

# Topological Interference Management With User Admission Control via Riemannian Optimization

Yuanming Shi, *Member, IEEE*, Bamdev Mishra, and Wei Chen, *Senior Member, IEEE*

**Abstract**—Topological interference management (TIM) provides a promising way to manage interference only based on the network connectivity information. Previous works on the TIM problem mainly focus on using the index coding approach and graph theory to establish conditions of network topologies to achieve the feasibility of topological interference management. In this paper, we propose a novel user admission control approach via sparse and low-rank optimization to maximize the number of admitted users for achieving the feasibility of topological interference management. However, the resulting sparse and low-rank optimization problem is non-convex and highly intractable, for which the *conventional convex relaxation approaches* are inapplicable, e.g., a simple  $\ell_1$ -norm relaxation approach yields the objective unbounded and non-convex. To assist efficient algorithms design for the formulated rank-constrained (i.e., degrees-of-freedom (DoFs) allocation)  $\ell_0$ -norm maximization (i.e., user capacity maximization) problem, we propose a novel *non-convex but smoothed  $\ell_1$ -regularized minimization* approach to induce sparsity pattern with bounded objective values. We further develop a Riemannian trust-region algorithm to solve the resulting rank-constrained smooth non-convex optimization problem via exploiting the quotient manifold of fixed-rank matrices. Simulation results demonstrate the effectiveness and optimality of the proposed Riemannian algorithm to maximize the number of admitted users for topological interference management.

**Index Terms**—Topological interference alignment, user admission control, sparse and low-rank modeling, Riemannian optimization, quotient manifold.

## I. INTRODUCTION

THE popularization of innovative applications and new services, such as Internet of Things (IoT) and wearable devices [1], is driving the era of wireless big data [2], thereby revolutionizing the segments of the society. In particular, with

Manuscript received January 10, 2017; revised June 19, 2017; accepted August 21, 2017. Date of publication September 6, 2017; date of current version November 9, 2017. This work was supported in part by Shanghai Sailing Program under Grant 16YF1407700, the National Nature Science Foundation of China under Grant 61601290, the National Natural Science Foundation of China under Project 61671269 and Project 61621091, and in part by the Chinese National 973 Program under Project 2013CB336600. The associate editor coordinating the review of this paper and approving it for publication was V. Cadambe. (*Corresponding author: Yuanming Shi.*)

Y. Shi is with the School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China (e-mail: shiym@shanghaitech.edu.cn).

B. Mishra is with the Core Machine Learning Team, Amazon, Bangalore 560055, India (e-mail: bamdevm@amazon.com).

W. Chen is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: wchen@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2017.2747539

ultra-low latency and ultra-reliable requirements, Tactile Internet [3] enables a new paradigm shift from content-delivery to skill-set delivery networks. Network densification [4], supported by the advanced wireless technologies (e.g., massive MIMO [5], Cloud-RAN [6], [7], and small cells [8], [9]), becomes the key enabling technology to accommodate the exponential mobile data traffic growth, as well as provides ubiquitous connectivity for massive devices. However, by adding more radio access points per volume, interference becomes the bottleneck to harness the benefits of wireless network densification. Although the recent development of interference alignment [10] and interference coordination [11] have been shown to be effective in the interference-limited communication scenarios, the significant signaling overhead of obtaining the global channel state information (CSI) limits its applicability in dense wireless networks [12].

To reduce the CSI acquisition overhead and make it scalable in dense wireless networks, the topological interference management (TIM) approach was proposed in [12] to manage interference only based on the network connectivity information. However, establishing the feasibility of topological interference management is a challenging task. In the slow fading scenario, i.e., channels stay constant during transmission, the TIM problem turns out to be equivalent to the index coding problem for the linear coding schemes [13]. However, the index coding problem is NP-hard in general and only some special cases have been solved [12], [14]. Furthermore, the topological interference management problems with transmitters cooperation and multiple transmitter antennas were investigated in [15] and [16], respectively. In the fast fading scenario, the graph theory and matroids theory were adopted to find the conditions of network topologies to achieve a certain amount of DoF allocation [17], [18]. A low-rank matrix completion approach with Riemannian algorithms has recently been proposed in [19] to find the minimum channel uses to achieve feasibility for *any* network topology.

In contrast, in this paper, we present a different viewpoint for the TIM problem: *given any network topology and DoF allocation for all the users, we aim at finding the maximum number of admitted users to achieve the feasibility of topological interference management, i.e., all the admitted users achieve the DoF requirements.* We call this problem as *user admission control* in topological interference management. Note that the user admission control problem is fundamentally different from the original TIM problem [19], where all the users are assumed to be admitted, and the

corresponding preceding and decoding schemes aim at maximizing the achievable DoFs for all the users. User admission control is critical in wireless communication networks (i.e., cognitive radio access networks [20], heterogeneous networks [21] and Cloud-RAN [22]) when quality-of-services (QoS) requirements are unsatisfied or the channel conditions are unfavorable [23]. Although the user admission control problems are normally non-convex mixed combinatorial optimization problems, a large body of recent work has demonstrated the effectiveness of convex relaxation for solving such problems [20]–[23] based on the sum-of-infeasibilities in optimization theory [24]. This is achieved by relaxing the original non-convex  $\ell_0$ -norm minimization problem for user admission control to the convex  $\ell_1$ -norm minimization problem [24], [25].

Unfortunately, the user admission control problem in topological interference management turns out to be highly intractable, which needs to optimize over continuous and combinatorial variables. To address the intractability, in this paper, we propose a sparse and low-rank modeling framework to compute the proposed solutions within polynomial time. In this model, sparsity of the diagonal entries of the matrix (i.e., the number of non-zero entries) represents the number of the admitted users. The fixed low-rank constraint indicates the DoF allocation [19]. However, the unique challenges arise in the proposed sparse and low-rank optimization model including the non-convex fixed-rank constraint and user capacity maximization objective function, i.e.,  $\ell_0$ -norm objective *maximization*. A simple  $\ell_1$ -norm relaxation approach yields the objective unbounded and non-convex. Novel algorithms thus need to be developed.

#### A. Related Works

1) *User Admission Control*: In dense wireless networks, user admission control is critical to maximize the user capacity while satisfying the QoS requirements for all the admitted users. To address the NP-hardness of the mixed combinatorial optimization problem, sparse optimization (e.g.,  $\ell_0$ -norm minimization) approach, supported by the efficient algorithms (e.g.,  $\ell_1$ -norm convex relaxation [20], [21] and the iterative reweighted  $\ell_2$ -algorithm [22]), provided an efficient way to find high quality solutions. However, convex relaxation approach is inapplicable in our sparse and low-rank optimization problem due to the  $\ell_0$ -norm maximization as the objective. For the  $\ell_1$ -norm relaxation approach, it yields a  $\ell_1$ -norm maximization problem, which is still non-convex. Furthermore, maximizing  $\ell_1$ -norm shall yield unbounded values. We thus propose a novel quadratic term regularized  $\ell_1$ -norm to bound the objective function. Note that the regularized  $\ell_1$ -norm relaxation is still *non-convex*.

2) *Low-Rank Models*: Low-rank models [26], [27] inspire enormous applications in machine learning, recommendation systems, sensor localization, etc. Due to the non-convexity of low-rank constraint or objective, many heuristic algorithms with optimality guarantees have been proposed in the last few years. In particular, convex relaxation approach using nuclear norm [28] provides a polynomial time complexity algorithm with optimality guarantees via convex geometry and

conic integral geometry analysis [29]. The other popular way for low-rank optimization is based on matrix factorization, e.g., the alternating minimization [27] and Riemannian optimization method [30].

In particular, the Riemannian optimization framework has the capability of exploiting the Riemannian quotient manifold of the fixed-rank matrices in the search space. Furthermore, the Riemannian conjugate gradient and trust-region algorithms are globally convergent (i.e., they converge to first-order and second-order KKT points on manifolds [31]) with superlinear [32] and even quadratic convergence rates [31]. The Riemannian algorithms thus achieve faster convergence rates and higher accurate solutions compared with the alternating minimization and gradient descent algorithms [19]. However, due to the  $\ell_0$ -norm maximization objective and non-convex fixed-rank constraint, the sparse and low-rank optimization problem for user admission control reveals unique challenges. We thus propose novel regularized formulations that allow to exploit the Riemannian geometry of fixed-rank matrices and induce sparsity in the diagonal entries of the matrices.

3) *Riemannian Optimization*: Basically, the Riemannian optimization approach requires the smoothness of the objective function [33]. For the smooth non-convex optimization over manifolds, the Riemannian algorithms turn out to be able to achieve global optimality for some specific high-dimensional statistical optimization problems, e.g., dictionary learning [34], generalized phase retrieval [34], and community detection problems [35]. When the sample size is sufficiently large, all of these problems enjoy the benign geometric structure that all the local optimum are globally optimal and all the saddle points can be escaped by the Riemannian trust-region algorithms [34], [35]. However, due to the non-convex and non-smooth objective function, we can not directly apply the existing Riemannian optimization approaches to solve the proposed sparse and low-rank optimization framework for user admission control. In this paper, we thus propose a regularized smoothed  $\ell_1$ -minimization approach supported by the Riemannian trust-region algorithm to find the maximal number of admitted users.

Based on the above discussions, in contrast to the previous works on user admission control [20]–[23],  $\ell_1$ -norm relaxation approaches, and low-rank optimization problems [26], [27], [30], [36], we need to address the following *coupled* challenges to solve the sparse and low-rank optimization for user admission control in topological interference management:

- The objective of *maximizing* the non-convex  $\ell_0$ -norm to maximize the user capacity, i.e., the number of admitted users;
- Non-convex fixed-rank constraint to achieve a certain amount of DoF allocation.

Therefore, unique challenges arise in the user admission control problem for topological interference management. We need to re-design the sparsity-inducing function and the efficient approaches to deal with the fixed-rank constraint.

## B. Contributions

In this paper, we propose a sparse and low-rank optimization framework for user admission control in topological interference management. The Riemannian trust-region algorithm is developed to solve the proposed regularized smoothed  $\ell_1$ -norm sparsity inducing minimization problem, thereby guiding user selection. The main contributions are summarized as follows:

- 1) We propose a novel sparse and low-rank optimization framework to maximize the number of admitted users for achieving the feasibility of topological interference management.
- 2) To address the difficulty that a simple  $\ell_1$ -norm relaxation approach yields the objective in the sparse and low-rank optimization unbounded and non-convex, we propose a novel quadratic regularized approach to bound the objective in the procedure of the sparsity inducing. Note that the regularized  $\ell_1$ -norm minimization is still *non-convex*.
- 3) We further propose novel regularized formulations that allow to exploit the Riemannian geometry of fixed-rank matrices and induce sparsity in matrices. This is achieved by relaxing the  $\ell_0$ -norm maximization problem as the quadratic regularized smoothed  $\ell_1$ -norm minimization problem and regularizing the affine constraints into the objective.
- 4) A Riemannian trust-region algorithm is further developed to solve the resulting rank-constrained smooth optimization problem for sparsity inducing. The Riemannian trust-region algorithm is globally convergent with superlinear convergence rate, i.e., it converges to the second-order KKT points starting from any random initialization.
- 5) Numerical results demonstrate the effectiveness and optimality of the proposed Riemannian trust-region algorithm to maximize the user capacity for topological interference management.

## C. Organization

The remainder of the paper is organized as follows. Section II presents the system model and problem formulation. A sparse and low-rank optimization framework for user admission control is proposed in Section III. The Riemannian optimization algorithm is developed in Section IV. The ingredients of optimization on quotient manifold are presented in Section V. Numerical results are illustrated in Section VI. Finally, conclusions and discussions are presented in Section VII.

