Interference alignment schemes for *k*-user interference channel based on manifold optimization

Chen Zhang[1,2]* , Ziwei Liu[1,2], Tao Hong[1,2] and Gengxin Zhang[1,2]

Abstract

Interference alignment (IA) is a key technology for achieving the capacity scaling required by next generation wireless networks, which is proved to obtain the maximum degrees of freedom (DoF). The aim of this paper is to propose interference alignment schemes through manifold optimization theory for *K*-user interference channel. We limit the optimization only at transmitters and relax the hypothesis of channel reciprocity to mitigate the overhead caused by alternation between the forward and reverse links significantly. Firstly, we introduce a classical algorithm based on the steepest descent (SD) algorithm in a multi-dimensional complex space to achieve feasible IA. Then, we reform the optimization problem on Stiefel manifold and propose a novel SD algorithm based on this manifold with lower dimensions. Moreover, aiming at further reducing the complexity, the Grassmann manifold is introduced to derive corresponding algorithm for reaching the perfect IA. Numerical simulations show that the proposed algorithms on manifolds have better performance both on system throughput and convergence than classical methods and also achieve the maximum DoF.

Keywords: Interference alignment, MIMO, Precoding, Manifold, Optimization

1 Introduction

Interference alignment (IA) has been envisioned as a promising technique [1, 2] to meet the overwhelming growth of data network traffic which is the main challenge of the next wireless networks. Compared with what is previously believed, it can obtain even more higher network capacity [3]. Generally speaking, there are three dominant interference alignment schemes: The first scheme is based on full channel state information (CSI), assuming the transmitters have the priori perfect CSI; the second one is based on limited (imperfect) CSI; and the third one do not need the CSI, which is called blind interference alignment. Except for the full CSI interference alignment scheme, the other schemes have no complete degrees of freedom (DoF) region even the total DoF for the *K*-user

interference channel [4, 5], thus the feasibility of imperfect CSI interference alignment and blind interference alignment is still an open problem. Consequently, in this paper, we focus on the full CSI interference alignment scheme. And the sum capacity of *K*-user $M \times M$ multiple-input multiple-output (MIMO) interference channel is

$$C_{\text{sum}} = \frac{KM}{2} \log(1 + \text{SNR}) + o(\log(\text{SNR})) \qquad (1)$$

where degrees of freedom (DoF) is defined as

$$\text{DoF} = \lim_{\text{SNR} \to \infty} \frac{C_{\text{sum}}}{\log(\text{SNR})} \qquad (2)$$

In this case, DoF is $KM/2$. It means each transmitter-receiver pair can communicate with $M/2$ DoF, irrespective of the number of interferers.

A feasible scheme to align interference is to design such a precoding that coordinates transmitting directions, for the purpose that the interference is forced to overlap as much as possible, and 1/2 signaling space is reserved for the desired signals at most. Based on the channel reciprocity's assumption, some pioneer studies such as [6–10] iteratively optimize both the precoding matrices and

*Correspondence: zhangchen@njupt.edu.cn
[1]National Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, People's Republic of China
[2]College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, People's Republic of China

interference suppression filters, by alternating between the uplinks and downlinks to align the interference in a distributed way. However, with the hypothesis of channel reciprocity, the application of these algorithms are only restricted within TDD systems. Moreover, at each node tight synchronization and feedback are needed for alternation between the downlinks and uplinks; when the channel varies fast, it may introduce too much overhead. Moreover, during each iteration process of optimization the transmitters and receivers exchange their "roles." Thus this scheme is improper for the receivers with limited ability of computing.

On the other hand, most of the previous works above employ traditional constrained optimization techniques that work in high dimensional complex space. Unavoidably, several shortcomings are accompanied with the traditional constrained optimization techniques such as low-converging speed and high-complexity.

