element. The integral (and summation) in (10) can be considered as a marginalization operation with hypothetic distributions given by the messages. Due to the Central Limit Theory (CLT), we approximate $\sum_{i \neq k} x_{d,it}\mathbf{h}_{li}$ to a Gaussian by using $\prod_{i \neq k} \mu_{\mathbf{h}_{li};\Psi_{2,lt}}(\mathbf{h}_{li}) \mu_{x_{d,it};\Psi_{2,lt}}(x_{d,it})$ as their joint distribution [10]. Therefore, (10) becomes

$$\begin{aligned} &b_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \propto \mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt}) \\ &\cdot \int\int p(\boldsymbol{y}_{d,lt}|x_{d,kt}\mathbf{h}_{lk} + \mathbf{z}_{d,lkt}) \mu_{\mathbf{z}_{d,lkt}}(\mathbf{z}_{d,lkt}) d\mathbf{z}_{d,lkt} \\ &\cdot \mu_{\mathbf{h}_{lk};\Psi_{2,lt}}(\mathbf{h}_{lk}) d\mathbf{h}_{lk}, \end{aligned} \quad (11)$$

where $\mu_{\mathbf{z}_{d,lkt}}(\mathbf{z}_{d,lkt}) = \mathcal{CN}(\mathbf{z}_{d,lkt}|\mathbf{m}_{\mathbf{z}_{d,lkt}}, \mathbf{C}_{\mathbf{z}_{d,lkt}})$ is computed as

$$\begin{aligned} \mathbf{m}_{\mathbf{z}_{d,lkt}} &= \sum_{i \neq k} m_{x_{d,it};\Psi_{2,lt}} \mathbf{m}_{\mathbf{h}_{li};\Psi_{2,lt}} \\ \mathbf{C}_{\mathbf{z}_{d,lkt}} &= \sum_{i \neq k} \tau_{x_{d,it};\Psi_{2,lt}} \mathbf{C}_{\mathbf{h}_{li};\Psi_{2,lt}} \\ &+ \tau_{x_{d,it};\Psi_{2,lt}} \mathbf{m}_{\mathbf{h}_{it};\Psi_{2,lt}} \mathbf{m}^H_{\mathbf{h}_{it};\Psi_{2,lt}} + |m_{x_{d,it};\Psi_{2,lt}}|^2 \mathbf{C}_{\mathbf{h}_{li};\Psi_{2,lt}} \end{aligned} \quad (12)$$

By applying the Gaussian reproduction lemma [8], and the fact that Gaussian distribution integrates to one, the belief (11) becomes

$$\begin{aligned} &b_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \propto \mathcal{CN}(\mathbf{0}|\boldsymbol{y}_{d,lt} - \mathbf{m}_{\mathbf{z}_{d,lkt}} - x_{d,kt}\mathbf{m}_{\mathbf{h}_{lk};\Psi_{2,lt}}, \\ &\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{d,lkt}} + |x_{d,kt}|^2 \mathbf{C}_{\mathbf{h}_{lk};\Psi_{2,lt}}) \cdot \mu_{x_{d,kt};\Psi_{2,lt}(x_{d,kt})}. \end{aligned} \quad (13)$$

This distribution is already a categorical distribution. Therefore, the outbound message is

$$\begin{aligned} &\mu_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \propto \mathcal{CN}(\mathbf{0}|\boldsymbol{y}_{d,lt} - \mathbf{m}_{\mathbf{z}_{d,lkt}} - x_{d,kt}\mathbf{m}_{\mathbf{h}_{lk};\Psi_{2,lt}}, \\ &\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{d,lkt}} + |x_{d,kt}|^2 \mathbf{C}_{\mathbf{h}_{lk};\Psi_{2,lt}}) \end{aligned} \quad (14)$$