*Notations:* Throughout this paper,  $\|\cdot\|_p$  is the  $\ell_p$ -norm. Boldface lower case and upper case letters represent vectors and matrices, respectively.  $(\cdot)^{-1}$ ,  $(\cdot)^T$ ,  $(\cdot)^H$  and  $\text{Tr}(\cdot)$  denote the inverse, transpose, Hermitian and trace operators, respectively. We use  $\mathbb{C}$  and  $\mathbb{R}$  to represent complex domain and real domain, respectively.  $\mathbb{E}[\cdot]$  denotes the expectation of a random variable.  $|\cdot|$  stands for either the size of a set or the absolute value of a scalar, depending on the context. We denote  $\mathbf{A} = \text{diag}\{x_1, \dots, x_N\}$  and  $\mathbf{I}_N$  as a diagonal matrix of order  $N$  and the identity matrix of order  $N$ , respectively.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we present the channel model, followed by the user admission control problem to achieve the feasibility of topological interference management.

### A. Channel Model

Consider the topological interference management problem in the partially connected  $K$ -user interference channel with each node equipped with a single antenna [12], [19]. Let  $\mathcal{V}$  be the index set of the connected transceiver pairs such that the channel coefficient  $h_{ij}$  between the transmitter  $j$  and receiver  $i$  is non-zero if  $(i, j) \in \mathcal{V}$ , and is zero otherwise. Each transmitter  $i$  wishes to send a message  $W_i$  to its corresponding receiver  $i$ . The message  $W_i$  is encoded into a vector  $\mathbf{x}_i \in \mathbb{C}^r$  of length  $r$ . Therefore, over the  $r$  channel uses, the received signal  $\mathbf{y}_i \in \mathbb{C}^r$  at receiver  $i$  is given by

$$\mathbf{y}_i = h_{ii}\mathbf{x}_i + \sum_{j \in \mathcal{V}, j \neq i} h_{ij}\mathbf{x}_j + \mathbf{z}_i, \quad \forall i = 1, \dots, K,$$

where  $\mathbf{z}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_r)$  is the additive noise at receiver  $i$ . We consider the block fading channel, where the channel coefficients stay constant during transmission, i.e., the channel coherence time is larger than channel uses  $r$  for transmission [12], [19]. We assume that each transmitter has an average power constraint, i.e.,  $\mathbb{E}[\|\mathbf{x}\|^2] \leq r\rho$  with  $\rho > 0$  as the maximum average transmit power.

The rate tuple  $(R_1, \dots, R_K)$  is said to be achievable if there exists a  $(2^{rR_1}, \dots, 2^{rR_K}, r)$  code scheme such that the average decoding error probability is vanishing as the code length  $r$  approaches infinity. Here, we assume that each message  $W_k$  is uniformly and independently chose over the  $K$  message sets  $\mathcal{W}_k := [1 : 2^{rR_k}]$ . In this paper, we choose our performance metric as the symmetric DoF [12], [15], i.e., the highest DoF achieved by all the users simultaneously,  $d_{\text{sym}} = \limsup_{\rho \rightarrow \infty} \sup_{(R_{\text{sym}}, \dots, R_{\text{sym}}) \in \mathcal{C}} \frac{R_{\text{sym}}}{\log \rho}$ , where  $\mathcal{C}$  is the capacity region defined as the set of all the achievable rate tuples. The metric of DoF gives the first-order measurement of data rates [37].

### B. Topological Interference Management

In this paper, we restrict the class of the linear interference management strategies [10], [12], [19]. Specifically, each transmitter  $i$  encodes its message  $W_i$  by a linear precoding vector  $\mathbf{v}_i \in \mathbb{C}^r$  over  $r$  channel uses:  $\mathbf{x}_i = \mathbf{v}_i s_i$ , where  $s_i \in \mathbb{C}$  is the transmitted data symbol.<sup>1</sup> Here the precoding vectors  $\mathbf{v}_i$ 's only depend on the knowledge of network topology  $\mathcal{V}$ . In this paper, we assume that the network connectivity information  $\mathcal{V}$

<sup>1</sup>For illustrative purpose, we only consider the scalar linear scheme, i.e., only one symbol for each message will be sent over  $r$  channel uses. The proposed approach can be extended to more general scenarios with vector linear coding scheme, i.e., each user sends multiple data symbols over  $r$  channel uses [19]. The basic ideas is to generalize the interference alignment conditions as (1)  $\det(\mathbf{U}_i^H \mathbf{V}_i) \neq 0, \forall i = 1, \dots, K$ ; (2)  $\mathbf{U}_i^H \mathbf{V}_j = \mathbf{0}, \forall i \neq j, (i, j) \in \mathcal{V}$ , where  $\mathbf{V}_j \in \mathbb{C}^{r \times M_j}$  and  $\mathbf{U}_i \in \mathbb{C}^{r \times M_i}$  be the precoding matrix at transmitter  $j$  and the receiver combining matrix at receiver  $i$ , respectively. Here, we assume that each message  $W_j$  is split into  $M_j$  independent scalar data streams for each user.

is available at the transmitters. Therefore, over the  $r$  channel uses, the received signal  $\mathbf{y}_i \in \mathbb{C}^r$  at receiver  $i$  can be rewritten as

$$\mathbf{y}_i = h_{ii}\mathbf{v}_i s_i + \sum_{j \in \mathcal{V}, j \neq i} h_{ij}\mathbf{v}_j s_j + \mathbf{z}_i, \quad \forall i = 1, \dots, K.$$

Let  $\mathbf{u}_i \in \mathbb{C}^r$  be the decoding vector for each message  $W_i$  at receiver  $i$ . In the regime of asymptotically high signal-to-noise ratio (SNR), to accomplish decoding, we impose the following interference alignment condition [12], [19] for the precoding and decoding vectors:

$$\mathbf{u}_k^H \mathbf{v}_k \neq 0, \quad \forall k = 1, \dots, K, \quad (1)$$

$$\mathbf{u}_k^H \mathbf{v}_i = 0, \quad \forall i \neq k, (i, k) \in \mathcal{V}, \quad (2)$$

where the first condition is to preserve the desired signal and the second condition is to align and cancel the interference signals. If conditions (1) and (2) are satisfied, the parallel interference-free channels can be obtained over  $r$  channel uses. Therefore, the symmetric DoF of  $1/r$  is achieved for each message  $W_i$  [12]. We call this problem as *topological interference management* [12], [19], as only network topology information is required to establish the interference alignment conditions.

However, establishing the conditions on  $r$ ,  $K$  and  $\mathcal{V}$  to achieve feasibility of the interference alignment conditions (1) and (2) is challenging. In particular, given a number of users  $K$  and channel uses  $r$  (or DoF allocation  $1/r$ ), the index coding approach [12] and graph theory [15], [17], [18] were adopted to establish the conditions on the network topologies  $\mathcal{V}$  to achieve feasibility for the interference alignment conditions (1) and (2). The low-rank matrix completion approach [19] has recently been proposed to find the minimum number of channel uses  $r$  satisfying conditions (1) and (2), given any network topology information  $\mathcal{V}$  and the number of users  $K$ . The feasibility conditions of antenna configuration for interference alignment in MIMO interference channel has also been extensively investigated using algebraic geometry [38]–[40].

In this paper, we put forth a different point of view on the feasibility conditions of topological interference management: given a number of  $K$  users with any network topology  $\mathcal{V}$  and the symmetric DoF allocation  $1/r$ , we present a novel user admission control approach to find the maximum number of the admitted users while satisfying the interference alignment conditions (1) and (2). Although user admission control has been extensively investigated in the scenarios of multiuser coordinated beamforming [23], cognitive radio networks [20], heterogeneous cellular networks [21] and Cloud-RAN [22], this is the first work using the principle of user admission control in the framework of topological interference management. This shall provide a systematic framework for efficient algorithms design, as well as provide numerical insights into this challenging problem of topological interference management.

### III. A SPARSE AND LOW-RANK OPTIMIZATION FRAMEWORK FOR USER ADMISSION CONTROL

In this section, we present a user admission control approach to maximize the user capacity, i.e., find the maximum number of admitted users while satisfying the interference alignment conditions (1) and (2). This viewpoint is different from the previous works on finding the conditions of network topologies and minimal channel uses [19] to achieve the feasibility of interference alignment [12], [15], [17], [18].

#### A. Feasibility of Interference Alignment

Given any network connectivity information  $\mathcal{V}$  for the partially connected  $K$ -user interference channel, we say that the symmetric DoF allocation  $1/r$  is feasible if there exists precoding vectors  $\mathbf{v}_i \in \mathbb{C}^r$  and decoding vectors  $\mathbf{u}_i \in \mathbb{C}^r$  such that the interference alignment conditions (1) and (2) are satisfied. Specifically, the feasibility of topological interference management problem can be formulated as

$$\begin{aligned} \mathcal{F} : \text{find } & \{\mathbf{v}_i\}, \{\mathbf{u}_i\} \\ \text{subject to } & \mathbf{u}_i^H \mathbf{v}_i \neq 0, \quad \forall i = 1, \dots, K, \\ & \mathbf{u}_i^H \mathbf{v}_j = 0, \quad \forall (i, j) \in \mathcal{V}, \end{aligned} \quad (3)$$

where  $\mathbf{v}_i \in \mathbb{C}^r$  and  $\mathbf{u}_i \in \mathbb{C}^r$  are optimization variables.

However, the solutions to the feasibility problem (3) are unknown in general. In particular, the index coding approach [12] and the graph theory [15], [17], [18] were adopted to establish the conditions on the network topology  $\mathcal{V}$  to achieve feasibility of interference alignment. On the other hand, the low-rank matrix completion approach was proposed in [19] to find the minimum number of channel uses to achieve interference alignment feasibility for any network topology.

In contrast, in this paper, our goal is to maximize the user capacity, i.e., finding the maximum number of admitted users while satisfying the interference alignment conditions:

$$\begin{aligned} \text{maximize } & |S| \\ & \text{subject to } \mathbf{u}_i^H \mathbf{v}_i \neq 0, \quad \forall i \in S, \\ & \mathbf{u}_i^H \mathbf{v}_j = 0, \quad \forall i \neq j, i, j \in S, (i, j) \in \mathcal{V}, \end{aligned} \quad (4)$$

where  $S \subseteq \{1, \dots, K\}$  is the admitted users,  $\mathbf{v}_i \in \mathbb{C}^r$  and  $\mathbf{u}_i \in \mathbb{C}^r$ . This problem is called as the *user admission control problem*. Unfortunately, it turns out to be highly intractable due to the non-convex quadratic constraints and the non-convex combinatorial objective function. To assist efficient algorithms design, in this paper, we propose a sparse and low-rank optimization for user admission control via exploiting the sparse and low-rank structures in problem (4).

#### B. Sparse and Low-Rank Optimization Paradigms for User Admission Control

Let  $\mathbf{X} = [X_{ij}] \in \mathbb{C}^{K \times K}$  with  $X_{ij} = \mathbf{u}_i^H \mathbf{v}_j \in \mathbb{C}$ . The interference alignment conditions (1) and (2) thus can be rewritten as

$$\begin{aligned} X_{kk} & \neq 0, \quad \forall k = 1, \dots, K, \\ X_{ki} & = 0, \quad \forall i \neq k, (i, k) \in \mathcal{V}. \end{aligned}$$

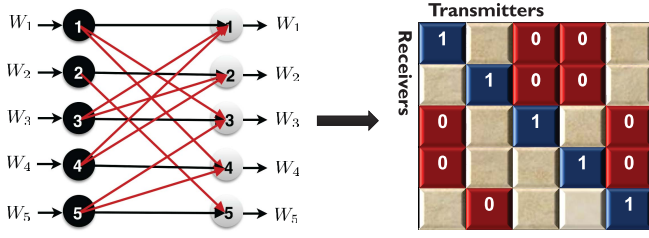


Fig. 1. (a) The topological interference alignment problem for the partially connected  $K$ -user interference channel with only the knowledge of the network connectivity information available. The interference links are marked as red while the desired links are marked as black. (b) The corresponding incomplete matrix with “0” indicating interference alignment and cancellation and “1” representing desired signal preserving.

