SPARSE BLIND DEMIXING FOR LOW-LATENCY SIGNAL RECOVERY IN MASSIVE IOT CONNECTIVITY

Jialin Dong, Yuanming Shi*, and Zhi Ding†, Fellow, IEEE*

*School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China
†Dept. of ECE, University of California, Davis, California 95616, USA
E-mail: {dongjl, shiym}@shanghaitech.edu.cn, zding@ucdavis.edu

ABSTRACT

Internet-of-Things (IoT) networks are envisioned to typically include a massive number of devices with sporadic and low-latency uplink service needs. This paper presents a *blind demixing* approach to support the data recovery of multiple simultaneous and unscheduled device transmissions without *a priori* channel state information (CSI). The proposed joint receiver leverages the group sparse bilinear characteristics of the underlying problem that involves active device detection and data recovery. We exploit the manifold geometry of rank-one matrices in the lifted bilinear equation and apply smoothed ℓ_1/ℓ_2-norm to induce the group sparsity for active device detection. We further develop a smoothed Riemannian algorithm to solve the sparse blind demixing optimization problem. Numerical results demonstrate the algorithmic advantage and desirable performance of the proposed algorithm.

Index Terms— Blind demixing, device activity detection, CSI, data recovery, manifold optimization.

1. INTRODUCTION

The significant success of wireless technologies has been achieved towards connecting sensors, machines and robots for novel applications, thereby establishing the bedrock for the Internet-of-Things (IoT). The IoT connectivity will bring remarkable benefits to our lives, e.g., smart home, smart city, health care, transportation system, etc. A typical IoT connectivity involves a massive number of machine-type communication devices where hundreds of or even thousands of devices need to be connected. However, in the typical scenarios, the IoT device traffic patterns are *sporadic* where devices are designed to rest most of the time in order to save energy and are activated only when triggered by external events [1]. Furthermore, for the emerging Tactile Internet services, e.g., immersive virtual reality and cooperative automated driving, the additional haptic information needs to be further delivered in ultra-low latency communication [2]. It is thus crucial to support massive connectivity with low-latency communication to satisfy the diversified and tight traffic requirements in the IoT networks.

A growing body of literature has recently proposed various methods to deal with massive connectivity problem in IoT networks. Typically, the massive IoT connectivity consists of three components: active device detection, channel estimation, and their data decoding. The advanced compressive sensing techniques based on message passing [3] and integral geometry [4] has recently been developed for joint device activity detection and channel estimation, followed by data decoding supported by massive MIMO [5]. It turns out that this two-stage grant-free random access paradigm offers the potential to reduce access latency compared with the grant-based access [1]. However, this framework essential needs to estimate the channel information for data decoding, which incurs latency. To reduce the channel signaling overhead, blind demixing has become a promising way to detect data for multiple users without the prior knowledge of channel information [6, 7], thereby providing low-latency communication by assuming all the users are active.

In contrast, in this paper, we shall propose a novel *sparse blind demixing* approach to support massive low-latency access for IoT networks. This is achieved by enabling sporadic device activity detection, as well as their channel estimation and data decoding into a unified framework. Specifically, we simultaneously detect active devices and decode their data without the knowledge of channel information. However, it turns out that the sparse blind demixing is a joint group sparse and low-rank estimation problem with a sparse bilinear model. Specifically, the group sparsity serves the purpose of detecting active devices, while the low-rankness models the bilinear measurements via lifting [6]. Although the convex relaxation approaches based on the ℓ_1/ℓ_2-norm [8], the nuclear norm [9], or the combination of the ℓ_1/ℓ_2-norm and the nuclear norm [10], they are normally computationally prohibitive due to the high computation and storage cost. The nonconvex paradigm thus becomes a promising way to solve the nonconvex bilinear systems [11, 12] and the sparse bilinear systems [13, 14] by solving the estimation the problem in the natural vector space using the gradient methods [12, 11, 14] or the alternating minimization algorithm [13].