C. Belief of $\mathbf{h}_{lk}$ at $\Psi_{2,lt}$

For the channel variable nodes, the message to $\mathbf{h}_{lk}$ is

$$\mu_{\Psi_{2,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \frac{\text{proj}[b_{\Psi_{2,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk})]}{\mu_{\mathbf{h}_{lk};\Psi_{2,lt}}(\mathbf{h}_{lk})}, \tag{15}$$

where the belief is defined as

$$\begin{aligned} & b_{\Psi_{2,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \propto \sum_{\boldsymbol{x}_{d,:t}} \int p(\boldsymbol{y}_{d,lt} | \sum_i x_{d,it}\mathbf{h}_{li}) \\ & \cdot \prod_i \mu_{\mathbf{h}_{li};\Psi_{2,lt}}(\mathbf{h}_{li}) \mu_{x_{d,it};\Psi_{2,lt}}(x_{d,it}) d\boldsymbol{H}_{l\overline{k}} \\ & = \mu_{\mathbf{h}_{lk};\Psi_{2,lt}}(\mathbf{h}_{lk}) \sum_{\boldsymbol{x}_{d,:t}} \int p(\boldsymbol{y}_{d,lt} | x_{d,kt}\mathbf{h}_{lk} + \sum_{i\neq k} x_{d,it}\mathbf{h}_{li}) \\ & \cdot \mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt}) \\ & \cdot \prod_{i\neq k} \mu_{\mathbf{h}_{li};\Psi_{2,lt}}(\mathbf{h}_{li}) \mu_{x_{d,it};\Psi_{2,lt}}(x_{d,it}) d\boldsymbol{H}_{l\overline{k}}. \end{aligned} \tag{16}$$

We use $\boldsymbol{H}_{l\overline{k}}$ to denote all the column vectors in $\boldsymbol{H}_l$ except the k-th column. By using the same approach from (10) to (13), and separating the terms that contains only $x_{d,kt}$ [10] [7], the belief (16) becomes

$$\begin{aligned} & b_{\Psi_{2,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \\ & = \mathbb{E}_{b_{\Psi_{2,lt};x_{d,kt}}} \{\mathcal{CN}[\mathbf{h}_{lk} | \mathbf{m}_{d,lkt}(x_{d,kt}), \mathbf{C}_{d,lkt}(x_{d,kt})]\} \end{aligned} \tag{17}$$

where the functions $\mathbf{m}_{d,lkt}(\cdot)$ and $\mathbf{C}_{d,lkt}(\cdot)$ are defined as

$$\begin{aligned} & \mathbf{C}_{d,lkt}(x) = [|x|^2(\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{d,lkt}})^{-1} + \mathbf{C}^{-1}_{\mathbf{h}_{lk};\Psi_{2,lt}}]^{-1} \\ & \mathbf{m}_{d,lkt}(x) = \mathbf{C}_{d,lkt}(x) \Big[|x|^2 (\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{d,lkt}})^{-1} \frac{\boldsymbol{y}_{d,lt} - \mathbf{m}_{\mathbf{z}_{d,lkt}}}{x} \\ & + \mathbf{C}^{-1}_{h_{lk};\Psi_{2,lt}} \mathbf{m}_{h_{lk};\Psi_{2,lt}} \Big]. \end{aligned} \tag{18}$$

Remark. *We should always use the latest belief of $x_{d,kt}$ when using* (17)*. If we use an ordering such that the message to $\Psi_{2,lt}$ is updated between the update of $\mu_{\Psi_{2,lt};x_{d,kt}}$ and the update of $\mu_{\Psi_{2,lt};\mathbf{h}_{lk}}$, we should recalculate* (13) *before using* (17)*.*

