Joint Active User Detection and Channel Estimation via Bayesian Learning Approaches in MTC Communications

Xiaoxu Zhang, *Member, IEEE*, Fabrice Labeau, *Senior Member, IEEE*,
Li Hao, *Member, IEEE*, and Jiaqi Liu

***Abstract*—To support the massive machine-type communications (mMTC) scenario for Internet of Things (IoTs) applications featured by large-scale device connectivity and low device activity, grant-free non-orthogonal multiple access (GF-NOMA) and compressive sensing (CS)-based multi-user detection methods (MUD) are developed. In this paper, we develop two Bayesian CS-based methods, i.e., sparse Bayesian Learning (SBL) and fast inverse-free sparse Bayesian Learning (FI-SBL), for joint MUD and channel estimation (CE) in GF-NOMA with Low-Activity Code Division Multiple Access (LA-CDMA) as the multiple access technology. SBL is investigated for robust MUD and CE by utilizing the parameterized Gaussian prior information. Then to resolve the high computational complexity of SBL, FI-SBL is proposed, which replaces matrix reversion operations with relaxed evidence lower bound. Simulation results show that the two proposed algorithms outperform the traditional methods, and FI-SBL reduces the computational complexity significantly.**

***Index Terms*—Machine-Type Communications, Spares Channel Estimation, Sparse Bayesian Learning, Fast Inverse-free Sparse Bayesian Learning.**

I. INTRODUCTION

Massive machine-type communications (mMTC) is an important scenario to support the proliferation of Internet of things (IoT) applications [1]. In mMTC, although a huge number of devices connect to a base station (BS), very few of them are active with uplink short package transmission at a time. If mMTC directly adopts the uplink procedure of the long time evolution (LTE) system, where orthogonal radio resource multiplexing schemes as well as handshaking between BS and user devices are adopted to avoid collision, then the large signaling overhead and high end-to-end latency are unacceptable. In this regard, non-orthogonal multiple access with grant-free transmission (GF-NOMA) is considered as is a promising technology to meet the requirements [2].

GF-NOMA uses non-orthogonal preamble sequences for user identification and channel estimation (CE), which enables the overload use of preamble sequences, i.e. the number of supported connections is larger than the preamble length, and eliminates preamble contention by preassigning a unique preamble sequence to each user device. Inspired by the fact of massive connectivity and sporadic activity in mMTC, many efficient compressive sensing (CS)-based sparse reconstruction methods have been adopted for multi-user detection (MUD) and CE [3]. Greedy sparse reconstruction algorithms, such as compressive sampling matching pursuit (CoSaMP) [4], orthogonal matching pursuit (OMP) [5], and subspace pursuit (SP) [6], have been adopted for joint MUD and CE in GF-NOMA systems. As IoT applications highly require for reliability, Zhang *et al.* [7] developed a sparse Bayesian learning (SBL)-based method to improve the robustness and accuracy of MUD and CE by exploiting the prior information on the statistic distribution of noise and sparsity [8]. However, a key challenge of SBL is its high computational complexity.

In this paper, we realize that the computational complexity of SBL mainly comes from the matrix inversion operation. Then we propose the FI-SBL method, which significantly reduces the computational complexity of SBL by replacing the matrix inversion operation with a relaxed evidence lower bound (relaxed-ELBO) operation [9], [10]. The novelty and contributions of this paper are summarized as follows:

- Two Bayesian learning strategies, i.e., SBL and fast inverse-free SBL (FI-SBL), are developed for joint MUD and CE in mMTC;
- SBL utilizes the prior information of the channel vector and estimates the hyperparameters via expectation maximization (EM) iterative method [8];
- FI-SBL is proposed to reduce the computational complexity of SBL;
- Simulation results show that both SBL and FI-SBL outperform the traditional methods and that FI-SBL has significantly reduced computational complexity.