To overcome these limitations, in this paper, we introduce optimization on matrix manifolds into the precoding scheme for interference alignment and limit the optimization only at the transmitters' side. Optimization n on manifolds consist the merits of lower complexity and better numerical properties. Firstly, for the sake of comparison, by employing classical constrained optimization method, a steepest descent (SD) algorithm in multi-dimensional complex space is provided to design the precoder of interference alignment. Then, we reformulate the constrained optimization problem to an unconstrained and non-degraded one on the complex Stiefel manifold with lower dimensions. We locally parameterize the manifold by Euclidean projection from the tangent space onto the manifold instead of the traditional method by moving descent step along the geodesic in [11, 12]. Thus, the SD algorithm on Stiefel manifold is proposed to achieve feasible interference alignment. To further reduce the computation complexity in terms of dimensions of manifold, we explore the unitary invariance property of our cost function and solve the optimization problem on the complex Grassmann manifold, then present the corresponding SD algorithm on the Grassmann manifold for interference alignment precoding design.

We not only generalize optimization algorithm on manifolds, but also turn the algorithm into an efficient numerical procedure to achieve perfect interference alignment. Moreover, by limiting the optimization algorithms performed at the transmitters' side only, all the three proposed algorithms are transparent at the receivers. Additionally, overhead generated by synchronization and feedback no longer exists since only transmitters participate in the iteration. Furthermore, by relaxing the assumption of channel reciprocity, our algorithms are applicable to both TDD and FDD systems. Furthermore, numerical simulation shows that the novel algorithms on manifolds

have better convergence performance and higher system capacity than previous methods. Finally, we prove the convergence of the proposed algorithms.

The paper is organized as follows: System model of interference alignment is presented in Section 2, followed by the detailed procedures of all three proposed SD methods for interference alignment in Section 3. In Section 4, numerical simulations and the corresponding discussion are stated. And the conclusion is given in the last section.

Notation: We use bold uppercase letters for matrices or vectors. X^{T} and $X^{\dagger}$ denote the transpose and the conjugate transpose (Hermitian) of the matrix X, respectively. Assuming the eigenvalues of a matrix X and their corresponding eigenvectors are sorted in ascending order, λ_X^i denotes the ith eigenvalue of the matrix X. Then, I represents the identity matrix. Moreover, $tr(\cdot)$ indicates the trace operation. And the Euclidean norm of X is $\|X\| = \sqrt{tr(X^{\dagger}X)}$. $\lfloor X \rfloor$ denotes the subspace spanned by the columns of X. $\mathbb{C}^{n \times p}$ represents the $n \times p$ dimensional complex space assuming $n > p$. $\mathbb{R}^+$ represents positive real number space. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex quantity, respectively. Finally $\kappa = \{1, \ldots, K\}$ is the set of integers from 1 to K.

2 System model

Consider the K-user wireless MIMO interference channel depicted in Fig. 1 where each transmitter and receiver are equipped with $M^{[k]}$ and $N^{[k]}$ antennas, respectively. Each transmitter communicates with its corresponding receiver and creates interference to all the other receivers. $d^{[k]}$ is the desired number of data streams between the kth transmitter-receiver pair. Additionally, $H^{[kj]}$ denotes the $N^{[k]} \times M^{[j]}$ channel coefficients matrix from the jth transmitter to the kth receiver and is assumed to have i.i.d.
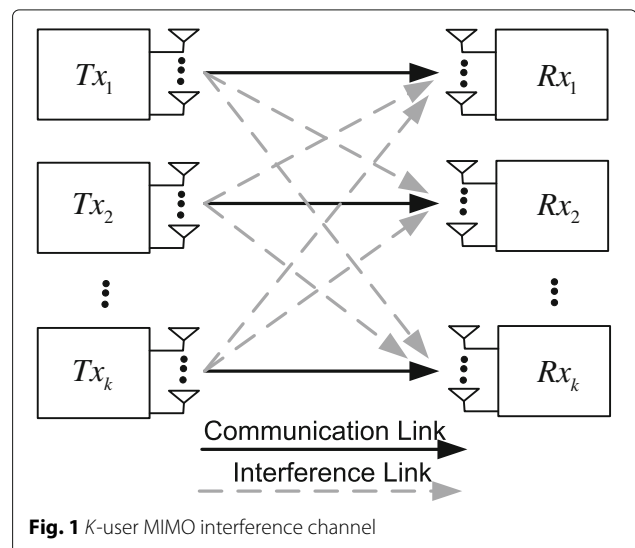