From the belief distribution (17), we can update the approximated posterior mean $\mathbf{m}_{\widehat{\mathbf{h}}^2_{lk}}$ and covariance of $\mathbf{C}_{\widehat{\mathbf{h}}^2_{lk}}$ by

$$\begin{aligned} & \mathbf{m}_{\widehat{\mathbf{h}}^2_{lk}} = \mathbb{E}_{b_{\Psi_{2,lt};x_{d,kt}}}[\mathbf{m}_{d,lkt}(x_{d,kt})] \\ & \mathbf{C}_{\widehat{\mathbf{h}}^2_{lk}} = \mathbb{E}_{b_{\Psi_{2,lt};x_{d,kt}}}[\mathbf{C}_{d,lkt}(x_{d,kt}) \\ & + \mathbf{m}_{d,lkt}(x_{d,kt})\mathbf{m}_{d,lkt}(x_{d,kt})^H] - \mathbf{m}_{\widehat{\mathbf{h}}^2_{lk}}\mathbf{m}^H_{\widehat{\mathbf{h}}^2_{lk}}. \end{aligned} \tag{19}$$

D. Belief of $\mathbf{h}_{lk}$ at $\Psi_{3,lt}$

The update of the message from $\Psi_{3,lt}$ to $\mathbf{h}_{lk}$ is carried out as

$$\mu_{\Psi_{3,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \frac{\text{proj}[b_{\Psi_{3,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk})]}{\mu_{\mathbf{h}_{lk};\Psi_{3,lt}}(\mathbf{h}_{lk})}, \tag{20}$$

This update operation uses the belief given by

$$\begin{aligned} & b_{\Psi_{3,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \propto \mu_{\mathbf{h}_{lk};\Psi_{3,lt}}(\mathbf{h}_{lk}) \\ & \cdot \int p(\boldsymbol{y}_{p,lt} | x_{p,kt}\mathbf{h}_{lk} + \sum_{i\neq k} x_{p,it}\mathbf{h}_{li}) \prod_{i\neq k} \mu_{\mathbf{h}_{il};\Psi_{3,lt}}(\mathbf{h}_{il}) d\boldsymbol{H}_{l\overline{k}}. \end{aligned} \tag{21}$$

Similar to (11), we can view $\sum_{i\neq k} x_{p,it}\mathbf{h}_{li}$ as a linear transformation of random vector $\mathbf{h}_{li}$ with joint distribution $\prod_{i\neq k} \mu_{\mathbf{h}_{il};\Psi_{3,lt}}(\mathbf{h}_{il})$ (difference is that here we don't need CLT). Like before, we denote the transformed random vector as $\mathbf{z}_{p,lkt}$ with distribution

$$\mu_{\mathbf{z}_{p,lkt}}(\mathbf{z}_{p,lkt}) = \mathcal{CN}(\mathbf{z}_{p,lkt} | \mathbf{m}_{\mathbf{z}_{p,lkt}}, \mathbf{C}_{\mathbf{z}_{p,lkt}}), \tag{22}$$

where

$$\begin{aligned} \mathbf{m}_{\mathbf{z}_{p,lkt}} &= \sum_{i\neq k} x_{p,it}\mathbf{m}_{\mathbf{h}_{il};\Psi_{3,lt}} \\ \mathbf{C}_{\mathbf{z}_{p,lkt}} &= \sum_{i\neq k} |x_{p,it}|^2 \mathbf{C}_{\mathbf{h}_{il};\Psi_{3,lt}}. \end{aligned} \tag{23}$$

By using the Gaussian reproduction lemma, we derive the belief as

$$b_{\Psi_{3,lt};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \mathcal{CN}(\mathbf{h}_{lk} | \mathbf{m}_{\widehat{\mathbf{h}}^3_{lk}}, \mathbf{C}_{\widehat{\mathbf{h}}^3_{lk}}), \tag{24}$$