For other entries  $X_{ki}, \forall (k, i) \notin \mathcal{V}$ , they can be any values. As  $\mathbf{X} = [\mathbf{u}_i^H \mathbf{v}_j] = \mathbf{U}^H \mathbf{V} \in \mathbb{C}^{K \times K}$  with  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_K]^H \in \mathbb{C}^{K \times r}$  and  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K] \in \mathbb{C}^{r \times K}$ , we have  $\text{rank}(\mathbf{X}) = r$ . The achievable symmetric DoF is thus given by [19]

$$\text{DoF} = 1/\text{rank}(\mathbf{X}) = 1/r,$$

a low-rank matrix completion problem was proposed in [19] to find the minimum channel uses while satisfying the interference alignment conditions. Fig. 1 demonstrates the procedure of transforming the topological interference alignment conditions (1) and (2) into the associated incomplete matrix  $\mathbf{X}$ .

Define  $\mathbf{X}(\mathcal{S}) \in \mathbb{C}^{|\mathcal{S}| \times |\mathcal{S}|}$  as the submatrix of  $\mathbf{X}$ , i.e.,  $\mathbf{X}(\mathcal{S}) = [X_{ij}]_{i,j \in \mathcal{S}}$ . The rank of the submatrix  $\mathbf{X}(\mathcal{S})$  equals  $r$ . The user admission control problem (4) can be further reformulated as follows:

$$\begin{aligned} & \underset{\mathbf{X} \in \mathbb{C}^{K \times K}, \mathcal{S}}{\text{maximize}} && |\mathcal{S}| \\ & \text{subject to} && \text{rank}(\mathbf{X}(\mathcal{S})) = r, \\ & && X_{ii} \neq 0, \quad \forall i \in \mathcal{S}, \\ & && X_{ij} = 0, \quad \forall i \neq j, i, j \in \mathcal{S}, (i, j) \in \mathcal{V}, \end{aligned} \quad (5)$$

where the first constraint is to preserve the symmetric DoF allocation as  $1/r$ . However, problem (5) is still a highly intractable mixed combinatorial optimization problem with a non-convex fixed-rank constraint and a combinatorial objective function.

To enable the capability of polynomial-time complexity algorithm design, we further reveal the sparsity structure in problem (5) for user admission control. Notice that the sparsity of the vector consisting of the diagonal entries in  $\mathbf{X}$  equals the number of admitted users  $|\mathcal{S}|$ , i.e.,

$$\|\text{diag}(\mathbf{X})\|_0 = |\mathcal{S}|, \quad (6)$$

where  $\text{diag}(\cdot)$  extracts the diagonal of a matrix and  $\|\cdot\|_0$  is the  $\ell_0$ -norm of a vector, i.e., the count of non-zero entries. Problem (5) can be further reformulated as the following sparse and low-rank optimization problem, i.e.,

$$\begin{aligned} \mathcal{P} : & \underset{\mathbf{X} \in \mathbb{R}^{K \times K}}{\text{maximize}} && \|\text{diag}(\mathbf{X})\|_0 \\ & \text{subject to} && \text{rank}(\mathbf{X}) = r, \\ & && X_{ij} = 0, \quad \forall i \neq j, (i, j) \in \mathcal{V}. \end{aligned} \quad (7)$$

The equivalence means that if  $\mathbf{X}^*$  is a solution to problem (7), then  $\{\mathbf{X}^*, \mathcal{S}^*\}$  with  $\mathcal{S}^* = \{i : X_{ii}^* \neq 0\}$  is a solution to problem (5), and vice versa. As we are particularly interested in the scenario with high DoF allocation requirements, where the interference alignment is infeasible, user admission control is thus critical. Therefore, low-rankness in matrix  $\mathbf{X}$  is from the high DoF requirements constraint (i.e.,  $\text{rank}(\mathbf{X}) = r < K$ ), and sparsity in the diagonal entries in matrix  $\mathbf{X}$  is from user admission control (i.e.,  $|\mathcal{S} = \{i : X_{ii} \neq 0\}| < K$ ). Notice that we only need to consider problem  $\mathcal{P}$  in the real field without losing any performance in terms of admitted users. The reason is that the affine constraint in (7) is restricted in real field and the diagonal entries of matrix  $\mathbf{X}$  can be further restricted to the real field while achieving the same value of  $\|\text{diag}(\mathbf{X})\|_0$  in the complex field. Furthermore, problem  $\mathcal{P}$  is always feasible, as  $\mathbf{X} = \mathbf{I}_r$  with  $\text{rank}(\mathbf{X}) = r$  and  $\mathbf{I}_r$  as the diagonal matrix with only  $r$  entries being one, is one trivial solution.

Sparse optimization has shown to be powerful for the user admission problems [20]–[23] via  $\ell_0$ -norm minimization using the sum-of-infeasibilities convex relaxation heuristic in optimization theory [24, Sec. 11.4]. In particular, to maximize the number of admitted users is equivalent to minimize the number of violated inequalities for the quality-of-service (QoS) constraints. Although problem  $\mathcal{P}$  adopts the same philosophy of  $\ell_0$ -norm to count the number of admitted users (6), it reveals unique challenges due to  $\ell_0$ -norm maximization and non-convex fixed-rank constraint. However, compared with the original formulation (5), the sparse and low-rank optimization formulation (7) holds algorithmic advantages, which are demonstrated in the sequel via the Riemannian optimization approach [33].

*Remark 1:* In [19], the low-rank matrix completion approach was proposed to maximize the achievable DoF, i.e., minimize the rank of matrix  $\mathbf{X}$ , given any network topology. In particular, a Riemannian pursuit approach was presented to solve the low-rank optimization problem by alternatively performing rank increase and solving the fixed-rank least-square problems, which can be solved via Riemannian optimization [33]. Note that the fixed-rank least-square problem has a convex objective function and a non-convex fixed-rank constraint. In our presented user admission control problem  $\mathcal{P}$ , however, we need to maximize a non-convex objective function, i.e.,  $\ell_0$ -norm, with a non-convex fixed-rank constraint. The coupled challenges with both the non-convex  $\ell_0$ -norm maximization objective and the non-convex fixed-rank constraint motivate us to propose novel non-convex regularized smoothed  $\ell_1$ -minimization approach to solve problem  $\mathcal{P}$  in Section IV, thereby applying the matrix manifold optimization technique [33]. Note that, with  $\ell_0$ -norm maximization as the objective, a simple  $\ell_1$ -norm relaxation still yields a non-convex and unbounded objective, as maximizing a convex function is non-convex.

### C. Problem Analysis

In this subsection, we reveal the unique challenges of solving the sparse and low-rank optimization problem  $\mathcal{P}$  for user admission control in topological interference management.

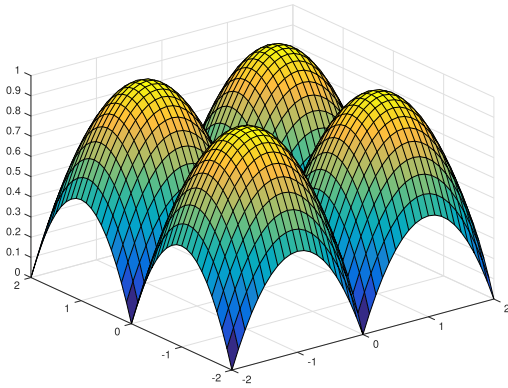


Fig. 2. The non-convex regularized sparsity inducing norm  $f(\mathbf{z}) = \|\mathbf{z}\|_1 - \lambda \|\mathbf{z}\|_2^2$  ( $\lambda = 0.5$ ) with bounded values in  $\mathbf{z} \in \mathbb{R}^2$ .

1) *Non-Convex Objective Function*: Although  $\ell_1$ -norm serves the convex surrogate for the non-convex  $\ell_0$ -norm [24], [25], it is inapplicable in problem  $\mathcal{P}$  for  $\ell_0$ -norm maximization, as it yields unbounded values, as well as non-convexity for  $\ell_1$ -norm maximization. To aid efficient algorithms design, we propose a novel non-convex regularized  $\ell_1$ -norm to induce sparsity with bounded values. This is achieved by adding a *negative* quadratic term in the  $\ell_1$ -norm as follows:

$$f(\mathbf{z}) = \|\mathbf{z}\|_1 - \lambda \|\mathbf{z}\|_2^2, \quad (8)$$

where  $\mathbf{z} \in \mathbb{R}^n$  and  $\lambda \geq 0$  is a weighting parameter. A typical example with  $f(\mathbf{z}) = \|\mathbf{z}\|_1 - 0.5 \|\mathbf{z}\|_2^2$  and  $\mathbf{z} \in \mathbb{R}^2$  is illustrated in Fig. 2, which upper bounds all the diagonal values by 1.

*Remark 2*: Although regularized  $\ell_1$ -norm approach has recently been intensively investigated in computational high-dimensional statistics [41], e.g., Lasso estimator, where the corresponding  $\ell_1$ -norm minimization problem normally is convex. However, a simple  $\ell_1$ -norm relaxation yields the objective in problem  $\mathcal{P}$  unbounded and non-convex, due to  $\ell_1$ -norm maximization. We thus propose to add a negative quadratic term in (8) to bound the objective value. Note that optimizing the proposed quadratic regularized  $\ell_1$ -norm  $f(\mathbf{z})$  becomes non-convex, as the sparsity inducing function  $f(\mathbf{z})$  is non-convex. In summary, the proposed non-convex quadratic regularized sparsity inducing norm (8) serves the purpose of inducing the sparsity and bounding the objective value.

2) *Non-Convex Fixed-Rank Constraint*: Matrix factorization serves a powerful way to address the non-convexity of the fixed-rank matrices. One popular way is to factorize a fixed rank- $r$  matrix  $\mathbf{X}$  (in real field) as  $\mathbf{UV}^T$  with  $\mathbf{U} \in \mathbb{R}^{K \times r}$  and  $\mathbf{V} \in \mathbb{R}^{K \times r}$ , followed by alternatively optimizing over  $\mathbf{U}$  and  $\mathbf{V}$  holding the other fixed [27], [36]. However, due to the non-convex objective function in problem  $\mathcal{P}$ , the resulting optimization problem over  $\mathbf{U}$  or  $\mathbf{V}$  is still non-convex. Furthermore, such factorization is not unique as  $\mathbf{X}$  remains unchanged under the transformation of the factors

$$(\mathbf{U}, \mathbf{V}) \mapsto (\mathbf{UM}^{-1}, \mathbf{VM}^T),$$

for all non-singular matrices  $\mathbf{M}$  of size  $r \times r$ . As a result, the critical points of an objective function parameterized with



Fig. 3. The proposed three-stage Riemannian framework for user admission control in topological interference alignment via sparse and low-rank optimization.  $\mathbf{z}^* \in \mathbb{R}^K$  is the induced sparsity pattern for user selection and  $\mathcal{S}^* \subseteq \{1, \dots, K\}$  is set of admitted users.

$\mathbf{U}$  and  $\mathbf{V}$  are not isolated on  $\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ . This profoundly affects the performance of second-order optimization algorithms which require non degenerate critical points, which is no longer the case here. We propose to address this issue by exploiting the *quotient manifold geometry* of the set of fixed-rank matrices [42]. The resulting non-convex optimization problem is further solved by exploiting the Riemannian optimization framework which provides systematic ways to develop algorithms on quotient manifolds [33].

In summary, in this paper, we propose a new regularized smooth formulation to induce the sparsity in the solution to problem  $\mathcal{P}$ , followed by the Riemannian optimization approach via exploiting the quotient manifold geometry of fixed-rank matrices. The induced sparsity pattern guides user selection for user admission control.

#### IV. REGULARIZED SMOOTHED $\ell_1$ -MINIMIZATION FOR SPARSE AND LOW-RANK OPTIMIZATION VIA RIEMANNIAN OPTIMIZATION

In this section, we present a Riemannian framework for sparse and low-rank optimization problem  $\mathcal{P}$  via regularized smoothed  $\ell_1$ -minimization by exploiting the quotient manifold geometry of fixed-rank matrices. The induced sparsity solution to problem  $\mathcal{P}$  provides guideline for user admission control, supported by a user selection procedure. In the final stage, a low-rank matrix completion approach with Riemannian optimization is adopted to design the linear topological interference management strategy. The proposed three-stage Riemannian framework for user admission control in topological interference management is presented in Fig. 3.