This paper is organized as follows. The system model of uplink GF-NOMA with low-activity code division multiple access (LA-CDMA) is presented in Section II. The proposed SBL and FI-SBL algorithms are respectively derived in Section III and Section IV. The simulations results are presented

This work was supported in part by the National Science Foundation of China under Grants 62001399, and in part by the Fundamental Research Funds for the Central Universities under Grant 2682020CX84. F. Labeau's work was partially funded by the Natural Sciences and Engineering Council of Canada, under the Discovery Grants program (RGPIN 2017-04728). L. Hao's work was partially funded by National Key R&D Program of China under Grant 2018YFB1801104. (Corresponding author: Xiaoxu Zhang.)

X. Zhang and L. Hao are with Southwest Jiaotong University, Chengdu, 611756, R. P. China (e-mail: xiaoxuzhang@swjtu.edu.cn, lhao@swjtu.edu.cn).

F. Labeau is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: fabrice.labeau@mcgill.ca).

J. Liu is with the Research and Development Department of Communication, Navigation, and Surveillance Equipments, China Electronics Technology Group Corporation Avionics Company Ltd., Chengdu 611731, R. P. China (e-mail: jq.liu@vip.163.com).

in Section V and the conclusions are drawn in Section VI.

II. Problem Formulation

We consider an uplink grant-free NOMA system with low activity code domain multiple access (LA-CDMA) [3] as the multiple access technology, in which N devices transmit signals to a base station. When a device is active, it transmits a signal in a certain frame; otherwise, it does not transmit. The spreading factor is M. Here we consider a flat fading channel. In the uplink LA-CDMA systems, we assume pilot symbol $b_{p,n}$ for device n and Binary Phase Shift Keying (BPSK) modulation with symbol alphabet ($\mathcal{B} = \{-1, +1\}$). The spreading sequence for device n is $\mathbf{s}_n \in \{-1/\sqrt{M}, +1/\sqrt{M}\}^M$. Thus the transmitted pilot signal for device n should be $b_{p,n}\mathbf{s}_n$. Due to the sporadic nature of the transmissions, only a small subset $D \subset \{1, 2, ..., N\}$ of users is active, and the cardinality is $K = |D| = p_a \cdot N$ on average, in which p_a is the percentage of active devices. In other words, when device n corresponding to channel coefficient h_n is active, $n \in D$; otherwise, $n \notin D$.

Without loss of generality, during pilot training, the received signal $\mathbf{y}$ at the base station can be formulated as

$$\begin{aligned}
\mathbf{y} &= \mathbf{s} \cdot \mathrm{diag}(\mathbf{b}_p) \cdot \mathbf{h} + \mathbf{w} \\
&= (\mathbf{s}_1, ..., \mathbf{s}_N) \begin{pmatrix} b_{p,1} & & \\ & \ddots & \\ & & b_{p,N} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_N \end{pmatrix} + \mathbf{w} \\
&= \mathbf{A}\mathbf{h} + \mathbf{w},
\end{aligned} \tag{1}$$

in which $\mathbf{A} = \mathbf{s}\cdot\mathrm{diag}(\mathbf{b}_p)$ is containing the spreading sequence and pilot signal information, $\mathbf{w}$ is a complex Gaussian noise and obeys $\mathbf{w} \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$. Note that in such transmission, if most devices are inactive, namely, most of the elements in $\mathbf{h}$ are zero, then the channel vector is considered as sparse. In the following, our goal for the base station is to do channel estimation based on the system model (1).

III. SBL for Channel Estimation

SBL was originally proposed as a machine learning algorithm by Tipping, and subsequently introduced into the field of CS. As a Bayesian algorithm, the SBL algorithm uses a parameterized Gaussian distribution as the prior distribution of the solution. The introduced hyperparameters are iteratively estimated via the EM strategy. In this section, we address the sparse channel estimation problem via SBL based on the EM iterative method [8]. The SBL method mainly exploits the sparsity of the channel vector $\mathbf{h}$, and has better recovery performance than traditional methods.