where

$$\begin{aligned} \mathbf{C}_{\widehat{\mathbf{h}}^3_{lk}} &= [|x_{p,kt}|^2(\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{p,lkt}})^{-1} + \mathbf{C}^{-1}_{\mathbf{h}_{lk};\Psi_{3,lt}}]^{-1} \\ \mathbf{m}_{\widehat{\mathbf{h}}^3_{lk}} &= \mathbf{C}_{\widehat{\mathbf{h}}^3_{lk}} \Bigg[|x_{p,kt}|^2 (\mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{p,lkt}})^{-1} \frac{\boldsymbol{y}_{p,lt} - \mathbf{m}_{\mathbf{z}_{p,lkt}}}{x_{p,kt}} \\ & \qquad + \mathbf{C}^{-1}_{\mathbf{h}_{lk};\Psi_{3,lt}} \mathbf{m}_{\mathbf{h}_{lk};\Psi_{3,lt}} \Bigg]. \end{aligned} \tag{25}$$

We observe that (25) has an identical structure as (18). The only difference is that (25) operates within the neighborhood of $\Psi_{3,lt}$ while (18) works in the neighborhood of $\Psi_{2,lt}$. An intuitive explanation is that if we consider $x_{d,kt}$ as a random variable with probability one of being a known value, then the derivation in Section IV-C degrades to steps in Section IV-D.

E. Belief of $\mathbf{h}_{lk}$ at $\Psi_{4,lk}$

The message from $\Psi_{4,lk}$ to $\mathbf{h}_{lk}$ is

$$\mu_{\Psi_{4,lk};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \frac{\text{proj}[b_{\Psi_{4,lk};\mathbf{h}_{lk}}(\mathbf{h}_{lk})]}{\mu_{\mathbf{h}_{lk};\Psi_{4,lk}}(\mathbf{h}_{lk})}, \tag{26}$$

where the belief is defined as

$$\begin{aligned} & b_{\Psi_{4,lk};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \propto \mu_{\mathbf{h}_{lk};\Psi_{4,lk}}(\mathbf{h}_{lk}) p(\mathbf{h}_{lk}) \\ & = \mathcal{CN}(\mathbf{h}_{lk} | \mathbf{m}_{\mathbf{h}_{lk};\Psi_{4,lk}}, \mathbf{C}_{\mathbf{h}_{lk};\Psi_{4,lk}}) \mathcal{CN}(\mathbf{h}_{lk} | \mathbf{0}, \mathbf{\Xi}_{lk}). \end{aligned} \tag{27}$$

Applying the Gaussian reproduction lemma results to

$$b_{\Psi_{4,lk};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \mathcal{CN}(\mathbf{h}_{lk} | \mathbf{m}_{\widehat{\mathbf{h}}^4_{lk}}, \mathbf{C}_{\widehat{\mathbf{h}}^4_{lk}}), \tag{28}$$

where

$$\begin{aligned} \mathbf{C}_{\widehat{\mathbf{h}}^4_{lk}} &= (\mathbf{C}^{-1}_{\mathbf{h}_{lk};\Psi_{4,lk}} + \mathbf{\Xi}^{-1}_{lk})^{-1} \\ \mathbf{m}_{\widehat{\mathbf{h}}^4_{lk}} &= \mathbf{C}_{\widehat{\mathbf{h}}^4_{lk}} (\mathbf{C}^{-1}_{\mathbf{h}_{lk};\Psi_{4,lk}} \mathbf{m}_{\mathbf{h}_{lk};\Psi_{4,lk}}). \end{aligned} \tag{29}$$

For low rank $\mathbf{\Xi}_{lk}$ we can use matrix inversion lemma to get $\mathbf{C}_{\widehat{\mathbf{h}}^4_{lk}}$.

