GENERALIZED MATRIX COMPLETION FOR LOW COMPLEXITY TRANSCEIVER PROCESSING IN CACHE-AIDED FOG-RAN VIA THE BURER-MONTEIRO APPROACH

Kai Yang, Yuanming Shi*, and Zhi Ding†*

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

ABSTRACT

The emerging concept of Dense Fog Radio Access Network (Fog-RAN) pushes computation and storage resources to network edges as an effective architecture for wirelessly storing, processing, and exchanging big data information. However, the corresponding traffic overhead for massive channel state information (CSI) acquisition poses a serious obstacle to the design objective of achieving low network latency. This paper studies a cache-aided dense Fog-RAN by proposing a topological content delivery approach to maximize the achievable degrees-of-freedom (DoF) based only on the network topologies and side message information. Specifically, we present a generalized low-rank matrix completion approach, for which, we propose a Riemannian trust-region algorithm based on Burer-Monteiro factorization to solve a fixed rank subproblem in *complex* field by semidefinite lifting. Our results demonstrate the performance benefit of the proposed algorithm over known methods.

Index Terms— Fog-RAN, mobile edge caching and computing, low rank matrix optimization, Burer-Monteiro approach, Riemannian optimization.

1. INTRODUCTION

The widespread deployment and permeation of high speed wireless networks continue to stimulate various innovative wireless applications in the modern society. These cutting edge technological advances such as internet-of-things (IoT), tele-medicine, cyber-physical systems, and virtual reality exert tremendous pressure on the computation and communication capacities of wireless systems [1]. By pushing computational resources and storage units into network edges, fog radio access network (Fog-RAN) [2] has emerged as a strong candidate architecture capable of significantly enhancing the network computation capability through computation offloading [3], wireless distributed computing [4], as well as network latency reduction through caching [5] during off-peak time.

One recognized drawback in dense Fog-RANs is the high overhead needed for acquiring global channel state information (CSI) [6]. Thus, our work in this paper focuses on the CSI overhead reduction for massive content delivery in the cache-enabled dense Fog-RANs. Without relying on the assumption of global CSI in many existing studies of caching networks [7, 8], we turn to the topological interference management [9] as a promising approach that relies only on the more efficient network connectivity information. We note that some encouraging results on topological caching have already appeared in [10] and [11] for transmitter caching and receiver caching, respectively. In this paper, we jointly consider both transmitter and receiver caching in topological interference management for achieving greater benefits.

In order to design efficient content delivery strategies, we shall present a novel framework for modeling generalized matrix completion in the complex field to maximize the achievable degrees-of-freedom (DoF). Despite the non-convexity of the low-rank optimization problems, a growing number of recent works have focused on convex approximated algorithms. Specifically, nuclear norm is a well-known convex surrogate for the non-convex rank function, though it does not fit the requirement of our problem solution since it always yields full-rank solutions [10, 11]. Another popular direction tackles non-convex algorithms for low rank problems via matrix factorization, followed by the alternating minimization [10] and gradient descent methods [12]. However, this alternating minimization approach is harder to scale to large problem sizes whereas their first-order algorithms have slow convergence and are sensitive to initial points.

In contrast, Riemannian optimization [13] approach has the capability of exploiting the non-uniqueness of low rank matrix factorization characterized by quotient manifold geometries. Furthermore, Riemannian trust region algorithm harnesses the second-order information on Riemannian manifold, thereby converging to an *approximate local minimum* [14] from any initial points. Thus far, however, no Riemannian algorithms have been developed for the general non-square complex fixed-rank smooth optimization problems. In this work, we propose to lift the original problem into the low-rank positive semidefinite (PSD) matrix optimization problem. We then apply the Burer-Monteiro factorization approach to recover the semidefinite factor matrix. This approach has recently been successfully exploited in the multicast beamforming, community detection, phase retrieval [15] and the low-rank matrix completion/recovery problems [16]. Through simulation tests, our proposed approach demonstrates strong transmitter cooperation gains over existing solutions to the topological interference management problem.

2. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we examine the mechanism of caching in K-user interference channel only based on network topology information and propose a general matrix completion formulation with cache-enabled transmitters and receivers/users.