During pilot training, the coefficients h_n of the channel vector are independent of each other. When the activity factor is very small, most coefficients in the channel vector are zero. In order to encourage the sparsity of channel vector, assume that $\mathbf{h}$ follows zero-mean Gaussian prior distribution

$$\begin{aligned}
p(\mathbf{h}; \boldsymbol{\alpha}) &= \prod_{n=1}^{N} \mathcal{CN}(h_n | 0, \alpha_n^{-1}) \\
&= (2\pi)^{-N} |\mathbf{\Lambda}| \exp(-\mathbf{h}^H \mathbf{\Lambda} \mathbf{h}),
\end{aligned} \tag{2}$$

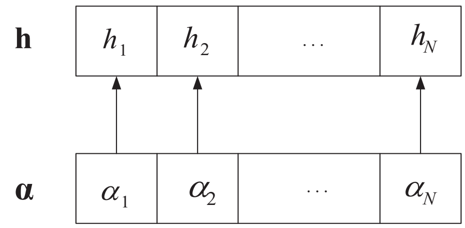


Fig. 1. The hyperparameter $\boldsymbol{\alpha}$ governs the estimated channel vector $\mathbf{h}$ in SBL algorithm.

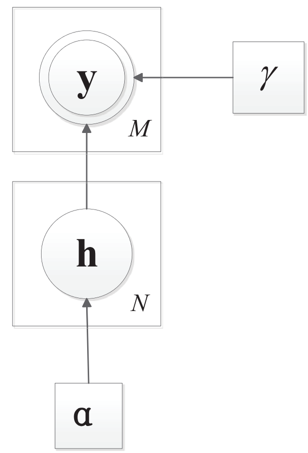


Fig. 2. The graphical model for SBL.

in which $\mathbf{\Lambda} = \mathrm{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$ is a diagonal matrix, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_N)^T$ is non-negative hyperparameter governing the estimated channel vector $\mathbf{h} = (h_1, h_2, ..., h_N)^T$, , as shown in Fig. 1. Each element in the estimated channel vector $\mathbf{h}$ is governed by the corresponding element in hyperparameter $\boldsymbol{\alpha}$. Suppose α_n approaches positive infinity, then the corresponding channel coefficient h_n becomes zero.

In SBL framework, we set γ as the inverse of the noise variance, i.e., $\gamma \triangleq 1/\sigma^2$, then the Gaussian noise vector is assumed to be zero-mean, and obeys $p(\mathbf{w}) = \mathcal{CN}(\mathbf{w}|0, \gamma^{-1}\mathbf{I})$. Thus the likelihood of received signal $\mathbf{y}$ is

$$p(\mathbf{y}; \mathbf{h}, \gamma) = \mathcal{CN}(\mathbf{y}|\mathbf{A}\mathbf{h}, \gamma^{-1}\mathbf{I}). \tag{3}$$

To show inner relationships of the parameters, we introduce the graphical model for SBL, as shown in Fig. 2. The SBL method considers $\boldsymbol{\alpha}$ and γ as hyperparameters of the system, $\mathbf{h}$ as hidden random variables. Here observed variable $\mathbf{y}$ is controlled by hidden variable $\mathbf{h}$ and hyperparameter γ, and the hidden variable $\mathbf{h}$ is controlled by hyperparameter $\boldsymbol{\alpha}$.

Based on the above assumptions, we calculate the posterior distribution of the channel vector $\mathbf{h}$, which is

$$\begin{aligned}
p(\mathbf{h}|\mathbf{y}; \boldsymbol{\alpha}, \gamma) &= \frac{p(\mathbf{y}|\mathbf{h}; \gamma) p(\mathbf{h}; \boldsymbol{\alpha})}{\int p(\mathbf{y}|\mathbf{h}; \gamma) p(\mathbf{h}; \boldsymbol{\alpha}) d\mathbf{h}} \\
&= \mathcal{CN}(\mathbf{h}|\boldsymbol{\mu}, \mathbf{\Sigma}),
\end{aligned} \tag{4}$$

with

$$\begin{aligned}
\mathbf{h}^{\mathrm{SBL}} &= \boldsymbol{\mu} = \gamma \mathbf{\Sigma} \mathbf{A}^H \mathbf{y} \\
\mathbf{\Sigma} &= (\gamma \mathbf{A}^H \mathbf{A} + \mathbf{\Lambda})^{-1}.
\end{aligned} \tag{5}$$