V. Projection Methods

To decrease the amount of matrix inversions when the number of antennas per AP is high, we can project all the messages to $\mathbf{h}_{lk}$ to Gaussian with multiple identities as covariance matrices. The projection from an arbitrary Gaussian to one with multiple of identities as a covariance matrix can be derived by setting the partial derivative [14] to zero:

$$\begin{aligned}&\forall \mathbf{m}_{\boldsymbol{x}} \forall \mathbf{C}_{\boldsymbol{x}}, \arg\min_{\mathbf{m},\tau} KLD[\mathcal{N}(\boldsymbol{x}|\mathbf{m}_{\boldsymbol{x}}, \mathbf{C}_{\boldsymbol{x}}) \| \mathcal{N}(\boldsymbol{x}|\mathbf{m}, \overline{\tau}\mathbf{I})] \\ &\qquad\qquad \Rightarrow \mathbf{m} = \mathbf{m}_{\boldsymbol{x}};\ \overline{\tau} = \frac{1}{N}\text{tr}(\mathbf{C}_{\boldsymbol{x}}).\end{aligned} \tag{30}$$

With this simplification the update at $\Psi_{4,lk}$ becomes $O(N^2)$ if the eigenvalues and eigenvectors of Ξ_{lk} are known. However, since $\mathbf{C}_{\mathbf{z}_{d,lkt}}$ contains a sum of non-diagonal matrices, the matrix inversion in (18) will have a complexity of $O(N^3)$.

A. Message to $\mathbf{h}_{lk}$

Since we have already derived the belief of $\mathbf{h}_{lk}$ at each factor, we only give the message from a general node Ψ to $\mathbf{h}_{lk}$ as an example. Suppose the belief of $\mathbf{h}_{lk}$ at Ψ to be

$$b_{\Psi;\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \mathcal{CN}(\mathbf{h}_{lk}|\mathbf{m}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}, \mathbf{C}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}). \tag{31}$$

The projection gives

$$\text{proj}[b_{\Psi;\mathbf{h}_{lk}}(\mathbf{h}_{lk})] = \mathcal{CN}(\mathbf{h}_{lk}|\mathbf{m}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}, \overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}\mathbf{I}), \tag{32}$$

where $\overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}} = \text{tr}(\mathbf{C}_{\widehat{\mathbf{h}}_{lk}^{\Psi}})/N$. Use the ansatz that the message from $\mathbf{h}_{lk}$ to Ψ is Gaussian with multiple of identities as covariance matrix, i.e. $\mathbf{C}_{\mathbf{h}_{lk};\Psi} := \overline{\tau}_{\mathbf{h}_{lk};\Psi}\mathbf{I}$. Thus, the message from Ψ to $\mathbf{h}_{lk}$ becomes

$$\mu_{\Psi;\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \mathcal{CN}(\mathbf{h}_{lk}|\mathbf{m}_{\Psi;\mathbf{h}_{lk}}, \overline{\tau}_{\Psi;\mathbf{h}_{lk}}\mathbf{I}), \tag{33}$$

where

$$\begin{aligned}\overline{\tau}_{\Psi;\mathbf{h}_{lk}} &= \frac{\overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}\overline{\tau}_{\mathbf{h}_{lk};\Psi}}{\overline{\tau}_{\mathbf{h}_{lk};\Psi} - \overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}} \\ \mathbf{m}_{\Psi;\mathbf{h}_{lk}} &= \frac{\overline{\tau}_{\mathbf{h}_{lk};\Psi}\mathbf{m}_{\widehat{\mathbf{h}}_{lk}^{\Psi}} - \overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}\mathbf{m}_{\mathbf{h}_{lk};\Psi}}{\overline{\tau}_{\mathbf{h}_{lk};\Psi} - \overline{\tau}_{\widehat{\mathbf{h}}_{lk}^{\Psi}}}.\end{aligned} \tag{34}$$

This operation may lead to negative results for the variance estimate. To avoid negative variance, we can set a lower bound ϵ, and clip $\overline{\tau}_{\Psi;\mathbf{h}_{lk}}$ above ϵ.