However, the presented sparse blind demixing brings

This work was supported in part by the National Nature Science Foundation of China under Grant 61601290 and in part by the Shanghai Sailing Program under Grant 16YF1407700.

unique challenge due to the group sparsity structure coupled with the sum of multiple bilinear measurements. To enjoy high computational efficiency and good performance for solving the sparse blind demixing problem, in this paper, we shall propose a smoothed Riemannian optimization algorithm on the product manifold. This is achieved by exploiting quotient manifold geometry of the product of rank-one matrices to address the multiple nonconvex bilinear constraints [7]. The smoothed ℓ_1/ℓ_2-norm is further developed to induce the group sparsity for active device detection, as well as fit the scope of the Riemannian optimization framework, as the objective function is normally required to be smooth for efficient manifold optimization algorithms design [15, 16].

2. PROBLEM FORMULATION

In this section, we present a sparse blind demixing for grant-free massive connectivity with low-latency communications.

2.1. System Model

We consider the Internet-of-Things (IoT) network consisting of one single-antenna base station (BS) and s single-antenna devices with sporadic traffic. Specifically, in each coherence block, only an unknown subset of devices are active, defined as $\mathcal{A} \subseteq \{1, 2, \cdots, s\}$ [1]. We assume that all the active users are synchronized by sending a beacon for the active devices from the BS.

Let $\boldsymbol{x}_k^\natural \in \mathbb{C}^N$ denote the original data symbols transmitted by the k-th user. Suppose that each device k is preassigned with a matrix $\boldsymbol{A}_k \in \mathbb{C}^{m \times N}$, followed by encoding the source signal $\boldsymbol{x}_k^\natural$ as $\boldsymbol{A}_k \boldsymbol{x}_k^\natural$ over the m time slots. Let $\boldsymbol{x}^{\mathsf{H}}$ denotes the conjugate transpose of the vector $\boldsymbol{x}$. Over m time slots, the received signals at the base station in the frequency domain are given as [6, 12]

$$y_j = \sum\nolimits_{k \in \mathcal{A}} \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{h}_k^\natural \boldsymbol{x}_k^{\natural\mathsf{H}} \boldsymbol{a}_{kj} + e_j,\ 1 \leq j \leq m, \tag{1}$$

where $\boldsymbol{h}_k^\natural \in \mathbb{C}^K$ are unknown channel vectors, $e_j \sim \mathcal{N}(0, \sigma^2) + i\mathcal{N}(0, \sigma^2)$ is the additive white complex Gaussian noise with $1/\sigma^2$ being the measurement of noise variance. Here, the first K columns of the unitary discrete Fourier transform (DFT) matrix $\boldsymbol{F} \in \mathbb{C}^{m \times m}$ with $\boldsymbol{F}\boldsymbol{F}^{\mathsf{H}} = \boldsymbol{I}_m$, form the known matrix $\boldsymbol{B} := [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m]^{\mathsf{H}} \in \mathbb{C}^{m \times K}$ with $\boldsymbol{b}_j \in \mathbb{C}^K$ for $1 \leq j \leq m$ [17, 12]. Furthermore, $\boldsymbol{a}_{kj} \in \mathbb{C}^N$ are known design vectors, which typically follow the i.i.d. complex Gaussian distribution [17, 12].

The massive IoT connectivity problem involves three components: sporadic device activity detection, channel estimation, and data decoding for the active devices. The advanced sparse signal estimation approach has recently been developed for joint activity detection and channel estimation [4, 3], followed by decoding their data [5]. By directly transmitting the metadata (e.g., preamble for active device detection and channel estimation) and data symbols to the BS without waiting for permission, this two-stage paradigm of grant-free random access has potential to reduce the access latency compared with the grant-based access scheme [1].

In contrast, in this paper, we propose a holistic view for grant-free massive IoT connectivity with low-latency communications. This is achieved by unifying the three components of active device detection $\mathcal{A}$, channel estimation $\boldsymbol{h}_k^\natural$'s and data decoding $\boldsymbol{x}_k^\natural$'s into a framework, yielding a novel *sparse blind demixing* approach. That is to say, we are able to offer the possibility for massive low-latency access without the knowledge of channel information via sparse blind demixing.