The estimate of channel vector $\mathbf{h}$ is the posterior mean value, our problem therefore becomes to updating the hyperparameters $\boldsymbol{\alpha}$ and γ. The EM and variational expectation maximization (V-EM) algorithms are iterative optimization strategies for maximum likelihood estimation, which are applicable for addressing statistic problems [8]. Each iteration of the calculation methods is divided into two steps: the expectation step and the maximization step. Here we use the EM strategy. EM treats the channel vector $\mathbf{h}$ as hidden variable, and maximizes the evidence lower bound (ELBO) of posterior probability. Due to the space limit, we give the results directly:

$$\begin{aligned} {\alpha_i}^{(t+1)} &= \tfrac{1}{\mu_i^2+\Sigma_{ii}} \\ \gamma^{(t+1)} &= \tfrac{M}{\left\|\mathbf{y}-\mathbf{A}\boldsymbol{\mu}^{(t)}\right\|^2+\text{tr}[\mathbf{A}\boldsymbol{\Sigma}^{(t)}\mathbf{A}^H]}, \end{aligned} \tag{6}$$

where μ_i and Σ_{ii} respectively stand for the ith entry of $\boldsymbol{\mu}^{(t)}$ and the ith diagonal element of convariance matrix $\boldsymbol{\Sigma}^{(t)}$. As a whole, the SBL approach involves updating the mean and covariance values of the channel vector $\mathbf{h}$ using (5) and (6) until convergence, the detail steps are given in Algorithm 1.

Algorithm 1 **Sparse Bayesian Learning**

Input: $\mathbf{y}$, $\mathbf{A}$, N and M

1. Calculate the posterior distribution over $\mathbf{h}$ according to (5);
2. Update the hyperparameters according to (6);
3. Continue the iterations until $\left\|\mu^{(t)} - \mu^{(t-1)}\right\|_2 \leq 10^{-6}$.

Output: $\mathbf{h}$, $\boldsymbol{\alpha}$ and γ

Using such a SBL method could be encourage sparsity and more accurate than many other compressive sensing methods. It's worth mentioning that the Bayesian method obtains a better recovery performance when the the sparsity of channel is improved, and the Bayesian compressed sensing method is an important issue worth studying.

IV. FI-SBL for Channel Estimation

SBL is one of the most popular methods in compressive sensing, and has better recovery performance than many other algorithms. Nevertheless, the main drawback is that we require computing the inverse operation of $N \times N$ covariance matrix when updating the channel vector $\mathbf{h}$. Then computational complexity is $\mathcal{O}(MN^2)$, which makes the SBL method difficult to implement in practice. To address the issue, in this section, we introduce an efficient algorithm called FI-SBL to reduce to computational complexity [10]. In a nutshell, the method tries to maximize a relaxed evidence lower bound (relaxed-ELBO), then uses V-EM to update the corresponding parameters.

The assumptions for noise variance and the channel coefficients prior are same as the SBL part. The graphical model for FI-SBL is shown in Fig. 3. In detail, the FI-SBL method considers $\mathbf{y}$ as the observed variable, $\mathbf{h}$ as the hidden random variable, $\{\boldsymbol{\alpha}, \gamma\}$ as the hyperparameters. The introduced hierarchical graphic model shows that all the parameters are statistically dependent.

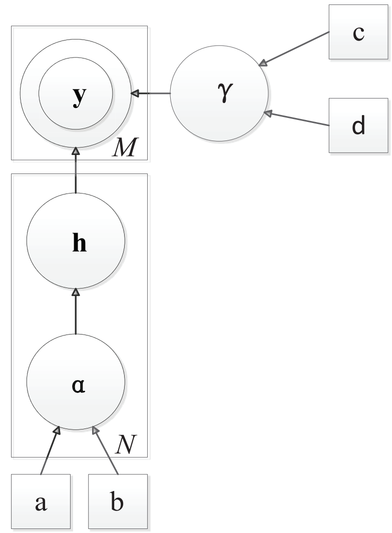


Fig. 3. The graphical model for FI-SBL.