2.1. System Model

In Fog-RAN, caches at both transmitters and receivers can be used to manage interferences [8] thereby reducing network latency. In particular, we consider a file library of N messages $W_1, W_2, \cdots, W_N$ where each message has entropy of F digits so that it carries one degree-of-freedom (DoF). Consider K pairs of transmitters and receivers, each equipped with a single antenna for communicating through a K-user interference channel. We focus on the one-shot content delivery problem with linear precoding schemes, where each receiver requests exactly one message from the library.

Let side information $\mathcal{S}_k$ and $\mathcal{W}_k$ denote the indexed sets of messages which are available at transmitter k and receiver k, respectively. We also let $\mathcal{V}_k$ denote the network topology information such that the channel coefficient from transmitter j to receiver k is non-zero $h_{kj} \neq 0$ for $j \in \mathcal{V}_k$ and $h_{kj} = 0$ otherwise. Here $\mathcal{S}_k, \mathcal{W}_k, \mathcal{V}_k$ are all subsets of $[N] = \{1, 2, \cdots, N\}$. An illustrative example of 4-user case is shown in Fig. 1. Let $\boldsymbol{v}_{ji} \in \mathbb{C}^r$ be the precoding vector at transmitter j for message i with r channel uses. Then the signal transmitted by the j-th transmitter will be $\boldsymbol{x}_j = \sum_{i \in \mathcal{S}_j} \boldsymbol{v}_{ji} s_i$ where $s_i \in \mathbb{C}$. We denote the vector of demands $\boldsymbol{d} \in [N]^K$ of K receivers as $\boldsymbol{d} = [d_1, \cdots, d_K]^{\mathsf{T}}$ and $d_k \notin \mathcal{W}_k$. So the received signal at receiver k is given by the $r \times 1$ received vector of

$$\boldsymbol{y}_k = \sum_{j \in \mathcal{V}_k, d_k \in \mathcal{S}_j} h_{kj} \boldsymbol{v}_{jd_k} s_{d_k} + \sum_{i \neq d_k} \sum_{j \in \mathcal{V}_k, i \in \mathcal{S}_j} h_{kj} \boldsymbol{v}_{ji} s_i + \boldsymbol{n}_k, \quad (1)$$

where $\boldsymbol{n}_k \in \mathbb{C}^r$ is complex addictive Gaussian noise vector.

To decode message d_k, the following constraints must hold

$$\sum_{j \in \mathcal{V}_k, d_k \in \mathcal{S}_j} \boldsymbol{u}_k^{\mathsf{H}} \boldsymbol{v}_{jd_k} h_{kj} \neq 0, \ \forall k = 1, \cdots, K \quad (2)$$

$$\sum_{j \in \mathcal{V}_k, i \in \mathcal{S}_j} \boldsymbol{u}_k^{\mathsf{H}} \boldsymbol{v}_{ji} h_{kj} = 0, \ \ i \notin \mathcal{W}_k \cup \{d_k\}, \quad (3)$$

in which $\boldsymbol{u}_k$ is the k-th receiver decoding vector. Then message d_k can be estimated by

$$\tilde{s}_{d_k} = \Big(\sum_{j \in \mathcal{V}_k, d_k \in \mathcal{S}_j} h_{kj} \boldsymbol{u}_k^{\mathsf{H}} \boldsymbol{v}_{jd_k}\Big)^{-1} \boldsymbol{u}_k^{\mathsf{H}} \Big(\boldsymbol{y}_k - \sum_{j \in \mathcal{V}_k} \sum_{i \in \mathcal{W}_k \cap \mathcal{S}_j} h_{kj} \boldsymbol{v}_{ji} s_i\Big). \quad (4)$$

However, in this interference cancellation scheme, the global instantaneous values of channel coefficients h_{kj} are required which leads to huge traffic overhead and is detrimental to dense Fog-RAN. To significantly lower such overhead for managing interferences, we consider the topological interference management (TIM) approach [9] that aims to align

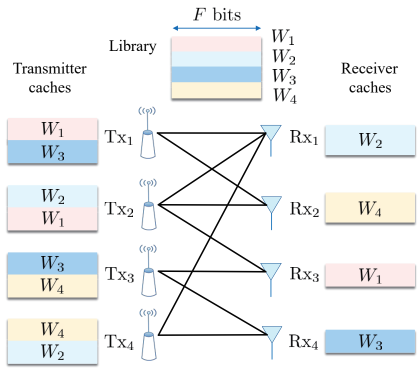

Fig. 1. System model of a 4-user case in which $d_k = k$ for $k = 1, \cdots, 4$. For example, side information for transmitter 1 and receiver 1 is $\mathcal{S}_1 = \{1, 3\}$ and $\mathcal{W}_1$=$\{2\}$. Receiver 1 is connected with transmitters $1, 2, 4$ thus $\mathcal{V}_1 = \{1, 2, 4\}$.

interferences based on mere connectivity information, i.e. the set of channels $h_{kj} \neq 0$.

