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

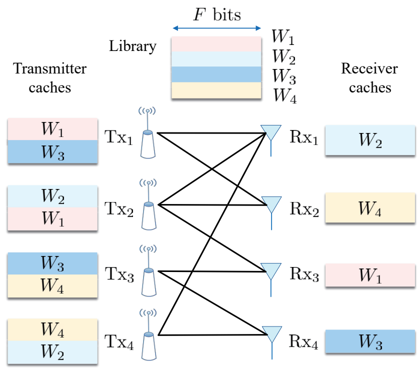


**Fig. 1**. System model of a 4-user case in which $d_k = k$ for $k = 1, \cdots, 4$. For example, side information for transmitter 1 and receiver 1 is $\mathcal{S}_1 = \{1, 3\}$ and $\mathcal{W}_1$=$\{2\}$. Receiver 1 is connected with transmitters $1, 2, 4$ thus $\mathcal{V}_1 = \{1, 2, 4\}$.

### 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

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

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

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

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