Motivated by [8], we place Gamma distributions over $\boldsymbol{\alpha}$, and the hyperpriors can be expressed as

$$p(\boldsymbol{\alpha}) = \prod_{n=1}^{N} \text{Gamma}(\alpha_n|a,b) = \prod_{n=1}^{N} \Gamma(a)^{-1} b^a \alpha_n^{a-1} e^{-b\alpha_n}, \tag{7}$$

where $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$ is called Gamma function. Furthermore, we assume that the prior of the noise follows a Gamma function

$$p(\gamma; c, d) = \text{Gamma}(\gamma|c,d). \tag{8}$$

Traditionally, in order to make the gamma prior without information, we always assign very small values to the parameters, i.e., $a = b = c = d = 10^{-6}$.

1. Relaxed-ELBO:

As we mentioned above, the method tries to maximize relaxed-ELBO. Here the ELBO for this model can be expressed as

$$\begin{aligned} L(q) &= \int q(\boldsymbol{\Theta}) \ln\left(\frac{p(\mathbf{y}, \boldsymbol{\Theta})}{q(\boldsymbol{\Theta})}\right) d\boldsymbol{\Theta} \\ &= \int q(\boldsymbol{\Theta}) \ln\left(\frac{p(\mathbf{y}|\mathbf{h},\gamma)p(\mathbf{h}|\boldsymbol{\alpha})p(\boldsymbol{\alpha})p(\gamma)}{q(\boldsymbol{\Theta})}\right) d\boldsymbol{\Theta}, \end{aligned} \tag{9}$$

in which $q(\boldsymbol{\Theta})$ is a probability density function, and one lower bound of $p(\mathbf{y}|\mathbf{h},\gamma)$ can be obtained by

$$\begin{aligned} p(\mathbf{y}|\mathbf{h};\gamma) &= \mathcal{CN}(\mathbf{A}\mathbf{h}, \gamma^{-1}\mathbf{I}) \\ &= (2\pi)^{-M}\gamma^M \exp(-\gamma \left\|\mathbf{y} - \mathbf{A}\mathbf{h}\right\|_2^2) \\ &\geq (2\pi)^{-M}\gamma^M \exp(-\gamma g(\mathbf{h},\mathbf{z})) \\ &\triangleq F(\mathbf{y},\mathbf{h},\mathbf{z},\gamma), \end{aligned} \tag{10}$$

where $g(\mathbf{h},\mathbf{z})$ is defined as

$$\begin{aligned} g(\mathbf{h},\mathbf{z}) \triangleq{}& \left\|\mathbf{y} - \mathbf{A}\mathbf{z}\right\|_2^2 + 2(\mathbf{h}-\mathbf{z})^H\mathbf{A}^H(\mathbf{A}\mathbf{z}-\mathbf{y}) \\ &+ \frac{T}{2}\left\|\mathbf{h}-\mathbf{z}\right\|_2^2, \end{aligned} \tag{11}$$

in which T is a constant. The relaxed operation in (10) is based on an operation for smooth function [9], which is widely used in fast first-order algorithm.

Based on (9) and (11), we have

$$L(q) \geq \tilde{L}(q, \mathbf{z}) = \int q(\mathbf{\Theta}) \ln(\frac{G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z})}{q(\mathbf{\Theta})}) d\mathbf{\Theta}, \tag{12}$$

where $G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z})$ is defined as

$$G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z}) \triangleq F(\mathbf{y}, \mathbf{h}, \mathbf{z}, \gamma) p(\mathbf{h}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\gamma). \tag{13}$$

Eventually, we get the relaxed-ELBO

$$\begin{aligned} \tilde{L}(q, \mathbf{\Theta}) &= \int q(\mathbf{\Theta}) \ln(\frac{G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z})}{q(\mathbf{\Theta})}) d\mathbf{\Theta} \\ &= \int q(\mathbf{\Theta}) \ln(\frac{G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z}) h(\mathbf{z})}{q(\mathbf{\Theta}) h(\mathbf{z})}) d\mathbf{\Theta} \\ &= \int q(\mathbf{\Theta}) \ln(\frac{G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z}) h(\mathbf{z})}{q(\mathbf{\Theta})}) d\mathbf{\Theta} - \ln h(\mathbf{z}), \end{aligned} \tag{14}$$

in which we define $\tilde{G}(\mathbf{y}, \mathbf{\Theta}, \mathbf{z}) \triangleq G(\mathbf{y}, \mathbf{\Theta}, \mathbf{z}) h(\mathbf{z})$, and $h(\mathbf{z})$ is a normalizing term to ensure $\tilde{G}(\mathbf{y}, \mathbf{\Theta}, \mathbf{z})$ to be a rigorous distribution.