2.2. Problem Formulation

Specifically, without knowing the values of channel coefficients h_{ij} at transmitters, the following interference alignment conditions are proposed to preserve each desired signal and cancel interferences at receivers [10]

$$\sum_{j \in \mathcal{V}_k, d_k \in \mathcal{S}_j} \boldsymbol{u}_k^{\mathsf{H}} \boldsymbol{v}_{jd_k} \neq 0, \ \forall k = 1, \cdots, K \quad (5)$$

$$\boldsymbol{u}_k^{\mathsf{H}} \boldsymbol{v}_{ji} = 0, \ \forall j \in \mathcal{V}_k, i \in \mathcal{S}_j, i \notin \mathcal{W}_k \cup \{d_k\} \quad (6)$$

which can ensure (2) (3) almost surely. We first define

$$\boldsymbol{U} = [\boldsymbol{u}_k] \in \mathbb{C}^{r \times K}, \ \boldsymbol{V}_k = [\boldsymbol{v}_{ki}] \in \mathbb{C}^{r \times N}$$
$$\bar{\boldsymbol{V}} = [\boldsymbol{V}_1, \cdots, \boldsymbol{V}_K] \in \mathbb{C}^{r \times KN}, \ \boldsymbol{X} = [X_{k,ji}] = \boldsymbol{U}^{\mathsf{H}} \bar{\boldsymbol{V}}.$$

Without loss of generality, the following low rank optimization problem

$$\begin{aligned} \underset{\boldsymbol{X} \in \mathbb{C}^{K \times KN}}{\text{minimize}} \ \ & \text{rank}(\boldsymbol{X}) \\ \text{subject to} \ & \sum_{j \in \mathcal{V}_k, d_k \in \mathcal{S}_j} X_{k,jd_k} = 1, \ \forall k = 1, \cdots, K \\ & X_{k,ji} = 0, \ \forall j \in \mathcal{V}_k, i \in \mathcal{S}_j, i \notin \mathcal{W}_k \cup \{d_k\}. \end{aligned} \quad (7)$$

is proposed to obtain minimal channel uses for interference-free message delivery. We call it as a *generalized matrix completion* problem as it is the generalization for low rank matrix completion problem when only receivers are cache-enabled [11]. It can be re-captured as

$$\begin{aligned} \underset{\boldsymbol{X} \in \mathbb{C}^{K \times KN}}{\text{minimize}} \ \ & \text{rank}(\boldsymbol{X}) \\ \text{subject to} \ \ & \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}, \end{aligned} \quad (8)$$

where $\mathcal{A}(\cdot)$ is a linear operator. And the symmetric degree-of-freedom is given as $\text{DoF} = \lim_{\text{SNR} \to \infty} C_{\text{sum}}(\text{SNR}) / \log \text{SNR} =$

$1/r$, in which C_{sum} is the sum rate of the system and SNR denotes the signal-to-noise-ratio. Note that the resulting generalized matrix completion problem (8) in the complex field is highly intractable because of the rank function.

3. BURER-MONTEIRO APPROACH FOR THE GENERALIZED MATRIX COMPLETION PROBLEM

In this section, we develop a general positive semidefinite (PSD) reparametrization framework for (8) based on the Burer-Monteiro approach [17].

Despite of the intractability of problems like (8) in general, both convex relaxation and non-convex algorithms have been proposed to find achievable solutions effectively. Unfortunately, the nuclear norm $\|\boldsymbol{X}\|_*$ [18] as a well known convex surrogate of rank function, fails for our problem and it always yields a full rank solution [10, 11]. Thus, we investigate non-convex methods based on low rank factorization $\boldsymbol{X} = \boldsymbol{U}^H\boldsymbol{V}$, such as alternating minimization (AltMin) [10] algorithm and bi-factored gradient descent (BFGD) [12] method, which can significantly reduce the storage and computation cost. Here $\boldsymbol{U} \in \mathbb{C}^{r\times K}, \boldsymbol{V} \in \mathbb{C}^{r\times KN}$ where r is the rank. Specifically, AltMin approach optimizes $\boldsymbol{U}$ and $\boldsymbol{V}$ in alternation by fixing the other. It is still costly in computation and fails to scale to large problem sizes. In addition, applying gradient descent approach where $\boldsymbol{U}, \boldsymbol{V}$ are updated separately, BFGD is sensitive to initial points and has slow convergence rate.