Fig. 1 *K*-user MIMO interference channel

complex Gaussian random variables, drawn from a continuous distribution. Moreover, H is prior known at the transmitters. Finally the received signal vector at receiver k after zero-forcing the interference is denoted by

$$\overline{Y}^{[k]} = U^{[k]\dagger} Y^{[k]} = U^{[k]\dagger} \left(\sum_{j=1}^{K} H^{[kj]} V^{[j]} S^{[j]} + W^{[k]} \right), k \in \kappa \tag{3}$$

where each element of the $d^{[j]} \times 1$ vector $S^{[j]}$ represents an independently encoded Gaussian symbol with power $P^{[j]}/d^{[j]}$ that beamformed with the corresponding $M^{[j]} \times d^{[j]}$ precoding matrix $V^{[j]}$ and then transmitted by the transmitter j. $U^{[k]}$ is the $N^{[k]} \times d^{[k]}$ interference zero-forcing filter at the receiver k. And $W^{[k]}$ is the i.i.d. complex Gaussian noise with zero mean unit variance.

2.1 Feasibility of interference alignment

The quality of alignment is measured by the interference power remaining in the intended signal subspace at each receiver. From [6], it can be obtained that the $d^{[k]}$-dimensional received signal subspace that contains the least interference is the space spanned by the eigenvectors corresponding to the $d^{[k]}$-smallest eigenvalues of the interference covariance matrix $Q^{[k]}$. Consequently, we try to minimize the sum of interference power spilled to the desired signal subspaces, by minimizing the sum of the absolute value of $d^{[k]}$-smallest eigenvalues of the interference covariance matrix at each receiver to create $d^{[k]}$-dimensional interference-free subspace for the desired signal.

2.2 Cost function

As previously stated, we try to minimize the sum of the $d^{[k]}$-smallest eigenvalues (in absolute value) of the interference covariance matrix at each receiver over the set of precoding matrices $V^{[1]}, \ldots, V^{[K]}$ [7]. Therefore, we define the cost function as follows:

$$\min_{V^{[1]}, \ldots, V^{[K]}} f = \sum_{k=1}^{K} \sum_{i=1}^{d^{[k]}} \left| \lambda_{Q^{[k]}}^{i} \right|, \ k, j \in \kappa \tag{4}$$

subject to $V^{[j]\dagger} V^{[j]} = I_{d^{[j]}}$

where

$$Q^{[k]} = \sum_{\substack{j=1 \\ j \neq k}}^{K} \frac{P^{[j]}}{d^{[j]}} H^{[kj]} V^{[j]} V^{[j]\dagger} H^{[kj]\dagger} \tag{5}$$

is the interference covariance matrix at receiver k. With the assumption that all the eigenvalues are sorted in ascending order, $\lambda_{Q^{[k]}}^{i}$ represents the ith eigenvalue of the corresponding interference covariance matrix $Q^{[k]}$. And because $Q^{[k]}$ is a Hermitian matrix, all its eigenvalues are

real. Therefore, the cost function $f(V), f : \mathbb{C}^{n \times p} \to \mathbb{R}^{+}$ is built.

3 Methods on different topologies for interference alignment

3.1 The steepest descent algorithm in complex space for iA

Since our cost function: $f(V), f : \mathbb{C}^{n \times p} \to \mathbb{R}^{+}$ is differentiable, intuitively, the steepest descent method can be employed to make the cost function converge to a local optimal point efficiently. Therefore, we will first find the closed-form expression of the steepest descent direction in $\mathbb{C}^{n \times p}$, then employ a property step size rule for each iteration.