2. Parameters updating:

In the following, we update the hidden variables $\mathbf{\Theta} = \{\mathbf{h}, \boldsymbol{\alpha}, \gamma\}$ and $\mathbf{z}$ based on the V-EM method. Due to the space limit, we give the results directly:

1) Update of $q(\mathbf{h})$:

$$q(\mathbf{h}) = \mathcal{CN}(\mathbf{h}|\boldsymbol{\mu}, \mathbf{\Sigma}), \tag{15}$$

with

$$\begin{aligned} &\mathbf{h}^{\text{FI-SBL}} = \boldsymbol{\mu} = \langle\gamma\rangle \mathbf{\Sigma}(\mathbf{A}^H \mathbf{A} \mathbf{z} - \mathbf{A}^H \mathbf{y} - \tfrac{T}{2} \mathbf{z}) \\ &\mathbf{\Sigma} = (\mathbf{\Lambda} + \tfrac{T\langle\gamma\rangle}{2} \mathbf{I})^{-1}. \end{aligned} \tag{16}$$

2) Update of $q(\boldsymbol{\alpha})$:

$$q(\boldsymbol{\alpha}) = \prod_{n=1}^{N} \text{Gamma}(\alpha_n; \tilde{a}, \tilde{b}_n), \tag{17}$$

with

$$\begin{aligned} \tilde{a} &= a + 1 \\ \tilde{b}_n &= b + \left\langle b_n^2 \right\rangle. \end{aligned} \tag{18}$$

3) Update of $q(\gamma)$:

$$q(\gamma) = \text{Gamma}(\gamma; \tilde{c}, \tilde{d}), \tag{19}$$

with

$$\begin{aligned} \tilde{c} &= c + N \\ \tilde{d} &= d + \langle g(\mathbf{h}, \mathbf{z}) \rangle. \end{aligned} \tag{20}$$

4) Update of $\mathbf{z}$:

$$\mathbf{z} = \boldsymbol{\mu}. \tag{21}$$

In summary, the FI-SBL method involves updating the mean and covariance values of the channel vector $\mathbf{h}$ using (15), (17), (19) and (21) until convergence, the detail steps are given in Algorithm 2. Referring to (16), we need to compute the inverse operation of $N \times N$ covariance matrix when updating $q(\mathbf{h})$. However, the covariance matrix is a diagonal matrix, the proposed FI-SBL method does not need to calculate any inverse operation, thus the computational complexity is largely reduced.

Algorithm 2 **Fast Inverse-free Sparse Bayesian Learning**

Input: $\mathbf{y}$, $\mathbf{A}$, N and M

1. Calculate the posterior distribution over $\mathbf{h}$ according to (15);
2. Update the hyperparameter $\boldsymbol{\alpha}$ according to (17);
3. Update the hyperparameter γ according to (19);
4. Update the hyperparameter $\mathbf{z}$ according to (21);
5. Continue the iterations until $\left\| \mu^{(t)} - \mu^{(t-1)} \right\|_2 \leq 10^{-6}$.

Output: $\mathbf{h}$, $\boldsymbol{\alpha}$, γ and $\mathbf{z}$

V. SIMULATIONS

In this section, we verify the MUD performance of LS, MMSE, OMP [11], BP [12], SBL, and FI-SBL for grant-free NOMA systems with LA-CDMA. The computer simulation uses BPSK constellations in MTC communications. We consider flat fading channels. For FI-SBL, we set $a = b = c = d = 10^{-6}$ throughout our experiments. Once the normalize square error value is smaller than 10^{-6}, the iterative loop will stop. The metric to evaluate the reconstruction ability is mean square error (MSE), which is defined as