In contrast, Riemannian optimization [13] exploits the non-uniqueness of low ran matrix factorization characterized by quotient manifold geometries. Moreover, by harnessing the second order information on Riemannian manifold, Riemannian trust region algorithm achieves a superlinear convergence rate and converges to an *approximate local minimum* [19] with arbitrary initial points. Note that in order to find the minimal rank of (8), there remains the rank parameter r to be determined for non-convex factorization-based algorithms. Therefore, in this paper we propose to alternatively solve the following fixed rank subproblem and rank increase

$$\begin{aligned} &\underset{\boldsymbol{X}\in\mathbb{C}^{K\times KN}}{\text{minimize}} && \frac{1}{2}\|\mathcal{A}(\boldsymbol{X}) - \boldsymbol{b}\|_2^2 \\ &\text{subject to} && \text{rank}(\boldsymbol{X}) = r. \end{aligned} \tag{9}$$

However, no available Riemannian algorithm has been developed for general non-square complex fixed-rank smooth optimization problems (9) so far. In this paper, we propose to reformulate (9) by semidefinite lifting [20] and Burer-Monteiro factorization, then develop a Riemannian trust region algorithm. Specifically, by semidefinite lifting $\boldsymbol{X}$ to

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_{11} & \boldsymbol{Z}_{12} \\ \boldsymbol{Z}_{21} & \boldsymbol{Z}_{22} \end{bmatrix} := \begin{bmatrix} \boldsymbol{U}^{\mathsf{H}} \\ \boldsymbol{V}^{\mathsf{H}} \end{bmatrix} \cdot [\boldsymbol{U}\ \boldsymbol{V}], \tag{10}$$

we can reformulate (9) as

$$\begin{aligned} &\underset{\boldsymbol{Z}\in\mathbb{S}_+^T}{\text{minimize}} && \frac{1}{2}\|\mathcal{B}(\boldsymbol{Z}) - \boldsymbol{b}\|_2^2 \\ &\text{subject to} && \text{rank}(\boldsymbol{Z}) = r \end{aligned} \tag{11}$$

by defining linear operation $\mathcal{B}(\boldsymbol{Z}) = \mathcal{A}(\boldsymbol{X})$. Therein, $\boldsymbol{Z} \in \mathbb{S}_+^T$ means $\boldsymbol{Z}$ is Hermitian positive semidefinite and $T = K + KN$. Furthermore, via Burer-Monteiro approach $\boldsymbol{Z} = \boldsymbol{Y}\boldsymbol{Y}^{\mathsf{H}}$ where $\boldsymbol{Y}^{\mathsf{H}} = [\boldsymbol{L}^{\mathsf{H}}\ \boldsymbol{R}^{\mathsf{H}}]$, (11) can be rewritten as

$$\mathscr{P}: \underset{\boldsymbol{Y}\in\overline{\mathcal{M}}}{\text{minimize}}\ \ f(\boldsymbol{Y}) = f_1(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{H}}), \tag{12}$$

which minimizes a smooth (C^∞) objective function on the non-compact Stiefel manifold $\overline{\mathcal{M}} = \mathbb{C}_*^{T\times r}$, i.e. the set of all $T \times r$ full column rank matrices in the complex field.

Lately, PSD matrices problem with the non-convex Burer-Monteiro approach has been exploited in the multicast beamforming, community detection, phase synchronization, phase retrieval [15, 19] and low rank matrix completion/recovery [21, 22] problems, wherein some global optimality guarantees can be achieved.

4. MANIFOLD OPTIMIZATION ALGORITHM FOR GENERAL MATRIX COMPLETION PROBLEM

In this section, we develop a systematic manifold optimization approach for the generalized matrix completion problem (8). Specifically, trust region algorithm on the Riemannian manifold for the fixed rank subproblem (12) driven by Burer-Monteiro approach is developed.