2.2. Spare Blind Demixing

Let $\boldsymbol{x} = [\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_s^T]^T \in \mathbb{C}^{Ns}$ with the index set $\mathcal{V} = \{1, 2, \ldots, Ns\}$ being the aggregative data vector to be detected. Consider the collection of groups $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_s\}$ with $\mathcal{G}_k = \{N(k-1)+1, \ldots, Nk\}$ and $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$ for $i \neq j$. Define the support of the data vector as $\mathrm{Supp}(\boldsymbol{x}) = \{i | x_i \neq 0, \forall i \in \mathcal{V}\}$. Therefore, the sparse blind demixing problem can be formulated as follows:

$$\begin{aligned} \mathscr{P}: \underset{\{\boldsymbol{x}_k\},\{\boldsymbol{h}_k\}}{\text{minimize}} \quad & \sum_{k=1}^{s} I(\mathrm{Supp}(\boldsymbol{x}) \cap \mathcal{G}_k \neq \emptyset) \\ \text{subject to} \quad & \sum_{j=1}^{m} \Big| \sum_{k=1}^{s} \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{h}_k \boldsymbol{x}_k^{\mathsf{H}} \boldsymbol{a}_{kj} - y_j \Big|^2 \leq \epsilon, \end{aligned}$$

where parameter $\epsilon > 0$ is assumed to be known *a priori*, $I(\mathrm{Supp}(\boldsymbol{x}) \cap \mathcal{G}_k \neq \emptyset)$ is an indicator function that takes value 1 if $\mathrm{Supp}(\boldsymbol{x}) \cap \mathcal{G}_k \neq \emptyset$ and 0 otherwise, $\boldsymbol{x}_k \in \mathbb{C}^N$ and $\boldsymbol{h}_k \in \mathbb{C}^K$ for $k = 1, \cdots, s$ are the optimization parameters. Let $\boldsymbol{x}^\star$ be a solution of problem $\mathscr{P}$, the set of active devices is given as $\mathcal{A}^\star = \{k : \mathrm{Supp}(\boldsymbol{x}) \cap \mathcal{G}_k \neq \emptyset\}$.

However, problem $\mathscr{P}$ turns out to be highly intractable due to the combinatorial objective function and the nonconvex bilinear constraint, for which we shall propose efficient algorithms with good performance.

2.3. Problem Analysis

A natural way to deal with the sparse blind demixing problem is to lift the bilinear model into the linear model with a low-rank matrix [18, 6], i.e., $\boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{h}_k \boldsymbol{x}_k^{\mathsf{H}} \boldsymbol{a}_{kj} = \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{W}_k \boldsymbol{a}_{kj}$ with $\boldsymbol{W}_k \in \mathbb{C}^{K \times N}$ and $\mathrm{rank}(\boldsymbol{W}_k) = 1, \forall k = 1, \ldots, s$. To address the algorithmic challenge of the lifted problem $\mathscr{P}$ due to the sparsity and low-rankness in matrices $\boldsymbol{W}_k$'s, the natural approach is to adopt the idea of convex relaxation, yielding the following semidefinite programming (SDP) problem:

$$\begin{aligned} \underset{\{\boldsymbol{W}_k\}}{\text{minimize}} \quad & \lambda_1 \sum\nolimits_{k=1}^{s} \|\boldsymbol{W}_k\|_* + \lambda_2 \sum\nolimits_{k=1}^{s} \|\boldsymbol{W}_k\|_F, \\ \text{subject to} \quad & \sum\nolimits_{j=1}^{m} \Big| \sum\nolimits_{k=1}^{s} \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{W}_k \boldsymbol{a}_{kj} - y_j \Big|^2 \leq \epsilon, \end{aligned} \tag{2}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the regularization parameters. Here, $\|\boldsymbol{W}_k\|_*$ is the nuclear norm of matrix $\boldsymbol{W}_k$, i.e., the sum of its singular values, to relax the rank-one constraints of matrices $\boldsymbol{W}_k$'s [18, 6]. With the group sparsity structure in the data vector $\boldsymbol{x}$, we have group sparsity structure in the lifting vector $\mathrm{vec}(\boldsymbol{W}) = [\mathrm{vec}(\boldsymbol{W}_1)^{\mathsf{H}}, \cdots, \mathrm{vec}(\boldsymbol{W}_s)^{\mathsf{H}}]^{\mathsf{H}} \in \mathbb{C}^{KNs}$, where the operation $\mathrm{vec}(\boldsymbol{M})$ denotes the vectorization of matrix $\boldsymbol{M}$. The ℓ_1/ℓ_2-norm serves the convex surrogate of inducing group sparsity in the vector $\mathrm{vec}(\boldsymbol{W})$, i.e., $\|\mathrm{vec}(\boldsymbol{W})\|_{1,2} = \sum_{k=1}^{s} \|\mathrm{vec}(\boldsymbol{W}_k)\|_2 = \sum_{k=1}^{s} \|\boldsymbol{W}_k\|_F$, where $\|\boldsymbol{M}\|_F$ represents the Frobenius norm of matrix $\boldsymbol{M}$ [19, 10]. Although the convex relaxation approach (2) provides a polynomial time complexity algorithm to solve the sparse blind demixing problem $\mathscr{P}$ in the lifted matrix space, it is computationally prohibitive for solving large-scale SDP problem (2).