B. Message to $x_{d,kt}$

Since we define the target family for the projection of $b_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt})$ to be a categorical distribution, which is already one, we can directly derive the message to $x_{d,kt}$

$$\begin{aligned}\mu_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \propto \mathcal{CN}(\mathbf{0}|\boldsymbol{y}_{d,kt} - \mathbf{m}_{\mathbf{z}_{d,lkt}} - x_{d,kt}\mathbf{m}_{\mathbf{h}_{lk};\Psi_{2,lt}}, \\ \mathbf{C}_{\boldsymbol{v}} + \mathbf{C}_{\mathbf{z}_{d,lkt}} + |x_{d,kt}|^2\mathbf{C}_{\mathbf{h}_{lk};\Psi_{2,lt}}).\end{aligned} \tag{35}$$

VI. Message from variable nodes to factor nodes

The EP rule describing the message from variable nodes to factor nodes can be intuitively understood as two steps. At first, each variable node computes its belief (approximated posterior) by multiplying all the messages from its neighboring factor. It then sends back the extrinsic message (messages from all the neighboring factors except the receiver).

A. Message from $x_{d,kt}$ to $\Psi_{2,lt}$

The belief at variable node $x_{d,kt}$ is

$$b_{x_{d,kt}}(x_{d,kt}) \propto \mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt}) \prod_{l} \mu_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}). \tag{36}$$

We can derive the message from $x_{d,kt}$ to $\Psi_{2,lt}$

$$\mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt}) \propto \mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt}) \prod_{\bar{l}\neq l} \mu_{\Psi_{2,\bar{l}t};x_{d,kt}}(x_{d,kt}). \tag{37}$$

we shall see later that this message can be processed in a decentralized way.

B. Messages from $\mathbf{h}_{lk}$

The belief at variable node $\mathbf{h}_{lk}$ is

$$\begin{aligned}&b_{\mathbf{h}_{lk}}(\mathbf{h}_{lk}) = \mu_{\Psi_{4,lk};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \\ &\quad \cdot \prod_{t_1} \mu_{\Psi_{2,lt_1};\mathbf{h}_{lk}}(\mathbf{h}_{lk}) \prod_{t_2} \mu_{\Psi_{3,lt_2};\mathbf{h}_{lk}}(\mathbf{h}_{lk})\end{aligned} \tag{38}$$

The message from $\mathbf{h}_{lk}$ to $\Psi_{2,lt}$, $\Psi_{3,lt}$ and $\Psi_{4,lk}$ can then be directly obtained by using Gaussian reproduction lemma. We will only give here the message from $\mathbf{h}_{lk}$ to a general factor node Ψ, i.e., $\mu_{\mathbf{h}_{lk};\Psi}$ as an example,

$$\mu_{\mathbf{h}_{lk};\Psi}(\mathbf{h}_{lk}) = \mathcal{CN}(\mathbf{h}_{lk}|\mathbf{m}_{\mathbf{h}_{lk};\Psi}, \mathbf{C}_{\mathbf{h}_{lk};\Psi}), \tag{39}$$

where

$$\begin{aligned}\mathbf{C}_{\mathbf{h}_{lk};\Psi} &= \left(\sum_{\overline{\Psi}\in N(\mathbf{h}_{lk})/\{\Psi\}} \mathbf{C}_{\overline{\Psi};\mathbf{h}_{lk}}^{-1}\right)^{-1} \\ \mathbf{m}_{\mathbf{h}_{lk};\Psi} &= \mathbf{C}_{\mathbf{h}_{lk};\Psi}\left(\sum_{\overline{\Psi}\in N(\mathbf{h}_{lk})/\{\Psi\}} \mathbf{C}_{\overline{\Psi};\mathbf{h}_{lk}}^{-1}\mathbf{m}_{\overline{\Psi};\mathbf{h}_{lk}}\right).\end{aligned} \tag{40}$$

VII. Decentralized Asynchronous Method

To obtain the belief of $x_{d,kt}$, we need to combine the message from all the AP. We consider the case where all the L AP are connected and the AP network has a tree structure. A decentralized message passing method based on the framework of consensus propagation [12] can be used. If we consider the logarithmic scale, the product operation (36) indeed becomes summation.