We note that the factorization $\boldsymbol{Z} = \boldsymbol{Y}\boldsymbol{Y}^{\mathsf{H}}$ remains invariant under the transformation $\boldsymbol{Y} \to \boldsymbol{Y}\boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is any unitary matrix. Unfortunately, the convergence properties of second-order methods [13] are severely hampered by this invariance property. To address this issue, we consider optimizing the problem on the quotient manifold $\mathcal{M} = \mathbb{C}_*^{T\times r}/\mathcal{U}(r)$

$$\underset{[\boldsymbol{Y}]\in\mathcal{M}}{\text{minimize}}\ \ f([\boldsymbol{Y}]) = \frac{1}{2}\|\mathcal{B}(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{H}}) - \boldsymbol{b}\|_2^2, \tag{13}$$

where $[\boldsymbol{Y}] = \{\boldsymbol{Y}\boldsymbol{\Theta} : \boldsymbol{\Theta} \in \mathcal{U}(r)\}$ is called an equivalent class and $\mathcal{U}(r)$ is the set of all $r\times r$ unitary matrices. Now we let the horizontal space $\mathcal{H}_{\boldsymbol{Y}}\mathcal{M}$ and vertical space $\mathcal{V}_{\boldsymbol{Y}}\mathcal{M}$ denote the directions tangent and vertical to the set of equivalent classes $[\boldsymbol{Y}]$, respectively. Then, the tangent space $\mathcal{T}_{\boldsymbol{Y}}\mathcal{M}$ at a given point $\boldsymbol{Y}$ can be decomposed into two complementary spaces

$$\mathcal{T}_{\boldsymbol{Y}}\mathcal{M} = \mathcal{V}_{\boldsymbol{Y}}\mathcal{M} \oplus \mathcal{H}_{\boldsymbol{Y}}\mathcal{M}, \tag{14}$$

We now define $\text{grad}f(\boldsymbol{Y}), \text{Hess}f(\boldsymbol{Y})$, i.e. the horizontal representation of Riemannian gradient and Riemannian Hessian in the tangent space $\mathcal{T}_{\boldsymbol{Y}}\mathcal{M}$. Recall that the Riemannian gradient and Hessian represented in the horizontal space for quotient manifold are derived from Euclidean gradient and Hessian. Thus, the trust region method on the Riemannian manifold [13] is a second-order optimization algorithm, which sequentially solves the subproblem

$$\begin{aligned} &\underset{\boldsymbol{\xi}\in\mathcal{H}_{\boldsymbol{Y}}\mathcal{M}}{\text{minimize}} && f(\boldsymbol{Y}) + g_{\boldsymbol{Y}}(\boldsymbol{\xi}, \text{grad}f(\boldsymbol{Y})) + \frac{1}{2}g_{\boldsymbol{Y}}(\boldsymbol{\xi}, \text{Hess}f(\boldsymbol{Y})[\boldsymbol{\xi}]) \\ &\text{subject to} && g_{\boldsymbol{Y}}(\boldsymbol{\xi}, \boldsymbol{\xi}) \le \delta^2. \end{aligned} \tag{15}$$

In order to develop trust region algorithm on the Riemannian manifold, we should define Riemannian metric $g_{\boldsymbol{Y}}$. Consequently, the Riemannian iteration is given by

$$\boldsymbol{Y}_{k+1} = \mathcal{R}_{\boldsymbol{Y}_k}(\boldsymbol{\eta}_k), \tag{16}$$

Algorithm 1: Riemannian Algorithm for (8)

Input: $\mathcal{B}, \boldsymbol{b}, N, l$, accuracy ε, τ.
for $r = 1, \cdots, N$ **do**
  Initialize: $\mathbf{Y}_0^{[r]} \in \mathbb{C}_*^{N \times r}$
  while $\|grad f(\mathbf{Y}_k^{[r]})\|_F \geq \tau$ **do**
    Choose search direction $\eta_{\mathbf{k}}$ and step size α_k,
    $\mathbf{Y}_{\mathbf{k+1}}^{[\mathbf{r}]} = \mathcal{R}_{\mathbf{Y}_{\mathbf{k}}^{[\mathbf{r}]}}(\alpha_{\mathbf{k}} \eta_{\mathbf{k}})$
  end
  if $f(\mathbf{Y}^{[r]}) < \varepsilon$ **then**
    return $\mathbf{Y}^{[\mathbf{r}]}$
  end
end
Output: $\boldsymbol{Y}^{[r]}$ and rank r.

where retraction operator $\mathcal{R}_{\boldsymbol{Y}_k}(\cdot)$ defines a pullback from the tangent space onto the manifold, and $\boldsymbol{\eta}_k$ is determined from (15). Please refer to [13] for more details about the geometry of Riemannian manifold.