##### A. Stage One: Regularized Smoothed $\ell_1$ -Minimization for Sparsity Inducing

In order to make problem  $\mathcal{P}$  (7) numerically tractable, we relax the non-convex  $\ell_0$ -norm objective function to its convex surrogate  $\ell_1$ -norm, resulting in the following optimization problem:

$$\begin{aligned} & \underset{\mathbf{X} \in \mathbb{R}^{K \times K}}{\text{maximize}} && \|\text{diag}(\mathbf{X})\|_1 \\ & \text{subject to} && \text{rank}(\mathbf{X}) = r, \\ & && X_{ij} = 0, \quad \forall i \neq j, (i, j) \in \mathcal{V}. \end{aligned} \quad (9)$$

Although the  $\ell_1$ -norm is tractable, it is *unbounded* from above due to  $\ell_1$ -norm maximization, which makes problem (9) ill-posed. Note that maximizing a convex  $\ell_1$ -norm is still non-convex.

To circumvent the unboundness issue, we add the quadratic term  $-\lambda \|\text{diag}(\mathbf{X})\|_2^2$  to the objective function in problem (9),

where  $\lambda \geq 0$  is a weighting parameter that bounds the overall objective function from above leading to the formulation

$$\begin{aligned} & \underset{X \in \mathbb{R}^{K \times K}}{\text{maximize}} \quad \|\text{diag}(X)\|_1 - \lambda \|\text{diag}(X)\|_2^2 \\ & \text{subject to} \quad \text{rank}(X) = r, \\ & \quad X_{ij} = 0, \quad \forall i \neq j, (i, j) \in \mathcal{V}. \end{aligned} \quad (10)$$

For example, if  $\lambda = 0.5$ , then the diagonal values of  $X$  are upper bounded by 1. It should be emphasized that the role of  $\lambda$  in (10) is to upper bound the objective function and it does not affect the sparsity pattern that is expected from (9). This is further confirmed in Section IV-D via simulations. Additionally, if  $X^*$  is the solution to (7), then  $\alpha X^*$  is also a solution of (7) for all non-zero scalar  $\alpha$ . Equivalently, there exists continuum of solutions, which is effectively resolved by the objective function in (10).

Although problem (10) is still non-convex due to the non-convex objective and non-convex fixed-rank constraint, it has the algorithmic advantage that it can be solved efficiently (i.e., numerically) in the framework of Riemannian optimization [33].

*Riemannian Optimization for Fixed-Rank Optimization:* In this subsection, we propose a Riemannian optimization algorithm to solve the non-convex optimization problem (10), which is equivalent to

$$\begin{aligned} & \underset{X \in \mathbb{R}^{K \times K}}{\text{minimize}} \quad -\|\text{diag}(X)\|_1 + \lambda \|\text{diag}(X)\|_2^2 \\ & \text{subject to} \quad \text{rank}(X) = r, \\ & \quad X_{ij} = 0, \quad \forall i \neq j, (i, j) \in \mathcal{V}. \end{aligned} \quad (11)$$

However, the intersection of rank constraint and the affine constraint is challenging to characterize. We, therefore, propose to solve problem (11) via a *regularized* version as follows:

$$\begin{aligned} \mathcal{P}_{\text{RS}} : \underset{X \in \mathbb{C}^{K \times K}}{\text{minimize}} \quad & \frac{1}{2} \underbrace{\sum_{i \neq j, (i, j) \in \mathcal{V}} X_{ij}^2}_{\text{network topology}} \\ & + \rho \underbrace{\sum_{i=1}^K (\lambda X_{ii}^2 - (X_{ii}^2 + \epsilon^2)^{1/2})}_{\text{admission}} \\ & \text{subject to} \quad \text{rank}(X) = r, \end{aligned} \quad (12)$$

where  $\rho \geq 0$  is the regularization parameter and  $\epsilon$  is the parameter that approximates  $|X_{ii}|$  with the smooth term  $(X_{ii}^2 + \epsilon^2)^{1/2}$  that makes the objective function *differentiable*. A very small  $\epsilon$  leads to ill-conditioning of the objective function in (12). Since we intend to obtain the sparsity pattern of the optimal  $X$ , we set  $\epsilon$  to a high value, e.g., 0.01, to make problem (12) well conditioned. Problem  $\mathcal{P}_{\text{RS}}$  is an optimization problem over the set of fixed-rank matrices and can be solved via a Riemannian trust-region algorithm [33].

*Remark 3:* Although Riemannian optimization turns out to be effective for solving rank-constrained optimization, it normally requires that the objective function is smooth and the constraint set is a Riemannian manifold. Unfortunately, to utilize Riemannian optimization technique to solve problem (11), unique challenges arise. The reason is that the objective

function in problem (11) is nonsmooth and the constraint in problem (11) is not a Riemannian manifold due to the affine constraint. We thus contribute to smoothing the objective value by approximating  $|X_{ii}|$  with the smooth term  $(X_{ii}^2 + \epsilon^2)^{1/2}$ , followed by regularizing the affine constraint into the objective function, resulting a Riemannian manifold constraint with the fixed-rank constraint.

### B. Stage Two: Finding Sparsity Pattern for User Admission Control

Let  $X^*$  be the solution to the regularized smoothed  $\ell_1$ -minimization problem  $\mathcal{P}_{\text{RS}}$ . We order the diagonal entries of matrix  $X^*$ , i.e., the vector  $z^* = \text{diag}(X^*) \in \mathbb{R}^K$ , in the descending order:  $|z_{\pi_1}| \geq |z_{\pi_2}| \geq \dots \geq |z_{\pi_K}|$ . Intuitively, the large coefficient  $z_i$  indicates that the corresponding users are allocated high desired signal power. Therefore, the user with larger coefficient  $z_i$ 's has a higher priority to be admitted with higher desired signals. We adopt the bi-section search procedure to find the maximum number of admitted users. Specifically, let  $N_0$  be the maximum number of users that can be admitted while satisfying the interference alignment conditions. To determine the value of  $N_0$ , a sequence of the following size-reduced topological interference management feasibility problem needs to be solved,

$$\begin{aligned} \mathcal{F}(S^{[m]}) : \text{find} \quad & X(S^{[m]}) \\ & \text{subject to} \quad \text{rank}(X(S^{[m]})) = r, \\ & \quad X_{ii} = 1, \quad \forall i \in S^{[m]}, \\ & \quad X_{ij} = 0, \quad \forall i \neq j, i, j \in S^{[m]}, (i, j) \in \mathcal{V}, \end{aligned} \quad (13)$$

where  $S^{[m]} = \{\pi_1, \dots, \pi_m\}$ .

To check the feasibility, we rewrite problem (13) as follows:

$$\begin{aligned} & \text{minimize} \quad \|\mathcal{P}_{\Omega}(X(S^{[m]})) - \mathbf{I}_{|S^{[m]|}}\|_F^2 \\ & \text{subject to} \quad \text{rank}(X(S^{[m]})) = r, \end{aligned} \quad (14)$$

where  $\Omega = \{(i, j) | i, j \in S^{[m]}, (i, j) \in \mathcal{V}\}$  and  $\mathcal{P}_{\Omega}(Y) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  is the orthogonal projection operator onto the subspace of matrices which vanish outside  $\Omega$  such that the  $(i, j)$ -th component of  $\mathcal{P}_{\Omega}(Y)$  equals to  $Y_{ij}$  if  $(i, j) \in \Omega$  and zero otherwise. If the objective value approaches to zero, we say that the set of users  $S^{[m]}$  can be admitted. Problem (14) can be solved by Riemannian trust-region algorithms [32] via Manopt [43]. Note that, theoretically, the Riemannian algorithm can only guarantee convergence to a first-order critical point, but empirically, we observe convergence to critical points that are local minima.

### C. Stage Three: Low-Rank Matrix Completion for Topological Interference Management

Let  $S^* = \{\pi_1, \dots, \pi_{N_0}\}$  be the admitted users. We need to solve the following sized-reduced rank-constrained *matrix completion* problem:

$$\begin{aligned} \mathcal{P}_{\text{LRMC}}(S^*) : \text{minimize} \quad & \|\mathcal{P}_{\Omega}(X(S^*)) - \mathbf{I}_{|S^*|}\|_F^2 \\ & \text{subject to} \quad \text{rank}(X(S^*)) = r, \end{aligned} \quad (15)$$

to find the precoding vectors  $\mathbf{v}_i$ 's and decoding vectors  $\mathbf{u}_i$ 's for the admitted users in  $\mathcal{S}^*$ . Specifically, let  $\mathbf{X}^*$  be the solution of problem  $\mathcal{P}_{\text{LRMC}}(\mathcal{S}^*)$ . The precoding vectors  $\mathbf{v}_i$ 's and decoding vectors  $\mathbf{u}_i$ 's can be extracted by QR decomposition for matrix  $\mathbf{X}^* = [\mathbf{u}_i^H \mathbf{v}_j]$  using the Gram-Schmidt process.

Therefore, the proposed three-stage Riemannian optimization based user admission control algorithm is presented in Algorithm 1.

---

**Algorithm 1** User Admission Control for Topological Interference Management via Riemannian Optimization

---

**Step 0:** Solve the sparse inducing optimization problem  $\mathcal{P}_{\text{RS}}$  (12) using the Riemannian trust-region algorithm in Section V. Obtain the solution  $\mathbf{X}^*$  and sort the diagonal entries in the descending order:  $|z_{\pi_1}| \geq \dots \geq |z_{\pi_K}|$ , **go to Step 1.**

**Step 1:** Initialize  $N_{\text{low}} = 0$ ,  $N_{\text{up}} = K$ ,  $i = 0$ .

**Step 2:** Repeat

- 1) Set  $i \leftarrow \lfloor \frac{N_{\text{low}} + N_{\text{up}}}{2} \rfloor$ .

- 2) Solve problem  $\mathcal{F}(S^{[i]})$  (13) via (14) using the Riemannian trust-region algorithm in Section V: if it is feasible, set  $N_{\text{low}} = i$ ; otherwise, set  $N_{\text{up}} = i$ .

**Step 3:** Until  $N_{\text{up}} - N_{\text{low}} = 1$ , obtain  $N_0 = N_{\text{up}}$  and obtain the admitted users set  $\mathcal{S}^* = \{\pi_1, \dots, \pi_{N_0}\}$ .

**Step 4:** Solve problem  $\mathcal{P}_{\text{LRMC}}(\mathcal{S}^*)$  (15) to obtain the precoding and decoding vectors for the admitted users.

**End**

---

#### D. The Framework of Fixed-Rank Riemannian Manifold Optimization

The optimization problems (12), (14), and (15) are *least-square* optimization problems with fixed rank constraint. A rank- $r$  matrix  $\mathbf{X} \in \mathbb{R}^{K \times K}$  is parameterized as  $\mathbf{X} = \mathbf{U}\mathbf{V}^T$ , where  $\mathbf{U} \in \mathbb{R}^{K \times r}$  and  $\mathbf{V} \in \mathbb{R}^{K \times r}$  are full column-rank matrices. Such a factorization, however, is not unique as  $\mathbf{X}$  remains unchanged under the transformation of the factors

$$(\mathbf{U}, \mathbf{V}) \mapsto (\mathbf{U}\mathbf{M}^{-1}, \mathbf{V}\mathbf{M}^T), \quad (16)$$

for all non-singular matrices  $\mathbf{M} \in \text{GL}(r)$ , the set of  $r \times r$  non-singular matrices. Equivalently,  $\mathbf{X} = \mathbf{U}\mathbf{V}^T = \mathbf{U}\mathbf{M}^{-1}(\mathbf{V}\mathbf{M}^T)^T$  for all non-singular matrices  $\mathbf{M}$ . As a result, the critical points of an objective function parameterized with  $\mathbf{U}$  and  $\mathbf{V}$  are *not isolated* on  $\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ .

The classical remedy to remove this indeterminacy requires further (triangular-like) structure in the factors  $\mathbf{U}$  and  $\mathbf{V}$ . For example, LU decomposition is a way forward. In contrast, we encode the invariance map (16) in an abstract search space by optimizing directly over a set of equivalence classes

$$[(\mathbf{U}, \mathbf{V})] := \{(\mathbf{U}\mathbf{M}^{-1}, \mathbf{V}\mathbf{M}^T) : \mathbf{M} \in \text{GL}(r)\}. \quad (17)$$

The set of equivalence classes is termed as the *quotient space* and is denoted by

$$\mathcal{M}_r := \mathcal{M}/\text{GL}(r), \quad (18)$$

where the total space  $\mathcal{M}$  is the product space  $\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ .

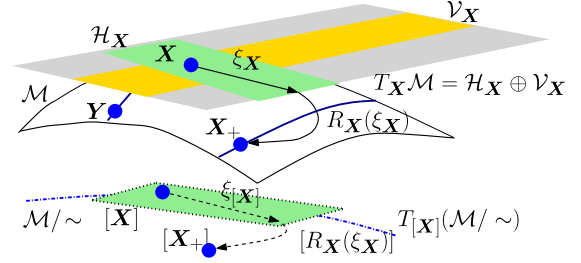


Fig. 4. Optimization on a quotient manifold. The dotted lines represent abstract objects and the solid lines are their matrix representations. The points  $\mathbf{X}$  and  $\mathbf{Y}$  in the total (computational) space  $\mathcal{M}$  belong to the same equivalence class (shown in solid blue color) and they represent a single point  $[\mathbf{X}] := \{\mathbf{Y} \in \mathcal{M} : \mathbf{Y} \sim \mathbf{X}\}$  in the quotient space  $\mathcal{M}/\sim$ . An algorithm by necessity is implemented in the computation space, but conceptually, the search is on the quotient manifold. Given a search direction  $\xi_{\mathbf{X}}$  at  $\mathbf{X}$ , the updated point on  $\mathcal{M}$  is given by the retraction mapping  $R_{\mathbf{X}}$ .

Consequently, if an element  $\mathbf{X} \in \mathcal{M}$  has the matrix characterization  $(\mathbf{U}, \mathbf{V})$ , then (12), (14), and (15) are of the form

$$\underset{[\mathbf{X}] \in \mathcal{M}_r}{\text{minimize}} f([\mathbf{X}]), \quad (19)$$

where  $[\mathbf{X}] = [(\mathbf{U}, \mathbf{V})]$  is defined in (17) and  $f : \mathcal{M} \rightarrow \mathbb{R} : \mathbf{X} \mapsto f(\mathbf{X})$  is a *smooth* function on  $\mathcal{M}$ , but now induced (with slight abuse of notation) on the quotient space  $\mathcal{M}_r$  (18).

The quotient space  $\mathcal{M}_r$  has the structure of a smooth *Riemannian* quotient manifold of  $\mathcal{M}$  by  $\text{GL}(r)$  [42]. The Riemannian structure conceptually transforms a rank-constrained optimization problem into an *unconstrained* optimization problem over the non-linear manifold  $\mathcal{M}_r$ . Additionally, it allows to compute objects like gradient (of an objective function) and develop a Riemannian trust-region algorithm on  $\mathcal{M}_r$  that uses second-order information for faster convergence [33].

## V. MATRIX OPTIMIZATION ON QUOTIENT MANIFOLDS

In this section, we present the Riemannian trust-region algorithm for the smooth optimization over manifolds by exploiting symmetry in the search space of fixed-rank matrices.

### A. Problem Structures

We exploit the symmetry in the fixed-rank constraint to design efficient Riemannian optimization algorithm.

1) *Quotient Manifold:* Consider an equivalence relation  $\sim$  in the *total* (computational) space  $\mathcal{M}$ . The quotient manifold  $\mathcal{M}/\sim$  generated by this equivalence property consists of elements that are *equivalence classes* of the form  $[\mathbf{X}] = \{\mathbf{Y} \in \mathcal{M} : \mathbf{Y} \sim \mathbf{X}\}$ . Equivalently, if  $[\mathbf{X}]$  is an element in  $\mathcal{M}/\sim$ , then its matrix representation in  $\mathcal{M}$  is  $\mathbf{X}$ . In the context of fixed-rank constraint,  $\mathcal{M}/\sim$  is identified with  $\mathcal{M}_r$ , i.e., the fixed-rank manifold. Fig. 4 shows a schematic viewpoint of optimization on a quotient manifold. In particular, we need the notion of “linearization” of the search space, “search” direction, and a way “move” on a manifold. Below we show the concrete development of these objects that allow to do develop a second-order trust-region algorithm on manifolds.

Since the manifold  $\mathcal{M}/\sim$  is an abstract space, the elements of its tangent space  $T_{[\mathbf{X}]}(\mathcal{M}/\sim)$  at  $[\mathbf{X}]$  also call for a matrix



representation in the tangent space  $T_X\mathcal{M}$  that respects the equivalence relation  $\sim$ . Equivalently, the matrix representation of  $T_{[X]}(\mathcal{M}/\sim)$  should be restricted to the directions in the tangent space  $T_X\mathcal{M}$  on the total space  $\mathcal{M}$  at  $X$  that do not induce a displacement along the equivalence class  $[X]$ . This is realized by decomposing  $T_X\mathcal{M}$  into complementary subspaces, the *vertical* and *horizontal* subspaces such that  $\mathcal{V}_X \oplus \mathcal{H}_X = T_X\mathcal{M}$ . The vertical space  $\mathcal{V}_X$  is the tangent space of the equivalence class  $[X]$ . On the other hand, the horizontal space  $\mathcal{H}_X$ , which is any complementary subspace to  $\mathcal{V}_X$  in  $T_X\mathcal{M}$ , provides a valid matrix representation of the abstract tangent space  $T_{[X]}(\mathcal{M}/\sim)$  [33, Sec. 3.5.8]. An abstract tangent vector  $\zeta_{[X]} \in T_{[X]}(\mathcal{M}/\sim)$  at  $[X]$  has a unique element in the horizontal space  $\zeta_X \in \mathcal{H}_X$  that is called its *horizontal lift*. Our specific choice of the horizontal space is the subspace of  $T_X\mathcal{M}$  that is the *orthogonal complement* of  $\mathcal{V}_X$  in the sense of a Riemannian metric (an inner product).

A Riemannian metric or an inner product  $g_X : T_X\mathcal{M} \times T_X\mathcal{M} \rightarrow \mathbb{R}$  at  $X \in \mathcal{M}$  in the total space defines a Riemannian metric  $g_{[X]} : T_{[X]}(\mathcal{M}/\sim) \times T_{[X]}(\mathcal{M}/\sim) \rightarrow \mathbb{R}$ , i.e.,

$$g_{[X]}(\zeta_{[X]}, \eta_{[X]}) := g_X(\zeta_X, \eta_X), \quad (20)$$

on the quotient manifold  $\mathcal{M}/\sim$ , provided that the expression  $g_X(\zeta_X, \eta_X)$  does not depend on a specific representation along the equivalence class  $[X]$ . Here  $\zeta_{[X]}$  and  $\eta_{[X]}$  are tangent vectors in  $T_{[X]}(\mathcal{M}/\sim)$ , and  $\zeta_X, \eta_X$  are their horizontal lifts in  $\mathcal{H}_X$  at  $X$ . Equivalently, if  $Y$  is another element that belongs to  $[X]$  and  $\zeta_Y$  and  $\eta_Y$  are the horizontal lifts of  $\zeta_{[X]}$  and  $\eta_{[X]}$  at  $Y$ , then the metric in (20) obeys the equality  $g_X(\zeta_X, \eta_X) = g_Y(\zeta_Y, \eta_Y)$ . Such a metric is then said to be *invariant* to the equivalence relation  $\sim$ .

2) *Riemannian Metric*: In the context of fixed-rank matrices, there exist metrics which are invariant. A particular invariant Riemannian metric on the total space  $\mathcal{M}$  that takes into account the symmetry (16) imposed by the factorization model and that is well suited to a least-squares objective function [44] is

$$g_X(\zeta_X, \eta_X) = \text{Tr}((V^T V)\zeta_U^T \eta_U) + \text{Tr}((U^T U)\zeta_V^T \eta_V), \quad (21)$$

where  $X = (U, V)$  and  $\zeta_X, \eta_X \in T_X\mathcal{M}$ . It should be noted that the tangent space  $T_X\mathcal{M}$  has the matrix characterization  $\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ , i.e.,  $\eta_X$  (and similarly  $\zeta_X$ ) has the matrix representation  $(\eta_U, \eta_V) \in \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ .

To show that (21) is invariant to the transformation (16), we assume that another element  $Y \in [X]$  has matrix representation  $(UM^{-1}, VM)$  for a non singular square matrix  $M$ . Similarly, we assume that the tangent vector  $\eta_Y$  (similarly  $\zeta_Y$ ) has matrix representation  $(\eta_{UM^{-1}}, \eta_{VM^T}) \in \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$ . If  $\eta_X$  and  $\eta_Y$  (similarly for  $\zeta_X$  and  $\zeta_Y$ ) are the *horizontal lifts* of  $\eta_{[X]}$  at  $X$  and  $Y$ , respectively. Then, we have  $\eta_{UM^{-1}} = \eta_U M^{-1}$  and  $\eta_{VM^T} = \eta_V M^T$  [33, Example 3.5.4]. Similar results can be obtained for  $\zeta_Y$ . A few computations then show that  $g_X(\zeta_X, \eta_X) = g_Y(\zeta_Y, \eta_Y)$ , which implies that the metric (21) is invariant to the transformation (16) along the equivalence class  $[X]$ . This implies that we have a unique metric on the quotient space  $\mathcal{M}/\sim$ .

Motivation for the metric (21) comes from the fact that it is induced from a *block diagonal* approximation of the

Hessian of a simpler cost function  $\|UV^T - I\|_F^2$ , which is strictly convex in  $U$  and  $V$  individually. This block diagonal approximation ensures that the cost of computing (21) depends linearly on  $K$  and the metric is well suited for least-squares problems. Similar ideas have also been exploited in [19], [45], and [46] which show robust performance of Riemannian algorithms for various least-squares problems.

### B. Matrix Representation for the Quotient Manifolds

Once the metric (21) is defined on  $\mathcal{M}$ , the development of the geometric objects required for second-order optimization follow [33], [44]. The matrix characterizations of the tangent space  $T_X\mathcal{M}$ , vertical space  $\mathcal{V}_X$ , and horizontal space  $\mathcal{H}_X$  are straightforward with the expressions:

$$\begin{aligned} T_X\mathcal{M} &= \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}, \\ \mathcal{V}_X &= \{(-U\Lambda, V\Lambda^T) : \Lambda \in \mathbb{R}^{r \times r}\}, \\ \mathcal{H}_X &= \{(\zeta_U, \zeta_V) : U^T \zeta_U V^T V = U^T U \zeta_V^T V\}, \end{aligned} \quad (22)$$

where  $\zeta_U, \zeta_V \in \mathbb{R}^{K \times r}$ .