Define the message from AP l to AP l' to be

$$\nu_{l\to l'}(x_{d,kt}) = \mu_{\Psi_{2,lt};x_{d,kt}}(x_{d,kt}) \prod_{\overline{l'}\in N(l)/\{l'\}} \nu_{\overline{l'}\to l}(x_{d,kt}). \tag{41}$$

At convergence, the belief in (36) can be obtained by any AP l as

$$b_{x_{d,kt}}(x_{d,kt}) \propto \mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt})\nu_{l\to l'}(x_{d,kt})\nu_{l'\to l}(x_{d,kt}). \tag{42}$$

Therefore, we can update $\mu_{x_{d,kt};\Psi_{2,lt}}$ by

$$\mu_{x_{d,kt};\Psi_{2,lt}}(x_{d,kt}) = \mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt}) \prod_{l'\in N(l)} \nu_{l'\to l}(x_{d,kt}). \tag{43}$$

Algorithm 1 Decentralized EP

Require: All prior distributions and likelihoods
Initialize all Messages, esp. $\mu_{\Psi_{1,kt};x_{d,kt}}(x_{d,kt}) = p(x_{d,kt})$
while Stopping criteria not met **do**
  for $l \in 1, \ldots, L$ **do**
    [Pilot Based Channel Estimation]
    for $k = 1 : K$ **do**
      Compute $\mu_{\mathbf{h}_{lk};\Psi_{4,lk}}$ based on (39)
      Compute $\mu_{\Psi_{4,lk};\mathbf{h}_{lk}}$ based on (33)
      for $t_2 = 1 : P$ **do**
        Compute $\mu_{\mathbf{h}_{lk};\Psi_{3,lt_2}}$ based on (39)
        Compute $\mu_{\Psi_{3,lt_2};\mathbf{h}_{lk}}$ based on (33)
      end for
    end for
    [Signal Estimation]
    for $k = 1 : K$ **do**
      for $t_1 = 1 : T$ **do**
        [Message from Other APs]
        Compute $\mu_{x_{d,kt_1};\Psi_{2,lt_1}}$ based on (43)
        Compute $\mu_{\Psi_{2,lt_1};x_{d,kt_1}}$ based on (14)
        [Message to Other APs]
        for $l' \in N(l)$ **do**
          Compute $\nu_{l \to l'}$ based on (41)
        end for
        [Data Based Channel Estimation]
        Compute $\mu_{\mathbf{h}_{lk};\Psi_{2,lt_1}}$ based on (39)
        Compute $\mu_{\Psi_{2,lt_1};\mathbf{h}_{lk}}$ based on (33)
      end for
    end for
  end for
end while

Adding this consensus-style message passing method between the APs within the conventional EP algorithm stated in the previous section is equivalent to a non-consensus-style EP algorithm with a modified update order and some $\delta(x^l_{d,kt} - x^{l'}_{d,kt})$ [15] factor nodes. One possible ordering method is concluded in Algorithm 1.

VIII. Simulation Results

In our simulation, we adopt a setup similar to that described in [7]. Each AP is equipped with 2 antennas. Our simulation environment is a 400×400 square meter area, populated with 16 APs and 8 User Terminals (UTs). The APs are strategically positioned at coordinates for each (i, j) in $\{0, 1, 2, 3\}^2$, specifically at $(50 + 100 \times i, 50 + 100 \times j)$ meters. UTs are uniformly distributed across the area. The distance between each UT, denoted as k, and AP, denoted as l, is represented by d_{lk}. Channel covariances are modeled using diagonal matrices, defined as $\sigma^2_{lk}\mathbf{I}$, where σ^2_{lk} is calculated using the formula $-30 - 36.7\log_{10}(d_{lk})$, following the approach of [7]. We use a 4QAM constellation for the transmitted signals, with a power of 30 dBm, and assume a noise power of -96 dBm. The results are based on 25 distinct realizations and are depicted in Figure 1. In our genie-aided scenario, the data symbol $x_{d,kt}$ is presumed known during the update of $\mu_{\Psi_{2,lt};\mathbf{h}_{lk}}$.