5. SIMULATION

In this section, we provide results from numerical experiments to show the benefit of both transmitter and receiver caches by comparing several known algorithms. Specifically we compare the following algorithms and their settings:

- AltMin [10]: AltMin is adopted when only transmitters are cache-enabled [10]. We use it to solve (9) given r.
- BFGD [12]: We use the proposed initialization strategy in [12] to solve (9) for each r.
- Proposed RTR: The proposed Riemannian trust region algorithm is implemented with *Manopt* [23] toolbox.
- RTR-UC [11]: Same Riemannina algorithm is used but only receivers are cache-enabled [11].

For all algorithms, we let rank r increase from 1 and terminate the iteration when cost function $C = \|\mathcal{A}(\boldsymbol{X}) - \boldsymbol{b}\|/K$ falls below 10^{-3} with randomized initializations strategy except for BFGD.

In our simulation, we consider the setting that $N = K$, $d_k = k$, and $k \in \mathcal{S}_k \cap \mathcal{V}_k, k \notin \mathcal{W}_k$ to ensure that each message can always be decoded by increasing the channel uses. The remaining elements of side information and network topology are chosen uniformly. Specifically, $\mathcal{S}_k = \{k\}$ for RTR-UC.

To exploit the benefit of caches for both transmitters and users, we perform two experiments with $K = 20$ and full cooperation, i.e. $\mathcal{S}_k = [K]$ for all algorithms except for RTR-UC. In the first experiment (see Fig. 2), we evaluate four algorithms for fixed network topology by varying the cache size for each user. In the second experiment (see Fig. 3), we fix the receiver cache size by varying the number of connected links. From both numerical experiments, we can see that the proposed approach (marked as "Proposed RTR" in the figures) outperforms other benchmark algorithms and the non-transmitter-cooperation case greatly. These results confirm that our approach can take advantages of caches in both transmitters and receivers.

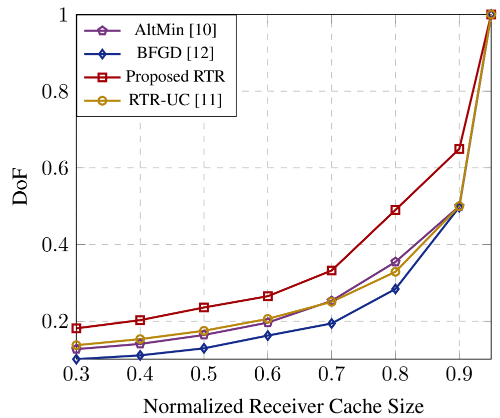

Fig. 2. Normalized receiver cache size ($|\mathcal{W}_k|/K$) versus DoF while $K = 20$ with full transmitter cooperation, i.e. $|\mathcal{S}_k| = 20$. In this experiment, we set the number of connected links as 300, i.e. $\sum_k |\mathcal{V}_k| = 300$.

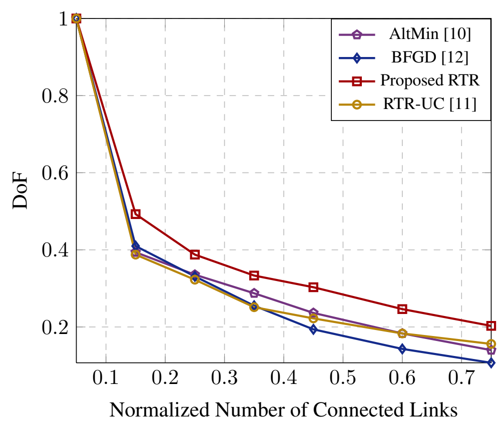

Fig. 3. Normalized link number ($\sum_k |\mathcal{V}_k|/K^2$) versus DoF while $K = 20$, $|\mathcal{W}_k| = 8$ under full transmitter cooperation.

6. CONCLUSION

In this work, we present a generalized matrix completion approach to large scale content delivery and sharing in cache-aided dense Fog-RAN based only on the network topology information. This novel framework provides a new and flexible way to exploit the benefits of data caching at both the transmitters and receivers for cooperative content delivery. To design efficient algorithms for the non-convex generalized matrix completion problem, we propose a positive semidefinite lifting approach to reformulate the original problem into a low-rank positive semi-definite matrix optimization problem. The lifting allows us to develop a Burer-Monteiro approach based Riemannian trust-region algorithm for solving a sequence of fixed-rank optimization problems through the procedure of rank increment, thereby finding a minimum rank solution. Test results showed that the proposed approach outperforms the benchmark methods in the literature and enjoys considerable transmitter cooperation gains.