Apart from the characterization of the horizontal space, we need a linear mapping  $\Pi_X : T_X\mathcal{M} \mapsto \mathcal{H}_X$  that projects vectors from the tangent space onto the horizontal space. Projecting an element  $\eta_X \in T_X\mathcal{M}$  onto the horizontal space is accomplished with the operator

$$\Pi_X(\eta_X) = (\eta_U + U\Lambda, \eta_V - V\Lambda^T), \quad (23)$$

where  $\Lambda \in \mathbb{R}^{r \times r}$  is uniquely obtained by ensuring that  $\Pi_X(\eta_X)$  belongs to the horizontal space characterized in (22). Finally, the expression of  $\Lambda$  is given by

$$\begin{aligned} &U^T (\eta_U + U\Lambda) V^T V \\ &= U^T U (\eta_V - V\Lambda^T)^T V \\ &\Rightarrow \Lambda = 0.5(\eta_V^T V (V^T V)^{-1} - (U^T U)^{-1} U^T \eta_U). \end{aligned}$$

1) *Gradient and Hessian Computations*: The choice of the metric (21) and of the horizontal space (as the orthogonal complement of  $\mathcal{V}_X$ ) turns the quotient manifold  $\mathcal{M}/\sim$  into a *Riemannian submersion* of  $(\mathcal{M}, g)$  [33, Sec. 3.6.2]. This special construction allows for a convenient matrix representation of the gradient [33, Sec. 3.6.2] and the Hessian [33, Proposition 5.3.3] on the quotient manifold  $\mathcal{M}/\sim$ . Below we show the gradient and Hessian computations for the problem (19).

The Riemannian gradient  $\text{grad}_{[X]}f$  of  $f$  on  $\mathcal{M}/\sim$  is uniquely represented by its horizontal lift in  $\mathcal{M}$  which has the matrix representation

$$\begin{aligned} \text{horizontal lift of } \text{grad}_{[X]}f &= \text{grad}_X f \\ &= \left(\frac{\partial f}{\partial U} (V^T V)^{-1}, \frac{\partial f}{\partial V} (U^T U)^{-1}\right), \end{aligned} \quad (24)$$

where  $\text{grad}_X f$  is the gradient of  $f$  in  $\mathcal{M}$  and  $\partial f/\partial U$  and  $\partial f/\partial V$  are the *partial derivatives* of  $f$  with respect to  $U$  and  $V$ , respectively.

In addition to the Riemannian gradient computation (24), we also require the directional derivative of the gradient along a search direction. This is captured by a *connection*  $\nabla_{\zeta_X} \eta_X$ ,

which is the *covariant derivative* of vector field  $\eta_X$  with respect to the vector field  $\zeta_X$ . The Riemannian connection  $\nabla_{\zeta_{[X]}}\eta_{[X]}$  on the quotient manifold  $\mathcal{M}/\sim$  is uniquely represented in terms of the Riemannian connection  $\nabla_{\zeta_X}\eta_X$  in the total space  $\mathcal{M}$  [33, Proposition 5.3.3] which is

$$\text{horizontal lift of } \nabla_{\zeta_{[X]}}\eta_{[X]} = \Pi_X(\nabla_{\zeta_X}\eta_X), \quad (25)$$

where  $\zeta_{[X]}$  and  $\eta_{[X]}$  are vector fields in  $\mathcal{M}/\sim$  and  $\zeta_X$  and  $\eta_X$  are their horizontal lifts in  $\mathcal{M}$ . Here  $\Pi_X(\cdot)$  is the projection operator defined in (23). It now remains to find out the Riemannian connection in the total space  $\mathcal{M}$ . We find the matrix expression by invoking the *Koszul* formula [33, Th. 5.3.1]. After a routine calculation, the final expression is given by [44]

$$\nabla_{\zeta_X}\eta_X = D\eta_X[\zeta_X] + (A_U, A_V),$$

where

$$\begin{aligned} A_U &= \eta_U \text{Sym}(\zeta_V^T V)(V^T V)^{-1} + \zeta_U \text{Sym}(\eta_V^T V)(V^T V)^{-1} \\ &\quad - U \text{Sym}(\eta_V^T \zeta_V)(V^T V)^{-1} \\ A_V &= \eta_V \text{Sym}(\zeta_U^T U)(U^T U)^{-1} + \zeta_V \text{Sym}(\eta_U^T U)(U^T U)^{-1} \\ &\quad - V \text{Sym}(\eta_U^T \zeta_U)(U^T U)^{-1} \end{aligned} \quad (26)$$

and  $D^\zeta[\eta]$  is the Euclidean directional derivative  $D^\zeta[\eta] := \lim_{t \rightarrow 0} (\zeta_{X+t\eta_X} - \zeta_X)/t$ .  $\text{Sym}(\cdot)$  extracts the symmetric part of a square matrix, i.e.,  $\text{Sym}(Z) = (Z + Z^T)/2$ .

The directional derivative of the Riemannian gradient in the direction  $\zeta_{[X]}$  is given by the *Riemannian Hessian operator*  $\text{Hess}_{[X]}f[\zeta_{[X]}]$  which is now directly defined in terms of the Riemannian connection  $\nabla$ . Based on (25) and (26), the horizontal lift of the Riemannian Hessian in  $\mathcal{M}/\sim$  has the matrix expression:

$$\text{horizontal lift of } \text{Hess}_{[X]}f[\zeta_{[X]}] = \Pi_X(\nabla_{\zeta_X}\text{grad}_X f), \quad (27)$$

where  $\zeta_{[X]} \in T_{[X]}(\mathcal{M}/\sim)$  and its horizontal lift  $\zeta_X \in \mathcal{H}_X$ .  $\Pi_X(\cdot)$  is the projection operator defined in (23).

2) *Retraction*: An iterative optimization algorithm involves computing a search direction (e.g., negative gradient) and then “moving in that direction”. The default option on a Riemannian manifold is to move along geodesics, leading to the definition of the *exponential map*. Because the calculation of the exponential map can be computationally expensive, it is customary in the context of manifold optimization to relax the constraint of moving along geodesics. To this end, we define *retraction*  $R_X : \mathcal{H}_X \rightarrow \mathcal{M} : \zeta_X \mapsto R_X(\zeta_X)$  [33, Definition 4.1.1]. A natural update on the manifold  $\mathcal{M}$  is, therefore, based on the update formula  $X_+ = R_X(\zeta_X)$ , i.e., defined as

$$\begin{aligned} R_U(\zeta_U) &= U + \zeta_U \\ R_V(\zeta_V) &= V + \zeta_V, \end{aligned} \quad (28)$$

where  $\zeta_X = (\zeta_U, \zeta_V) \in \mathcal{H}_X$  is a search direction and  $X_+ \in \mathcal{M}$ . It translates into the update  $[X_+] = [R_X(\zeta_X)]$  on  $\mathcal{M}/\sim$ .

### C. Riemannian Trust-Region Algorithm

Analogous to trust-region algorithms in the Euclidean space [47, Ch. 4], trust-region algorithms on a Riemannian quotient manifold with guaranteed superlinear rate convergence and global convergence have been proposed in [33, Ch. 7]. At each iteration we solve the *trust-region sub-problem* on the quotient manifold  $\mathcal{M}/\sim$ . The trust-region sub-problem is formulated as the minimization of the *locally-quadratic* model of the objective function, say  $f : \mathcal{M} \rightarrow \mathbb{R}$  at  $X \in \mathcal{M}$ ,

$$\begin{aligned} &\underset{\zeta_X \in \mathcal{H}_X}{\text{minimize}} && g_X(\zeta_X, \text{grad}_X f) + \frac{1}{2}g_X(\zeta_X, \text{Hess}_X f[\zeta_X]) \\ &\text{subject to} && g_X(\zeta_X, \zeta_X) \leq \Delta^2, \end{aligned} \quad (29)$$

where  $\Delta$  is the trust-region radius,  $g_X$  is the Riemannian metric (21), and  $\text{grad}_X f$  and  $\text{Hess}_X f$  are the Riemannian gradient and Riemannian Hessian operations defined in (24) and (27), respectively.

Solving the above trust-region sub-problem (29) leads to a direction  $\zeta_X$  that minimizes the quadratic model. Depending on whether the decrease of the cost function is sufficient or not, the potential iterate is accepted or rejected. The concrete matrix characterizations of Riemannian gradient (24), Riemannian Hessian (27), projection operator (23), and retraction (28) allow to use an *off-the-shelf* trust-region implementation on manifolds, e.g., in Manopt [43], which implements [33, Algorithm 1] that solves the trust-region sub-problem inexactly at every iteration.

The Riemannian trust-region algorithm is *globally convergent*, i.e., it converges to a critical point starting from any random initialization. The rate of convergence analysis of the algorithm is in [33, Ch. 7]. Theoretically, the algorithm converges to a critical point, but often in practice the convergence is observed to a local minimum. Under certain regularity conditions, the trust-region algorithm shows a *superlinear* rate of convergence locally near a critical point. The recent work [31] also establishes *worst-case* global rates (i.e., number of iterations required to obtain a fixed accuracy) of convergence over manifolds, it converges to the second-order KKT points starting from any random initial points. In practice, however, we observe better rates.

In summary, the concrete manifold-related ingredients are shown in Table I, which are based on the developments in [44].

### D. Computational Complexity

The numerical complexity of the algorithm in Algorithm 1 depends fixed-rank Riemannian optimization algorithm for solving (12), (14), and (15) and sorting the diagonal entries of rank- $r$  matrix. The sorting operation depends linearly with  $K$  (and logarithmic factors of  $K$ ). The computational cost of the Riemannian algorithm depends on i) the computational cost of the computing the partial derivatives of the objective functions in (12), (14), and (15) and ii) the manifold-related operations. The computational cost of the manifold-related ingredients are shown below.

- 1) Computation of partial derivatives of the objective functions in (12), (14), and (15) with respect to  $U$  and  $V$ :  $O(|\mathcal{V}|r)$ .

TABLE I  
MANIFOLD-RELATED INGREDIENTS

	minimize $\mathbf{X} \in \mathcal{M}_r$ $f(\mathbf{X})$ with $\mathbf{X} = \mathbf{U}\mathbf{V}^T$
Matrix representation	$\mathbf{X} = (\mathbf{U}, \mathbf{V})$
Total space $\mathcal{M}$	$\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$
Group action	$(\mathbf{U}\mathbf{M}^{-1}, \mathbf{V}\mathbf{M}^T)$ , $\mathbf{M} \in \text{GL}(r)$
Quotient space $\mathcal{M}/\sim$	$\mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r} / \text{GL}(r)$
Vectors in the ambient space	$(\mathbf{Z}_U, \mathbf{Z}_V) \in \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$
Matrix representation of a tangent vector $\xi_{\mathbf{X}}$ in $T_{\mathbf{X}}\mathcal{M}$	$(\xi_U, \xi_V) \in \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r}$
Metric $g_{\mathbf{X}}(\xi_{\mathbf{X}}, \eta_{\mathbf{X}})$ for any $\xi_{\mathbf{X}}, \eta_{\mathbf{X}} \in T_{\mathbf{X}}\mathcal{M}$	$\text{Tr}((\mathbf{V}^T \mathbf{V}) \xi_U^T \eta_U) + \text{Tr}((\mathbf{U}^T \mathbf{U}) \xi_V^T \eta_V)$
Vertical tangent vectors in $\mathcal{V}_{\mathbf{X}}$	$\{(-\mathbf{U}\mathbf{A}, \mathbf{V}\mathbf{A}^T) : \mathbf{A} \in \mathbb{R}^{r \times r}\}$
Horizontal tangent vectors in $\mathcal{H}_{\mathbf{X}}$	$\{(\zeta_U, \zeta_V) \in \mathbb{R}^{K \times r} \times \mathbb{R}^{K \times r} : \mathbf{U}^T \zeta_U \mathbf{V}^T \mathbf{V} = \mathbf{U}^T \mathbf{U} \zeta_V^T \mathbf{V}\}$
Projection of a tangent vector $\eta_{\mathbf{X}} \in T_{\mathbf{X}}\mathcal{M}$ on the horizontal space $\mathcal{H}_{\mathbf{X}}$	$\Pi_{\mathbf{X}}(\eta_{\mathbf{X}}) = (\eta_U + \mathbf{U}\mathbf{A}, \eta_V - \mathbf{V}\mathbf{A}^T)$ , where $\mathbf{A} = 0.5(\eta_V^T \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} - (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \eta_U)$ .
Retraction of a horizontal vector $\xi_{\mathbf{X}}$ onto the manifold	$R_{\mathbf{X}}(\xi_{\mathbf{X}}) = (\mathbf{U} + \xi_U, \mathbf{V} + \xi_V)$
Matrix representation of the Riemannian gradient $\text{grad}_{\mathbf{X}} f$	$(\frac{\partial f}{\partial \mathbf{U}}(\mathbf{V}^T \mathbf{V})^{-1}, \frac{\partial f}{\partial \mathbf{V}}(\mathbf{U}^T \mathbf{U})^{-1})$ , where $\partial f / \partial \mathbf{U}$ and $\partial f / \partial \mathbf{V}$ are the partial derivatives of $f$ with respect to $\mathbf{U}$ and $\mathbf{V}$ , respectively.
Matrix representation of the Riemannian Hessian $\text{Hess}_{\mathbf{X}} f[\xi_{\mathbf{X}}]$ along a horizontal vector $\xi_{\mathbf{X}}$	$\Pi_{\mathbf{X}}(\nabla_{\xi_{\mathbf{X}}} \text{grad}_{\mathbf{X}} f)$ , where $\text{grad}_{\mathbf{X}} f$ has the representation shown above. The matrix representation of the Riemannian connection $\nabla_{\xi_{\mathbf{X}}} \eta_{\mathbf{X}}$ is shown in (26). Finally, the projection operator $\Pi_{\mathbf{X}}$ is defined in (23).