IX. Conclusions

In this paper, we proposed an EP-based algorithm to perform bilinear detection. The algorithm also tries to decentralize the network by using message passing between the APs. CLT is used to avoid the introduction of auxiliary variables.

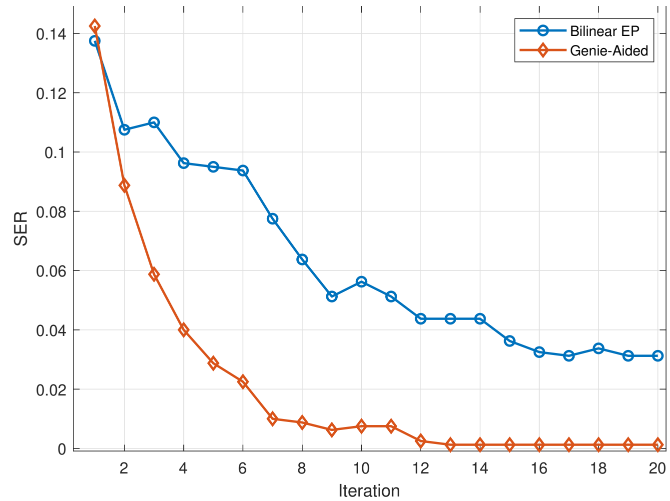

Fig. 1. The dynamic of proposed algorithm.

Acknowledgements EURECOM's research is partially supported by its industrial members: ORANGE, BMW, SAP, iABG, Norton LifeLock, and by the Franco-German projects CellFree6G.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
[2] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive mimo: A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 492–523, 2021.
[3] R. Gholami, L. Cottatellucci, and D. Slock, "Tackling pilot contamination in cell-free massive mimo by joint channel estimation and linear multi-user detection," in *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2021, pp. 2828–2833.
[4] T. P. Minka, "A family of algorithms for approximate bayesian inference," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.
[5] T. Heskes, M. Opper, W. Wiegerinck, O. Winther, and O. Zoeter, "Approximate inference techniques with expectation constraints," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 11, p. P11015, 2005.
[6] R. Gholami, L. Cottatellucci, and D. Slock, "Message passing for a bayesian semi-blind approach to cell-free massive mimo," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021, pp. 1237–1241.
[7] A. Karataev, C. Forsch, and L. Cottatellucci, "Bilinear Expectation Propagation for Distributed Semi-Blind Joint Channel Estimation and Data Detection in Cell-Free Massive MIMO," *IEEE Open Journal of Signal Processing*, 2024.
[8] Q. Zou and H. Yang, "A concise tutorial on approximate message passing," *arXiv preprint arXiv:2201.07487*, 2022.
[9] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—part i: Derivation," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5839–5853, 2014.
[10] Q. Zou, H. Zhang, C.-K. Wen, S. Jin, and R. Yu, "Concise derivation for generalized approximate message passing using expectation propagation," *IEEE Signal Processing Letters*, vol. 25, no. 12, pp. 1835–1839, 2018.
[11] K. Takeuchi, "Decentralized generalized approximate message-passing for tree-structured networks," *arXiv preprint arXiv:2311.01032*, 2023.
[12] C. C. Moallemi and B. Van Roy, "Consensus propagation," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4753–4766, 2006.
[13] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, and P. Schniter, "Multi-user detection based on expectation propagation for the non-coherent simo multiple access channel," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 6145–6161, 2020.
[14] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, p. 510, 2008.
[15] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6664–6684, 2019.