As previously stated, the steepest descent method is tightly related to derivative and differentiation. In order to get the derivative of $f(V)$ over V, two Jacobian matrices blocks are employed as:

$$df = \begin{bmatrix} D_R^{[1]} & \ldots & D_R^{[K]} \end{bmatrix} \begin{bmatrix} dV_R^{[1]} \\ \cdot \\ \cdot \\ dV_R^{[K]} \end{bmatrix}$$
$$+ \begin{bmatrix} D_I^{[1]} & \ldots & D_I^{[K]} \end{bmatrix} \begin{bmatrix} dV_I^{[1]} \\ \cdot \\ \cdot \\ dV_I^{[K]} \end{bmatrix} \tag{6}$$

where $V_R^{[j]} = \Re\left\{V^{[j]}\right\}$, and $V_I^{[j]} = \Im\left\{V^{[j]}\right\}$. $D_R^{[j]}$ and $D_I^{[j]}$ are the $d^{[j]} \times M^{[j]}$ Jacobian matrices which denote the partial differential relation of the cost function over the real and imaginary parts of $V^{[j]}$, respectively. The detail of mathematical derivations can be found in [7, 13]. Thus, the derivative of f over $V^{[j]}$ is given by

$$D_V^{[j]} = \left(D_R^{[j]} + iD_I^{[j]} \right)^T \tag{7}$$

The inner product typically defined in the Euclidean multi-dimensional space is given as follows:

$$\langle Z_1, Z_2 \rangle = tr \left(Z_2^{\dagger} Z_1 \right) \tag{8}$$

Then, under the given inner product, the steepest descent direction is:

$$Z^{[j]} = -D_V^{[j]} = -\left(D_R^{[j]} + iD_I^{[j]} \right)^T \tag{9}$$

Once the formulation of steepest descent direction $Z^{[j]}$ is defined in (9), it is necessary to choose a suitable positive step size $\beta^{[j]}$ for each iteration. The Armijo step size rule [14] states that $\beta^{[j]}$ should be chosen to satisfy the following inequalities:

$$f(V) - f\left(V + \beta Z^{[j]}\right) \geq \frac{1}{2} \beta^{[j]} \left\langle Z^{[j]}, Z^{[j]} \right\rangle \tag{10}$$

$$f(V) - f\left(V + 2\beta Z^{[j]}\right) < \beta^{[j]} \left\langle Z^{[j]}, Z^{[j]} \right\rangle \tag{11}$$

Rule (10) guarantees that the step $\beta^{[j]}Z^{[j]}$ will expressively decrease the cost function, whereas (11) undertakes that the step $2\beta^{[j]}Z^{[j]}$ would not be a better choice. A direct procedure for acquiring a suitable $\beta^{[j]}$ is to keep on doubling $\beta^{[j]}$ until (11) no longer holds and then halving $\beta^{[j]}$ until it satisfies (10). It can be proved that such $\beta^{[j]}$ can always be found [15].

Consolidating all the ideas stated above, we present our algorithm in Algorithm 1. In Step 3 and Step 4, the Armijo step rule is performed to find a proper convergence step length. Generally speaking, Step 3 ensures the chosen step $\beta^{[j]}$ will significantly reduce the cost function while Step 4 prevents $\beta^{[j]}$ from being too large that may miss the potential optimal point. The operator $gs(\cdot)$ means Gram-Schmidt Orthogonalization [16] of a matrix, which guarantees the newly computed solution $V^{[j]}$ (or $B_1^{[j]}$, $B_2^{[j]}$) still satisfies the unitary constraint.