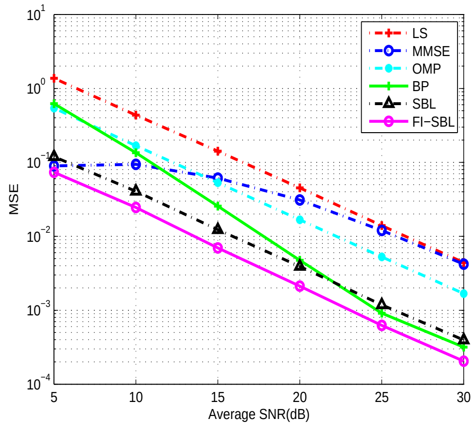
$$\text{MSE} = \frac{\overline{\left\| \mathbf{h} - \hat{\mathbf{h}} \right\|_2^2}}{N}, \tag{22}$$

where $\mathbf{h}$ and $\hat{\mathbf{h}}$ are the true and estimated channels, $\overline{a}$ is an average value over 1000 Monte Carlo trials, and N is the number of devices.

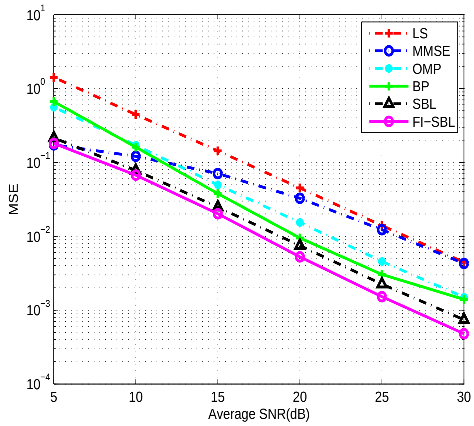
We evaluate the SNR versus the MSE of LS, MMSE, OMP, BP, SBL, and FI-SBL. The CDMA spreading factor and the number of devices are assigned to be $M = 128$ and $N = 100$, respectively. We select different user activity factor p_a, $p_a = 0.1$ and $p_a = 0.3$, the results are depicted in Fig. 4. From Fig. 4, we find that the SBL and FI-SBL methods outperform the LS, MMSE, OMP, and BP detectors. This is because LS and MMSE do not utilize any sparsity property. Due to the unknown sparsity, the greedy OMP algorithm has some performance loss. The BP method converts the optimization problem into minimal 1-norm, and we associate the relaxed 1-norm problem with the help of CVX tool box [13], which involves more effort. Meanwhile, it is noticed that the SBL method has good performance. The reason is that the SBL method exploits the prior information on the Gaussian distributions of the channel vector and is well suited for the condition when the activity factor is small. Fig. 4 depicts FI-SBL has the best recovery performance. Compared with SBL, the FI-SBL strategy has a much faster rate of convergence. The inverse of a matrix in SBL maybe inaccurate when the condition number of this matrix is large, and FI-SBL repalces the matrix inverse with relaxed-ELBO. Moreover, the performance gain of SBL and FI-SBL against other algorithms increases when SNR gets lower. This indicates that SBL and FI-SBL have better robustness against noise.

Table I records the running time of different algorithms. From the table, FI-SBL goes through much shorter running time, compared with SBL. This is because that the relaxed-ELBO involved by FI-SBL has much lower computational complexity than the matrix inverse involved by SBL.

Fig. 5 illustrates the impact of the number of users N on the MSE performance of LS, MMSE, OMP, BP, SBL, and



(a)



(b)

Fig. 4. MSE verse SNR of different algorithms when (a) $p_a = 0.1$; (b) $p_a = 0.3$.

TABLE I
RUNNING TIME OF DIFFERENT ALGORITHMS UNDER $M = 128$ AND $N = 100$ BASED ON 1000 MONTE CARLO AVERAGING

Algorithm	runtime for $p_a = 0.1$ (s)	runtime for $p_a = 0.3$ (s)
LS	0.5827	0.7008
MMSE	0.5108	0.5427
OMP	0.9771	1.0145
BP	77.9687	71.2468
SBL	**264.9341**	**239.5753**
FI-SBL	**1.4215**	**1.2214**

FI-SBL, with $p_a = 0.1$, SNR $= 20$ dB and $M = 128$. From the figure, we observe that SBL and FI-SBL outperform LS, MMSE, OMP, and BP. Moreover, with the increase of device number, the MSE of the SBL and FI-SBL algorithms are relatively steady. This indicates that SBL and FI-SBL have better robustness against the increase of device number.