- 2) Computation of Riemannian gradient with the formula (24):  $O(Kr^2 + r^3)$ .
- 3) Computation of the projection operator (23):  $O(Kr^2 + r^3)$ .
- 4) Computation of retraction  $R_{\bar{\mathbf{x}}}$  in (28):  $O(Kr)$ .
- 5) Computation of Riemannian Hessian with the formulas (25), (26), and (27):  $O(r^3 + Kr^2)$ .

It is clear that all the manifold-related operations are of linear complexity in  $K$  and cubic in  $r$ . Overall, the cost per iteration of the proposed algorithm in Algorithm 1 is *linear* with  $|\mathcal{V}|$ .

## VI. SIMULATION RESULTS

In this section, we simulate our proposed Riemannian trust-region (RTR) algorithm for the user admission control problem  $\mathcal{P}$  in topological interference management. All simulations are performed in Matlab on a 2.4 GHz octa-core Intel Xeon E5-2630 v3 machine (2 processors) with 64 GB RAM. All the simulated algorithms are initialized randomly. The RTR algorithms for the rank-constrained optimization problems (12), (14) and (15) are implemented based on the manifold optimization toolbox Manopt [43]. Furthermore, the RTR algorithm is terminated when either the Frobenius norm of the Riemannian gradient is below  $10^{-8}$  or the number of iterations exceeds the maximal iteration number, which is set to be 300. We set  $\frac{1}{2} \|\mathcal{P}_{\Omega}(\mathbf{X}(s^{[m]})) - \mathbf{I}_{|s^{[m]|}}\|_F^2 = 10^{-6}$  in (14) as the certificate of satisfaction for the affine constraint  $\mathcal{P}_{\Omega}(\mathbf{X}(s^{[m]})) = \mathbf{I}_{|s^{[m]|}}$  for all the algorithms. In the proposed RTR algorithm, we choose a reasonable small  $\epsilon$  to achieve good approximation for  $\ell_1$ -norm as well as avoid numerical issues. Any value between 0 and 1 is a good candidate for parameter  $\lambda$  as indicated in Fig. 8. We choose a good value for parameter  $\rho$  via cross validation.

The proposed Riemannian trust-region algorithm are compared to the following algorithms:

- **Riemannian conjugate gradient (RCG) algorithm:** The RCG algorithm is adopted to solve the rank-constrained optimization problems (12), (14) and (15). Based on the

matrix representation of the Riemannian gradient in (24) and the retraction mapping operator in Section V-B.2, we implement the first-order RCG algorithm using the toolbox Manopt [43]. The RCG algorithm is terminated when either the Frobenius norm of the Riemannian gradient is below  $10^{-8}$ , i.e.,  $\|\text{grad}_{\mathbf{X}} f\| \leq 10^{-8}$ , or the number of iterations exceeds the maximal iteration number 1000.

- **Alternating minimization (AltMin) algorithm:** The AltMin algorithm [27] is used to solve the rank-constrained optimization problems (12), (14) and (15). This is achieved by fixing  $\mathbf{U}$  or  $\mathbf{V}$  alternatively and solving the resulting problems (maybe non-convex) using the gradient descent algorithm. For the gradient descent algorithm, backtracking line-search is used to determine the step size. The inner loop of the gradient descent algorithm will be terminated either when the relative difference of objective values between two consecutive iterations is less than  $10^{-8}$ , i.e.,  $|f(\mathbf{X}^{[j+1]}) - f(\mathbf{X}^{[j]})| / f(\mathbf{X}^{[j]}) < 10^{-6}$ , or the number of iterations exceeds 1000 (unless otherwise stated). The stopping criterion of the outer loop of AltMin algorithm is given as either when the relative difference of objective value between two consecutive iteration falls below  $10^{-6}$  or the number of iterations exceeds 50 (unless otherwise stated).

### A. Convergence Rates

Consider a 20-user partially connected interference channel with 169 interference links. Each cross link belongs to interference channel links  $\mathcal{V}$  with connectivity probability  $p = 0.4$ . We solve the fixed-rank optimization problem (14) with all the users admitted using difference algorithms. In this experiment, the maximal number of iteration is set to be 100 for both of the inner and outer loop in the AltMin algorithm. Fig. 5 demonstrates the convergence rates of all algorithms with rank  $r = 8$ . This figure shows that the RTR algorithm has the fast (quadratic) convergence rate while the RCG algorithm also converges faster than the AltMin algorithm by exploiting the second-order information in the

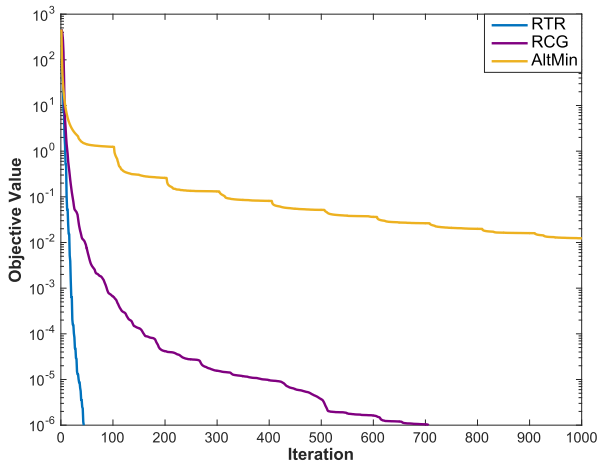


Fig. 5. Convergence rate results of different algorithms. All the algorithm are randomly initialized at the same point.

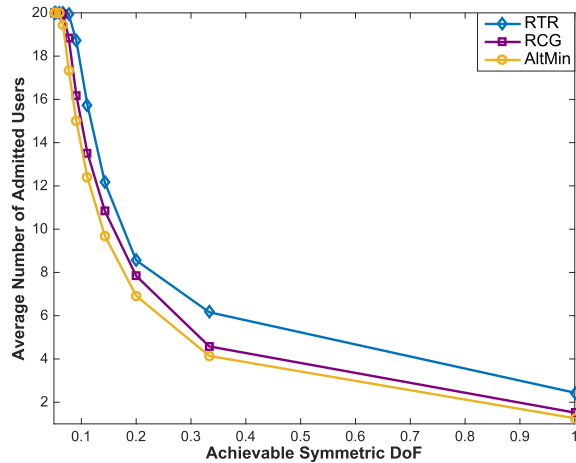


Fig. 6. The average number of admitted users versus the achievable DoFs with different algorithms.

Riemannian metric [32]. The rapid convergence rate of the RTR algorithm with high precision solutions in a few iterations yields better performance in the procedure of user admission control.

### B. Admitted Users Versus Achievable DoFs and Optimal Admitted Users

Consider a 20-user partially connected interference channel where each cross link belongs to interference channel set  $\mathcal{V}$  with connectivity probability  $p = 0.8$ . The proposed three-stage Riemannian trust-region algorithm based user admission approach is compared with the RCG algorithm and the alternating minimization algorithm. We set  $\lambda = 0.5$  and  $\epsilon = 0.001$  in the sparse inducing optimization problem (12). A good choice of  $\rho$  is 0.01, which is obtained by cross validation. Fig. 6 demonstrates the average number of admitted users with different symmetric DoF allocations. Each point in the simulations is averaged over 50 randomly generated network topology realizations  $\mathcal{V}$ . From Fig. 6, we can see that the proposed three-stage Riemannian trust-region algorithm achieves significantly outperforms the RCG algorithm and

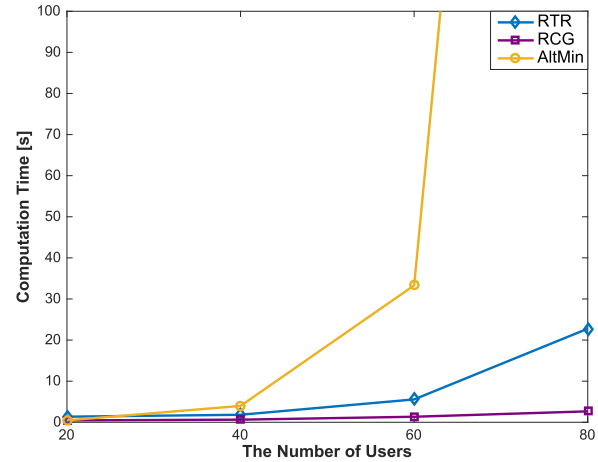


Fig. 7. The computation time in seconds versus the number of users  $K$  with different algorithms.

the AltMin algorithm, especially in the high DoF allocation regime.

To further justify the effectiveness of the Riemannian optimization framework, we numerically check that the proposed RTR algorithm can recover all the optimal user admission results for the class of deterministic TIM problems in [12]. That is, given a class of network topologies and optimal DoFs allocation, we verify the accessibility of the admitted users. Our work is the first work to systematically and numerically find the maximal number of admitted users to achieve the feasibility of topological interference alignment for *any* given DoF allocation and network topology.

### C. Time Results for Different Algorithms

Consider a  $K$ -user partially connected interference channel. The sets of the connected interference links are generated randomly with connectivity probability  $p = 0.1$ . We set  $\rho = 0.01$ ,  $\epsilon = 0.001$  and  $r = 8$  in the sparse inducing optimization problem (12). Each point in the simulations is averaged over 50 randomly generated network topology realizations  $\mathcal{V}$ . Fig. 7 shows the average computation time with different problem sizes for the user admission control problem with different algorithms. This figure shows that the RCG algorithm has the lowest computation time algorithm although it has a relatively slower convergence rate as shown in Fig. 5. Compared with the AltMin algorithm, the proposed RTR algorithm has much lower computation time compared with the AltMin algorithm. Meanwhile, the RTR algorithm achieves the best performance as shown in Fig. 6 with the comparable computation time as the RCG algorithm.