Discussion:

(i) The inner product and the gradient direction are defined in different topologies in [7]. However, it is considered to be inappropriate because the gradient is defined after the inner product is given only. In

---

Algorithm 1 The Steepest Descent Algorithm in Complex Space for IA

Start with arbitrary precoding matrices $V^{[1]},...,V^{[K]}$, set

initial step size $\beta^{[j]} = 1$ and begin iteration.

for $j = 1...K$

(1) Compute the Jacobian matrix $D_R^{[j]}$ and $D_I^{[j]}$

(2) Get the steepest descent direction
$$Z^{[j]} = -D_V^{[j]} = -(D_R^{[j]} + iD_I^{[j]})^T$$

(3) Compute $B_1^{[j]} = gs(V^{[j]} + 2\beta^{[j]}Z^{[j]})$,

if $f(V^{[1]},...,V^K) - f(V^{[1]},...,V^{[j-1]}, B_1^{[j]}, V^{[j+1]}...V^{[K]})$

$\geq \beta^{[j]}tr(Z^{[j]\dagger}Z^{[j]})$, then set $\beta^{[j]} := 2\beta^{[j]}$, repeat Step 3.

(4) Compute $B_2^{[j]} = gs(V^{[j]} + \beta^{[j]}Z^{[j]})$,

if $f(V^{[1]},...,V^{[K]}) - f(V^{[1]},...,V^{[j-1]}, B_2^{[j]}, V^{[j+1]}...V^{[K]})$

$< \frac{1}{2}\beta^{[j]}tr(Z^{[j]\dagger}Z^{[j]})$, then set $\beta^{[j]} := \frac{1}{2}\beta^{[j]}$, repeat Step 4.

(5) $V^{[j]} = gs(V^{[j]} + \beta^{[j]}Z^{[j]})$

(6) Continue till the cost function f is sufficiently small.

---

other words, the inner product and the gradient direction must be defined in the same topology. Our proposed algorithm rectifies the topology flaw in [7] and thus avoids the risk of non-convergence.

(ii) It can be concluded that Algorithm 1 belongs to the classical optimization method [17], which means it works in multi-dimensional space $\mathbb{C}^{n\times p}$ with the dimensions:

$$\dim(\mathbb{C}^{n\times p}) = np \qquad (12)$$

Obviously, the algorithm complexity increases with the dimensions. As previous discussed, optimization algorithms on manifolds work in an embedded or quotient space whose dimension will be much smaller than that of classical constrained optimization methods. Thus, optimization algorithms on manifolds not only have lower complexity, but also perform better numerical properties. The corresponding algorithms on manifolds will be stated for IA in the next two subsections.

3.2 The steepest descent algorithm on complex stiefel manifold for iA

Informally, a manifold is a space that is "modeled on" Euclidean space. It can be defined as a subset of Euclidean space which is locally the graph of a smooth function.

Conceptually, the simplest approach to optimize a differentiable function is to continuously translate a test point in the direction of the steepest descent on the constraint set until one reaches a point where the gradient is equal to zero. However, there are two challenges for optimization on manifolds. First, in order to define algorithms on manifolds, these operations above must be translated into the language of differential geometry. Second, once the test point shifts along the steepest descent direction, it must be retracted back to the manifold. Therefore, after reformulating the optimization problem on the Stiefel manifold, we introduce definitions about project operation and tangent space for retraction and gradient, respectively.

In many cases, the underlying symmetry property can be exploited to reformulate the problem as a non-degenerate optimization problem on manifolds associated with the original matrix representation. Thus, the constraint condition $V^{[j]\dagger}V^{[j]} = I$ in the cost function (4) inspires us to solve the problem on the complex Stiefel manifold. The complex Stiefel manifold [16] $St(n,p)$ is the set satisfying

$$St(n,p) = \{X \in \mathbb{C}^{n\times p} : X^{\dagger}X = I\} \qquad (13)$$

$St(n,p)$ naturally embeds in $\mathbb{C}^{n\times p}$ and inherits the usual topology of $\mathbb{C}^{n\times p}$. It is a compact manifold and from [18], we can get:

$$\dim(St(n,p)) = np - \frac{1}{2}p(p+1) \tag{14}$$

Another important definition is the projection. Assuming $Y \in \mathbb{C}^{n \times p}$ is a rank-p matrix, the projection operator $\pi_{st}(\cdot) : \mathbb{C}^{n \times p} \to St(n,p)$ is given by