Therefore, another line of literature has developed efficient nonconvex algorithms to solve the sparse and low-rank optimization problem [12, 13, 14] in the natural vector space. In particular, the provable Wirtinger flow algorithm was provided in [12] to solve the blind demixing problem without the additional sparsity structure. For the sparse blind deconvolution problem, i.e., the bilinear model with single user and the individual sparsity, the provable alternating minimization was developed in [13] and the simple descent algorithm was proven able to escape strict saddle points and then recover the ground truth signal [14]. However, the sparse blind demixing problem $\mathscr{P}$ brings unique challenges to adopt the nonconvex optimization paradigm due to the additional group sparsity structure and the sum of multiple bilinear measurements.

Instead, in this paper, we shall propose a smoothed Riemannian optimization approach to solve the sparse blind demixing problem to achieve the algorithmic advantages and admirable performance. This is achieved by exploiting the manifold geometry of the product of rank-one matrices to address the nonconvex bilinear constraints, and developing the smoothed ℓ_1/ℓ_2-norm to induce the group sparsity.

3. SMOOTHED RIEMANNIAN OPTIMIZATION ON PRODUCT MANIFOLD

In this section, we shall develop a smoothed Riemannian optimization algorithm to solve the sparse blind demixing problem $\mathscr{P}$ to improve computational efficiency and enjoy good performance.

3.1. Optimization on Product Manifold

To begin with, we reformulate problem $\mathscr{P}$ as a regularized optimization problem with fixed-rank constraints via lifting. Specifically, let $\mathbb{S}_+^{N+K}$ denote the set of $(N+K) \times (N+K)$ Hermitian positive semidefinite matrices. For $k = 1, \cdots, s$, $j = 1, \cdots, m$, we define $\boldsymbol{c}_j = [\boldsymbol{b}_j^{\mathsf{H}}, \boldsymbol{0}_N^{\mathsf{H}}]^{\mathsf{H}} \in \mathbb{C}^{N+K}$, $\boldsymbol{d}_{kj} = [\boldsymbol{0}_K^{\mathsf{H}}, \boldsymbol{a}_{kj}^{\mathsf{H}}]^{\mathsf{H}} \in \mathbb{C}^{N+K}$, hence it yields $\boldsymbol{c}_j^{\mathsf{H}} \boldsymbol{M}_k \boldsymbol{d}_{kj} = \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{h}_k \boldsymbol{x}_k^{\mathsf{H}} \boldsymbol{a}_{kj}$, where $\boldsymbol{M}_k = \boldsymbol{w}_k \boldsymbol{w}_k^{\mathsf{H}} \in \mathbb{S}_+^{N+K}$ is a Hermitian positive semidefinite matrix with $\boldsymbol{w}_k = [\boldsymbol{h}_k^{\mathsf{H}}, \boldsymbol{x}_k^{\mathsf{H}}]^{\mathsf{H}} \in \mathbb{C}^{N+K}$.

Therefore, problem $\mathscr{P}$ can be equivalently reformulated as the following optimization problem on the product of Hermitian positive semidefinite matrices:

$$\begin{aligned} \underset{\boldsymbol{M}}{\text{minimize}} \quad & \sum_{j=1}^{m} \Big| \sum_{k=1}^{s} \boldsymbol{c}_j^{\mathsf{H}} \boldsymbol{M}_k \boldsymbol{d}_{kj} - y_j \Big|^2 + \lambda f(\boldsymbol{M}) \\ \text{subject to} \quad & \mathrm{rank}(\boldsymbol{M}_k) = 1,\ k = 1, \cdots, s, \end{aligned} \tag{3}$$