VI. Conclusions

In this paper, we adopt LA-CDMA as the multiple access technology in grant-free NOMA systems for massive MTC communications, and develop two novel Bayesian inference algorithms, i.e., SBL and FI-SBL, for joint active user detection and channel estimation based on CS theory. SBL improves

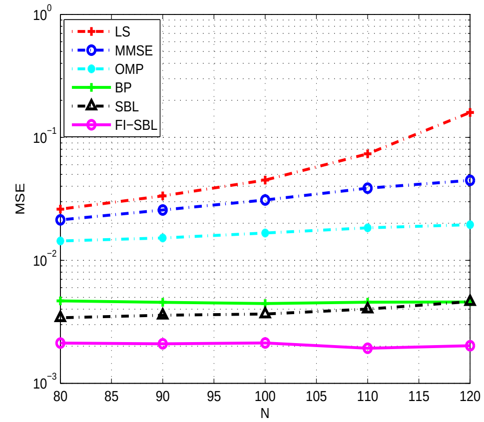


Fig. 5. MSE verse device number N of different algorithms.

the robustness of MUD and CE by exploiting the sparse prior information of the estimated channel vector. FI-SBL significantly reduces the computational complexity of SBL by replacing matrix reversion operations with relaxed evidence lower bound. Simulation results have presented that these two methods outperforms the classical strategies, and the FI-SBL method has superiority in running time.

References

[1] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy and G. Caire, “Massive access for future wireless communication systems,” *IEEE Wireless Communications*, doi: 10.1109/MWC.001.1900494.

[2] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, “Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the internet of things,” *IEEE Signal Processing Magazine*, vol. 35, pp. 88-99, Sep. 2018.

[3] X. Zhang, Y.-C. Liang and J. Fang, “Novel Bayesian inference algorithms for multiuser detection in M2M communications,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 7833-7848, 2017.

[4] B. Wang, L. Dai, Y. Yuan, and Z. Wang, “Compressive sensing based multi-user detection for uplink grant-free non-orthogonal multiple access,” in *IEEE Vehicular Technology Conference*, pp. 1-5, Sep. 2015.

[5] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA,” *IEEE Communications Letters*, vol. 20, pp. 2320-2323, Nov. 2016

[6] Y. Du, C. Cheng, B. Dong, Z. Chen, X. Wang, J. Fang, and S. Li, “Blocksparsity-based multiuser detection for uplink grant-free NOMA,” *IEEE Transactions on Wireless Communications*, vol. 17, pp. 7894-7909, Dec. 2018.

[7] Y. Zhang, Q. Guo, Z. Wang, J. Xi, and N. Wu, “Block sparse bayesian learning based joint user activity detection and channel estimation for grant-free NOMA systems,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9631-9640, 2018.

[8] D. G. Tzikas, A. C. Likas and N. P. Galatsanos, “The variational approximation for Bayesian inference,” *IEEE Signal Processing Magazine*, vol. 25, No. 6, pp. 131-146, Nov. 2008.

[9] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183-202, Mar. 2009.

[10] H. Duan, L. Yang, J. Fang and H. Li, “Fast inverse-free sparse Bayesian learning via relaxed evidence lower bound maximization,” *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 774-778, Mar. 2017.

[11] S. Adnan, Y. Fu, N. Junejo, Z.Chen, and H. Esmaiel, “Sparse detection with orthogonal matching pursuit in multiuser uplink quadrature spatial modulation MIMO system”, *IET Communications*, vol. 13, no. 20, pp. 3472-3478, 2019.

[12] S. Narayanan, S. K. Sahoo, and A. Makur, “Greedy pursuits assisted basis pursuit for compressive sensing”, *2015 23rd European Signal Processing Conference (EUSIPCO)*, Nice, France, 2015.

[13] S. Boyd, “Convex optimization,” *UK: Cambridge University*, Press, 2004.