$$\pi_{st}(Y) = \arg \min_{X \in St(n,p)} \|Y - X\|^2 \tag{15}$$

It can be proved that there exists a unique solution if Y has full column rank [18]. From (15), it can be acquired that the projection of an arbitrary rank-p matrix Y onto the Stiefel manifold is defined to be the point on the Stiefel manifold closest to Y in the Euclidean norm [19]. Moreover, if the singular value decomposition (SVD) of Y is $Y = U \sum V^\dagger$, then

$$\pi_{st}(Y) = UI_{n \times p} V^\dagger \tag{16}$$

Consider $X \in St(n,p)$ and its disturbing point $\pi_{st}(X + \varepsilon Y) \in St(n,p)$ for certain directions matrix $Y \in \mathbb{C}^{n \times p}$ and scalar $\varepsilon \in \mathbb{R}$. If Y satisfies $f(\pi_{st}(X + \varepsilon Y)) = f(X) + O(\varepsilon^2)$ which means certain directions Y do not cause $\pi_{st}(X + \varepsilon Y)$ to move away from X as ε increases. The collection of such directions Y is called the normal space at X of $St(n,p)$ [18]. The tangent space $T_X(n,p)$ is defined to be the orthogonal complement of the normal space, which can be roughly illustrated as Fig. 2. And the mathematical expression of the tangent space $T_X(n,p)$ at $X \in St(n,p)$ is defined by

$$T_X(n,p) = \left\{ Z \in \mathbb{C}^{n \times p} : Z = XA + X_\perp B, A \in \mathbb{C}^{p \times p}, \right.$$
$$\left. A + A^\dagger = 0, B \in \mathbb{C}^{(n-p) \times p} \right\} \tag{17}$$

in which $X_\perp \in \mathbb{C}^{n \times (n-p)}$ is defined to be any matrix satisfying $[X \, X_\perp]^\dagger [X \, X_\perp] = I$ and is the complement of $X \in St(n,p)$. Also from [16], it can be obtained that the gradient of our cost function is in the tangent space $T_X(n,p)$. And the dimension of $T_X(n,p)$ is:
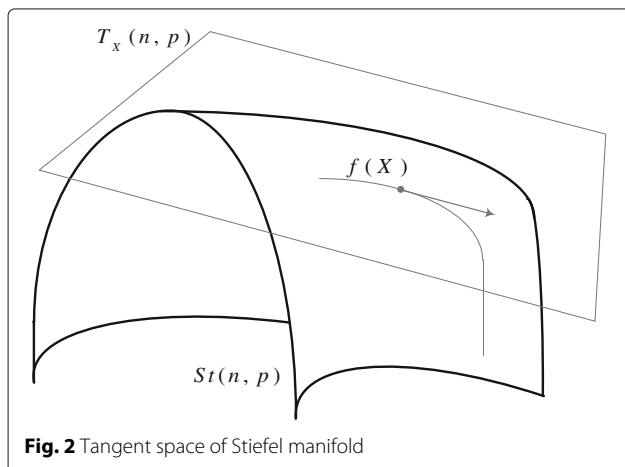
$$\dim(T_X(n,p)) = p(2n - p) \tag{18}$$

Fig. 2 Tangent space of Stiefel manifold

Obviously the steepest descent algorithm requires the computation of the gradient. As we previously emphasize, the gradient is only defined after $T_X(n,p)$ is given an inner product:

$$\langle Z_1, Z_2 \rangle = \Re \left\{ tr(Z_2^\dagger \left(I - \frac{1}{2} XX^\dagger \right) Z_1) \right\} \tag{19}$$

where $Z_1, Z_2 \in T_X(n,p)$ and $X \in St(n,p)$. The derivation of (19) can be found in [16]. Therefore, under the defined inner product, the steepest descent direction of the cost function $f(X)$ at the point $X \in St(n,p)$ is

$$Z = XD_X^\dagger X - D_X \tag{20}$$

where D_X is the derivative of $f(X)$.