where $\boldsymbol{M} = \{\boldsymbol{M}_k\}_{k=1}^{s}$ with $\boldsymbol{M}_k \in \mathbb{S}_+^{N+K}$, $f(\boldsymbol{M})$ is the function to induce the sparsity structure and $\lambda > 0$ is the regularization parameter. Note that each $\boldsymbol{M}_k$ lies on the manifold encoded by complex symmetric rank-one matrices, denoted as $\mathcal{M}_k$ [7]. We thus have $\boldsymbol{M} \in \mathcal{M}^s$, where $\mathcal{M}^s := \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_s$ represents the product of manifolds $\mathcal{M}_k$. We aim to develop computationally efficient Riemannian optimization algorithms on product manifold by exploiting the quotient manifold geometry of the product of complex symmetric rank-one matrices.

3.2. Smoothed Riemannian Optimization

To adopt the principle of Riemannian optimization for solving problem (3), smooth objective function is normally required [16, 20]. To achieve this goal, we shall present the smoothed ℓ_1/ℓ_2-norm to induce the group sparsity structure in vector $\mathrm{vec}(\boldsymbol{M}) = [\mathrm{vec}(\boldsymbol{M}_1)^{\mathsf{H}}, \cdots, \mathrm{vec}(\boldsymbol{M}_s)^{\mathsf{H}}]^{\mathsf{H}}$, i.e., $f_\epsilon(\boldsymbol{M}) = \sum_{k=1}^{s} \left(\|\boldsymbol{M}_k\|_F^2 + \epsilon^2 \right)^{1/2}$ with $\epsilon > 0$ as the smoothing parameter with a small value. Hence, the proposed smoothed Riemannian optimization approach over the product manifold $\mathcal{M}^s$ for solving the sparse blind demixing problem $\mathscr{P}$ is presented as

$$\underset{\boldsymbol{M} \in \mathcal{M}^s}{\text{minimize}} \sum_{j=1}^{m} \Big| \sum_{k=1}^{s} \boldsymbol{c}_j^{\mathsf{H}} \boldsymbol{M}_k \boldsymbol{d}_{kj} - y_j \Big|^2 + \lambda f_\epsilon(\boldsymbol{M}), \tag{4}$$

where the objective function is smooth and the constraint is a manifold.

Based on the geometry of the product manifold, the Riemannian optimization algorithms operated on the product manifold $\mathcal{M}^s$ can be elementwisely operated over the individual manifolds $\mathcal{M}_k$ [7]. Specifically, for each manifold $\mathcal{M}_k$, we search the update direction on the horizontal space of the manifold and then retract it on the manifold via retraction mapping. Here, searching the update direction can be achieved by the Riemannian optimization algorithms, e.g., conjugate gradient descent algorithm based on the Riemannian gradient [16].

4. SIMULATION RESULTS

In this section, we simulate our proposed smoothed Riemannian optimization algorithm for solving the sparse blind demixing problem $\mathscr{P}$ to demonstrate the algorithmic advantages and good performance.

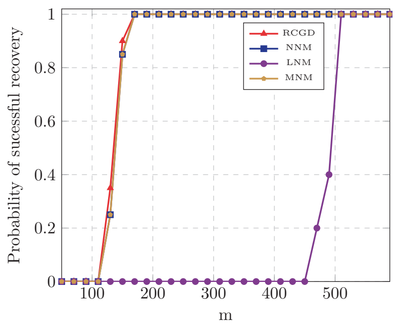

Fig. 1. Probability of successful recovery with different sample sizes m.

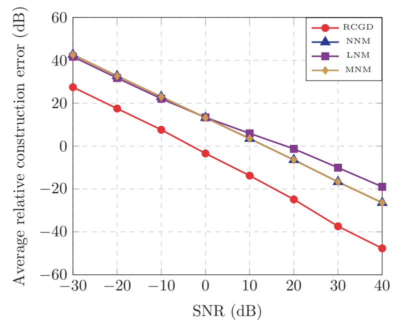

Fig. 2. Average relative construction error vs. SNR (dB).