7. REFERENCES

[1] S. Bi, R. Zhang, Z. Ding, and S. Cui, "Wireless communications in the era of big data," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 190–199, Oct. 2015.

[2] Y. Shi, J. Zhang, K. B. Letaief, B. Bai, and W. Chen, "Large-scale convex optimization for ultra-dense Cloud-RAN," *IEEE Commun. Mag.*, vol. 22, no. 3, pp. 84–91, June 2015.

[3] J. Cheng, Y. Shi, B. Bai, and W. Chen, "Computation offloading in Cloud-RAN based mobile cloud computing system," in *2016 IEEE Int. Conf. Commun. (ICC)*, May. 2016, pp. 1–6.

[4] Songze Li, Qian Yu, Mohammad Ali Maddah-Ali, and A Salman Avestimehr, "A scalable framework for wireless distributed computing," *arXiv preprint arXiv:1608.05743*, 2016.

[5] M. A. Maddah-Ali and U. Niesen, "Coding for caching: fundamental limits and practical challenges," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 23–29, Aug. 2016.

[6] C. Fan, Y. J. Zhang, and X. Yuan, "Dynamic nested clustering for parallel PHY-layer processing in Cloud-RANs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1881–1894, Mar. 2016.

[7] Fan Xu, Meixia Tao, and Kangqi Liu, "Fundamental tradeoff between storage and latency in cache-aided wireless interference networks," *arXiv preprint arXiv:1605.00203*, 2016.

[8] N. Naderializadeh, M. A. Maddah-Ali, and A. S. Avestimehr, "Fundamental limits of cache-aided interference management," in *2016 IEEE Int. Symp. Inform. Theory (ISIT)*, July 2016, pp. 2044–2048.

[9] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[10] X. Yi and G. Caire, "Topological coded caching," in *2016 IEEE Int. Symp. Inform. Theory (ISIT)*, July 2016, pp. 2039–2043.

[11] K. Yang, Y. Shi, and Z. Ding, "Low-rank matrix completion for mobile edge caching in Fog-RAN via Riemannian optimization," in *Proc. 2016 IEEE Global Commun. Conf.*, pp. 1–6.

[12] Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi, "Finding low-rank solutions to matrix problems, efficiently and provably," *arXiv preprint arXiv:1606.03168*, 2016.

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

[14] Nicolas Boumal, P-A Absil, and Coralia Cartis, "Global rates of convergence for nonconvex optimization on manifolds," *arXiv preprint arXiv:1605.08101*, 2016.

[15] Nicolas Boumal, Vladislav Voroninski, and Afonso S Bandeira, "The non-convex Burer-Monteiro approach works on smooth semidefinite programs," *arXiv preprint arXiv:1606.04970*, 2016.

[16] Rong Ge, Chi Jin, and Yi Zheng, "No spurious local minima in nonconvex low rank problems: A unified geometric analysis," *arXiv preprint arXiv:1704.00708*, 2017.

[17] Samuel Burer and Renato DC Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Math. Program.*, vol. 95, no. 2, pp. 329–357, 2003.

[18] Emmanuel J Candès and Benjamin Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[19] Afonso S Bandeira, Nicolas Boumal, and Vladislav Voroninski, "On the low-rank approach for semidefinite programs arising in synchronization and community detection," in *Proc. 29th Conf. Learn. Theory, COLT 2016, New York, NY, June*, 2016, pp. 23–26.

[20] Qinqing Zheng and John Lafferty, "Convergence analysis for rectangular matrix completion using Burer-Monteiro factorization and gradient descent," *arXiv preprint arXiv:1605.07051*, 2016.

[21] Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi, "Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach," *arXiv preprint arXiv:1609.03240*, 2016.

[22] Srinadh Bhojanapalli, Behnam Neyshabur, and Nati Srebro, "Global optimality of local search for low rank matrix recovery," in *Adv. Neural. Inf. Process. Syst. (NIPS)*, 2016, pp. 3873–3881.

[23] Nicolas Boumal, Bamdev Mishra, P-A Absil, and Rodolphe Sepulchre, "Manopt, a matlab toolbox for optimization on manifolds," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1455–1459, 2014.