### D. Different Values of the Weighting Parameter $\lambda$

Consider a 20-user partially connected interference channel. The sets of the connected interference links are generated randomly with connectivity probability  $p$  varying from 0.1 to 0.9. We set  $\epsilon = 0.001$  and  $r = 4$  in the sparse inducing optimization problem (12). Fig. 8 shows the average number of admitted users with different values of the weighting parameter  $\rho$  in the regularized smoothed  $\ell_1$ -norm in (12). Each point in the

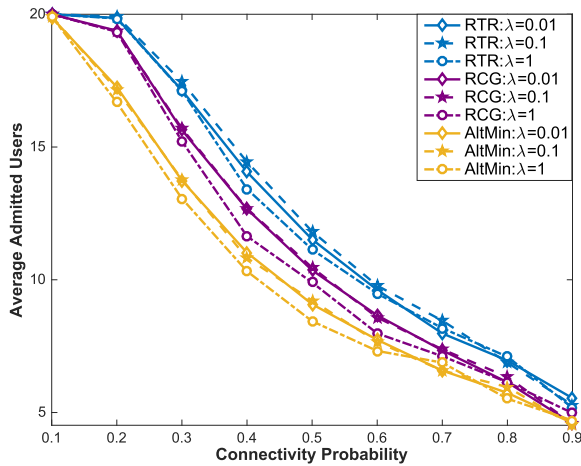


Fig. 8. Different weighting parameters  $\lambda$  in (12).

simulations is averaged over 50 randomly generated network topology realizations  $\mathcal{V}$ . Fig. 8 demonstrates that parameter  $\rho$  has little effect on the induced sparsity pattern in  $\text{diag}(\mathbf{X})$ , thereby yielding almost the same number of admitted users. The reason is that the the role of the weighting parameter  $\lambda$  in (12) only serves to upper bound the objective function. This figure further indicates that the proposed Riemannian trust region algorithm outperforms the Riemannian conjugate gradient algorithm and the alternating minimization algorithm.

## VII. CONCLUSIONS AND DISCUSSIONS

This paper presented a sparse and low-rank optimization framework for user admission control in topological interference management. A Riemannian optimization framework was further developed to solve the non-convex rank-constrained  $\ell_0$ -norm maximization problem, supported by a novel regularized smoothed  $\ell_1$ -norm sparsity inducing minimization approach. In particular, by exploiting the quotient manifold of fixed-rank matrices, we presented a Riemannian trust-region algorithm to find good solutions to the non-convex sparse and low-rank optimization problem. Simulation results illustrated the effectiveness and near-optimal performance of the proposed algorithms.

Several future directions of interest are as follows:

- It is desirable but challenging to theoretically establish the fundamental tradeoffs between the sparsity and low-rankness in the sparse and low-rank model  $\mathcal{P}$ . It is also interesting to extend the presented approaches in more generic scenarios, e.g., transmitter cooperation, finite SNRs and multiple antennas. The main challenge is establishing the channel independent interference alignment conditions.
- It is particularly interesting and also important to apply the sparse and low-rank modeling framework to other important problems including the index coding problem [48] (e.g., matrix completion over finite field [49]), caching networks [50], [51], and distributed computing systems [52], thereby investigating the fundamental limits of communication, computation and storage. However, as optimization on manifolds deeply relies on smoothness, the search space will become discrete in a finite

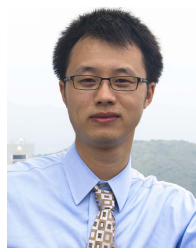
field. Therefore, the proposed Riemannian algorithm cannot be extended to the finite field in principle. Other numerical optimization techniques over finite field need to be introduced.

- It is also interesting to apply the Riemannian optimization technique to other important network optimization problems, e.g., hybrid precoding in millimeter wave systems [53].

## REFERENCES

- [1] M. R. Palattella *et al.*, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Mar. 2016.
- [2] 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.
- [3] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile Internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 460–473, Mar. 2016.
- [4] Y. Shi, J. Zhang, K. Letaief, B. Bai, and W. Chen, "Large-scale convex optimization for ultra-dense cloud-RAN," *IEEE Wireless Commun. Mag.*, vol. 22, no. 3, pp. 84–91, Jun. 2015.
- [5] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [6] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green cloud-RAN," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2809–2823, May 2014.
- [7] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.
- [8] T. Q. Quek, G. de la Roche, I. Güvenc, and M. Kountouris, *Small Cell Networks: Deployment, PHY Techniques, and Resource Management*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [9] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
- [10] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the  $K$ -user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [11] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
- [12] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.
- [13] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479–1494, Mar. 2011.
- [14] H. Maleki, V. R. Cadambe, and S. A. Jafar, "Index coding—An interference alignment perspective," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5402–5432, Sep. 2014.
- [15] X. Yi and D. Gesbert, "Topological interference management with transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6107–6130, Nov. 2015.
- [16] H. Sun and S. A. Jafar, "Topological interference management with multiple antennas," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 1767–1771.
- [17] A. E. Gamal, N. Naderialzadeh, and A. S. Avestimehr. (Jan. 2015). "When does an ensemble of matrices with randomly scaled rows lose rank?" [Online]. Available: <https://arxiv.org/abs/1501.07544>
- [18] N. Naderialzadeh and A. S. Avestimehr, "Interference networks with no CSIT: Impact of topology," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 917–938, Feb. 2015.
- [19] Y. Shi, J. Zhang, and K. B. Letaief, "Low-rank matrix completion for topological interference management by Riemannian pursuit," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4703–4717, Jul. 2016.
- [20] X. Zhai, L. Zheng, and C. W. Tan, "Energy-infeasibility tradeoff in cognitive radio networks: Price-driven spectrum access algorithms," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 528–538, Mar. 2014.
- [21] J. Zhao, T. Q. S. Quek, and Z. Lei, "Heterogeneous cellular networks using wireless backhaul: Fast admission control and large system analysis," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2128–2143, Oct. 2015.

- [22] Y. Shi, J. Cheng, J. Zhang, B. Bai, W. Chen, and K. B. Letaief, "Smoothed  $L_p$ -minimization for green cloud-RAN with user admission control," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 1022–1036, Apr. 2016.
- [23] E. Matakani, N. D. Sidiropoulos, Z. Q. Luo, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2682–2693, Jul. 2008.
- [24] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [25] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Found. Trends Mach. Learn.*, vol. 4, no. 1, pp. 1–106, Jan. 2012.
- [26] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 608–622, Jun. 2016.
- [27] M. Udell, C. Horn, R. Zadeh, and S. Boyd, "Generalized low rank models," *Found. Trends Mach. Learn.*, vol. 9, no. 1, pp. 1–118, 2016.
- [28] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
- [29] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: Phase transitions in convex programs with random data," *Inf. Inference*, vol. 3, pp. 224–294, Jun. 2014.
- [30] B. Vandereycken, "Low-rank matrix completion by Riemannian optimization," *SIAM J. Optim.*, vol. 23, pp. 1214–1236, Jun. 2013.
- [31] N. Boumal, P.-A. Absil, and C. Cartis. (May 2016). "Global rates of convergence for nonconvex optimization on manifolds." [Online]. Available: <https://arxiv.org/abs/1605.08101>
- [32] B. Mishra and R. Sepulchre, "Riemannian preconditioning," *SIAM J. Optim.*, vol. 26, pp. 635–660, Mar. 2016.
- [33] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, U.K.: Princeton Univ. Press, 2008.
- [34] J. Sun, Q. Qu, and J. Wright. (Oct. 2015). "When are nonconvex problems not scary?" [Online]. Available: <https://arxiv.org/abs/1510.06096>
- [35] N. Boumal, V. Voroninski, and A. S. Bandeira. (Jun. 2016). "The non-convex Burer–Monteiro approach works on smooth semidefinite programs." [Online]. Available: <https://arxiv.org/abs/1606.04970>
- [36] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. ACM Symp. Theory Comput.*, 2013, pp. 665–674.
- [37] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [38] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.
- [39] L. Ruan, V. K. N. Lau, and M. Z. Win, "The feasibility conditions for interference alignment in MIMO networks," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2066–2077, Apr. 2013.
- [40] G. Bresler, D. Cartwright, and D. Tse, "Feasibility of interference alignment for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5573–5586, Sep. 2014.
- [41] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning With Sparsity: The Lasso and Generalizations*. Boca Raton, FL, USA: CRC Press, 2015.
- [42] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre, "Fixed-rank matrix factorizations and Riemannian low-rank optimization," *Comput. Stat.*, vol. 29, nos. 3–4, pp. 591–621, 2014.
- [43] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a MATLAB toolbox for optimization on manifolds," *J. Mach. Learn. Res.*, vol. 15, pp. 1455–1459, Aug. 2014.
- [44] B. Mishra, K. A. Apuroop, and R. Sepulchre. (Nov. 2012). "A Riemannian geometry for low-rank matrix completion." [Online]. Available: <https://arxiv.org/abs/1211.1550>
- [45] B. Mishra and R. Sepulchre, "R3MC: A Riemannian three-factor algorithm for low-rank matrix completion," in *Proc. IEEE CDC*, Dec. 2014, pp. 1137–1142.
- [46] H. Kasai and B. Mishra, "Low-rank tensor completion: A Riemannian manifold preconditioning approach," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2016, pp. 1012–1021.
- [47] J. Nocedal and S. Wright, *Numerical Optimization*. New York, NY, USA: Springer, 2006.
- [48] M. Effros, S. El Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2478–2487, May 2015.
- [49] J. Saunderson, M. Fazel, and B. Hassibi, "Simple algorithms and guarantees for low rank matrix completion over  $F_2$ ," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 86–90.
- [50] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.
- [51] F. Xu, M. Tao, and K. Liu. (May 2016). "Fundamental tradeoff between storage and latency in cache-aided wireless interference networks." [Online]. Available: <https://arxiv.org/abs/1605.00203>
- [52] S. Li, M. A. Maddah-Ali, Q. Yu, and A. S. Avestimehr. (Apr. 2016). "A fundamental tradeoff between computation and communication in distributed computing." [Online]. Available: <https://arxiv.org/abs/1604.07086>
- [53] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.



**Yuanming Shi** (S'13–M'15) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2011, and the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology, in 2015. Since 2015, he has been a tenure-track Assistant Professor with the School of Information Science and Technology, ShanghaiTech University. He visited the University of California at Berkeley, Berkeley, CA, USA, from 2016 to 2017. His research areas include intelligent IoT, wireless networks, big data systems, mobile edge computing, by developing the tools from convex and nonconvex optimization, deep machine learning, statistics, and signal processing. He was a recipient of the 2016 IEEE Marconi Prize Paper Award in wireless communications and the 2016 Young Author Best Paper Award from the IEEE Signal Processing Society.



**Bamdev Mishra** received the B.Tech. and M.Tech. degrees in electrical engineering from the IIT Bombay, India, in 2010, and the Ph.D. degree from the Department of Electrical Engineering and Computer Science, University of Liège, Belgium, in 2014. He was a Post-Doctoral Fellow with the University of Liège and a Visiting Research Associate with the University of Cambridge in 2015. He is currently a Machine Learning Scientist with Amazon. His research interests include nonlinear optimization on matrix manifolds and machine learning.



**Wei Chen** (S'05–M'07–SM'13) received the B.S. and Ph.D. degrees (Hons.) from Tsinghua University in 2002 and 2007, respectively. From 2005 to 2007, he was also a visiting Ph.D. student with the Hong Kong University of Science and Technology. Since 2007, he has been on the faculty with Tsinghua University, where he is currently a Tenured Full Professor, the Director of Academic Degree Office, and a member of the University Council. From 2014 to 2016, he served as the Deputy Head with the Department of Electronic Engineering. He has also held visiting appointments with several universities, including Princeton University. His research interests are in the areas of wireless networks and information theory.

He is also supported by the NSFC Excellent Young Investigator Project, the New Century Talent Program of Ministry of Education, and the Beijing Nova Program. He is a member of the National 10000-Talent Program, a Cheung Kong Young Scholar, and a Chief Scientist of the National 973 Youth Project. He received the first prize of 14th Henry Fok Ying-Tung Young Faculty Award, the 17th Yi-Sheng Mao Science and Technology Award for Beijing Youth, the 2017 Young Scientist Award of the China Institute of Communications, and the 2015 Information Theory New Star Award of the China Institute of Electronics. He received the IEEE Marconi Prize Paper Award and the IEEE Comsoc Asia Pacific Board Best Young Researcher Award in 2009 and 2011, respectively. He was a recipient of the National May 1st Medal and Beijing Youth May 4th Medal. He serves as an Editor of the IEEE TRANSACTIONS COMMUNICATIONS and the IEEE TRANSACTIONS ON EDUCATION. He also serves as the Executive Chairman of the Youth Forum of the China Institute of Communications.