4.1. Simulation Settings and Performance Metric

The simulation settings are given as: 1) Recall that $\mathcal{A}$ denotes the index set of active devices. For $k \notin \mathcal{A}$, we set $\boldsymbol{x}_k^\natural = \mathbf{0}_N$. For $k \in \mathcal{A}$, entries of $\boldsymbol{x}_k^\natural$ are signals generated via quadrature phase shift keying (QPSK) modulation, and $\boldsymbol{h}_k^\natural \in \mathbb{C}^K$ are drawn i.i.d from the standard complex Gaussian distribution. 2) Measurement vectors $\{\boldsymbol{b}_j\}, \{\boldsymbol{a}_{kj}\}$: We generate the normalized discrete Fourier transform (DFT) matrix $\boldsymbol{F} \in \mathbb{C}^{m\times m}$ and the complex Gaussian vectors $\boldsymbol{a}_{kj} \in \mathbb{C}^N$ to construct the measurement vectors according to Section 2.1. 3) Performance metric: The average relative construction error is adopted to evaluate the performance of the algorithms [12].

The four algorithms are compared: 1) **Riemannian conjugate-gradient descent algorithm (RCGD)**: The Riemannian conjugate-gradient descent algorithm (RCGD) for solving (4) is implemented using the manifold optimization toolbox *Manopt* [21]. The RCGD algorithm adopts the initialization strategy in [7] and stops when the norm of Riemannian gradient falls below 10^{-8} or the number of iterations exceeds 500. The regularization parameter λ is determined via cross validation. 2) **Nuclear norm minimization (NNM)**: The algorithm is implemented with the toolbox CVX [22] to solve the convex relaxation problem (2) with $\lambda_1 = 1$, $\lambda_2 = 0$. 3) **ℓ_1/ℓ_2-norm minimization (LMN)**: The setting of this algorithm is similar to NNM except for $\lambda_1 = 0$, $\lambda_2 = 1$. 4) **Mixed norm minimization (MNM)**: The setting of this algorithm is similar to NNM except that λ_1 and λ_2 are chosen via cross validation.

We investigate the empirical recovery performance of above four algorithms, i.e., RCGD, NNM, LMN and MNM. We set $N = K = 10$, $s = 10$, $|\mathcal{A}| = 3$ with sample size m varying from 100 to 560. For each setting, 30 independent trails are performed and the recovery is treated as a success if the relative construction error $\mathrm{err}(\boldsymbol{x}) \leq 10^{-2}$. Fig. 1 shows the probability of successful recovery for different sample sizes m. Based on the phase transitions results in Fig. 1, we can see that the proposed smoothed Riemannian optimization algorithm achieves much better performance compared with the LMN algorithm, i.e., exactly recover the ground truth signals with less samples. Although the RCGD, NNM and MNM algorithms achieve almost the same performance, the RCGD algorithm has the capability to scale to large problem sizes. This is based on the fact that Riemannian optimization algorithm RCGD is essentially based on matrix factorization, while the convex relaxation algorithms NNM and MNM need to solve SDP problems, which are known computationally prohibitive for large-scale problems.

We further study the average relative construction error of the four algorithms and explore the robustness of the proposed smoothed Riemannian optimization algorithm against additive noise. We assume the additive noise in the formulation (1) satisfies $e_j = \sigma \cdot \|\sum_{k\in\mathcal{S}} \boldsymbol{b}_j^{\mathsf{H}} \boldsymbol{h}_k^\natural \boldsymbol{x}_k^{\natural\mathsf{H}} \boldsymbol{a}_{kj}\|_2 \cdot \frac{\omega}{|\omega|}$, $1 \leq j \leq m$, where $\omega \in \mathbb{C}$ is drawn from the standard complex Gaussian distribution [12]. We compare the four algorithms for each level of signal-to-noise ratio (SNR) $1/\sigma$ in the setting of $m = 550$, $N = K = 10$, $s = 10$, $|\mathcal{A}| = 3$. For each setting, 20 independent trails are performed and the condition of successful recovery is the same with the one aforementioned. The average relative construction error in dB against the SNR is illustrated in Fig. 2. It shows that the proposed algorithm is robust to the noise and achieves better performance than other algorithms in the noisy case.

5. CONCLUSION

In this paper, we proposed a novel *sparse blind demixing* approach to recover the source signals generated by active devices without the prior knowledge of channel information, thereby providing massive low-latency IoT connectivity. To address the unique challenge of inducing the group sparsity for multiple source signals with bilinear measurements, we developed a smoothed Riemannian optimization algorithm to solve the blind demixing problem. This is achieved by exploiting the quotient manifold geometry of the product of complex rank-one matrices to address bilinear constraints, followed by the smoothed ℓ_1/ℓ_2-norm to induce group sparsity. The simulation results showed that our proposed method outperforms the existing methods in terms of computational efficiency and signal recovery performance.

6. REFERENCES

[1] Liang Liu, Erik G Larsson, Wei Yu, Petar Popovski, Cedomir Stefanovic, and Elisabeth de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.

[2] K. S. Kim, D. K. Kim, C. Chae, S. Choi, Y. Ko, J. Kim, Y. Lim, M. Yang, S. Kim, B. Lim, K. Lee, and K. L. Ryu, "Ultrareliable and low-latency communication techniques for tactile internet services," *Proc. IEEE*, to appear 2018.

[3] Liang Liu and Wei Yu, "Massive connectivity with massive MIMO-part I: Device activity detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2933–2946, Jun. 2018.

[4] Tao Jiang, Yuanming Shi, Jun Zhang, and Khaled B Letaief, "Joint activity detection and channel estimation for iot networks: Phase transition and computation-estimation tradeoff," *arXiv preprint arXiv:1810.00720*, 2018.

[5] Liang Liu and Wei Yu, "Massive connectivity with massive MIMO-part II: Achievable rate characterization," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2947–2959, Jun. 2017.

[6] Shuyang Ling and Thomas Strohmer, "Blind deconvolution meets blind demixing: Algorithms and performance bounds," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4497–4520, May 2017.

[7] Jialin Dong, Kai Yang, and Yuanming Shi, "Blind demixing for low-latency communication," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Jun. 2018.

[8] Axel Flinth, "Sparse blind deconvolution and demixing through $\ell_{1,2}$-minimization," *Adv. Comput. Math.*, vol. 44, no. 1, pp. 1–21, Feb. 2018.

[9] Ali Ahmed and Laurent Demanet, "Leveraging diversity and sparsity in blind deconvolution," *IEEE Trans. Inf. Theory*, vol. 64, no. 6, pp. 3975–4000, Jun. 2018.

[10] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C Eldar, and Babak Hassibi, "Simultaneously structured models with application to sparse and low-rank matrices.," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2886–2908, May. 2015.

[11] Cong Ma, Kaizheng Wang, Yuejie Chi, and Yuxin Chen, "Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval and matrix completion," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Jul. 2018, vol. 80, pp. 3345–3354.

[12] Jialin Dong and Yuanming Shi, "Nonconvex demixing from bilinear measurements," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5152–5166, Oct. 2018.

[13] Kiryung Lee, Yihong Wu, and Yoram Bresler, "Near-optimal compressed sensing of a class of sparse low-rank matrices via sparse power factorization," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1666–1698, Mar. 2018.

[14] Yuqian Zhang, Han-Wen Kuo, and John Wright, "Structured local optima in sparse blind deconvolution," *arXiv preprint arXiv:1806.00338*, 2018.

[15] Yuanming Shi, Jun Zhang, Wei Chen, and Khaled B Letaief, "Generalized sparse and low-rank optimization for ultra-dense networks," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 42–48, 2018.

[16] P-A Absil, Robert Mahony, and Rodolphe Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, Apr. 2009.

[17] Shuyang Ling and Thomas Strohmer, "Regularized gradient descent: A nonconvex recipe for fast joint blind deconvolution and demixing," *Inf. Inference: J. IMA*, Mar. 2018.

[18] Ali Ahmed, Benjamin Recht, and Justin Romberg, "Blind deconvolution using convex programming," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1711–1732, Mar. 2014.

[19] Martin J Wainwright, "Structured regularizers for high-dimensional problems: Statistical and computational issues," *Annu. Rev. Stat. Appl.*, vol. 1, pp. 233–253, Jan. 2014.

[20] Y. Shi, B. Mishra, and W. Chen, "Topological interference management with user admission control via riemannian optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7362–7375, Nov. 2017.

[21] Nicolas Boumal, Bamdev Mishra, P.-A. Absil, and Rodolphe Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *J. Mach. Learn. Res.*, vol. 15, pp. 1455–1459, Jan. 2014.

[22] Michael Grant, Stephen Boyd, and Yinyu Ye, "CVX: Matlab software for disciplined convex programming," 2008.