The proposed SD algorithm on complex Stiefel manifold is presented in Algorithm 2. From (19) and (20), it can be easily obtained that the inner product needed for the Armijo step rule is

$$\left\langle Z^{[j]}, Z^{[j]} \right\rangle = \Re \left\{ tr \left(Z^{[j]\dagger} \left(I - \frac{1}{2} V^{[j]} V^{[j]\dagger} \right) Z^{[j]} \right) \right\} \tag{21}$$

which is used in Step 4 and 5, and the steepest descent on Stiefel manifold of our cost function is

$$Z^{[j]} = V^{[j]} D_V^{[j]\dagger} V^{[j]} - D_V^{[j]} \tag{22}$$

which is used in Step 3. Noticing that the project operation $\pi_{st}(\cdot)$ in Step 6 (Steps 4 and 5) guarantees the newly computed solution $V^{[j]}$ (or $B_1^{[j]}$, $B_2^{[j]}$) after iteration still satisfies $V^{[j]} \in St(n,p)$. Using the method of SVD, we can easily compute the project operation. *Discussion*:

(i) As previous stated, the algorithms in [11, 12] were performed by moving the descent step along the geodesic of the constrained surface within each iteration. A disadvantage of this method is the redundant computational cost for calculating the path of a geodesic. In this paper, we locally parameterize the manifold by Euclidean projection from the tangent space onto the manifold instead of moving along a geodesic, to achieve a modest reduction in the computational complexity of the algorithms.

(ii) Recall (12) and (18), it can be obtained that when we reformulate the problem from $\mathbb{C}^{n \times p}$ to $St(n,p)$, the dimension of the optimization problem decreases from np to $np - \frac{1}{2}p(p+1)$. Although such dimension-dissension can be observed clearly, we still intend to reduce the dimensions of the space which the optimization algorithm works in. Thus, the Grassmann manifold and its corresponding algorithm for IA are stated in the following subsection.

Algorithm 2 The Steepest Descent Algorithm on Complex Stiefel Manifold for IA

Start with arbitrary precoding matrices $V^{[1]}, ..., V^{[K]}$,

for $j = 1...K$

(1) Compute the Jacobian matrix $D_R^{[j]}$ and $D_I^{[j]}$

(2) Then get the derivative of f:

$$D_V^{[j]} = (D_R^{[j]} + iD_I^{[j]})^T$$

(3) Get the steepest descent direction

$$Z^{[j]} = V^{[j]}D_V^{[j]\dagger}V^{[j]} - D_V^{[j]}$$

(4) Compute $B_1^{[j]} = \pi_{st}(V^{[j]} + 2\beta^{[j]}Z^{[j]})$,

if $f(V^{[1]}, ..., V^K) - f(V^{[1]}, ..., V^{[j-1]}, B_1^{[j]}, V^{[j+1]}...V^{[K]}) \geq$

$\beta^{[j]}\Re\{tr(Z^{[j]\dagger}(I - \frac{1}{2}V^{[j]}V^{[j]\dagger})Z^{[j]})\}$, then set $\beta^{[j]} := 2\beta^{[j]}$,

and repeat Step 4.

(5) Compute $B_2^{[j]} = \pi_{st}(V^{[j]} + \beta^{[j]}Z^{[j]})$,

if $f(V^{[1]}, ..., V^{[K]}) - f(V^{[1]}, ..., V^{[j-1]}, B_2^{[j]}, V^{[j+1]}...V^{[K]}) <$

$\frac{1}{2}\beta^{[j]}\Re\{tr(Z^{[j]\dagger}(I - \frac{1}{2}V^{[j]}V^{[j]\dagger})Z^{[j]})\}$, then set $\beta^{[j]} := \frac{1}{2}\beta^{[j]}$,

and repeat Step 5.

(6) $V^{[j]} = \pi_{st}(V^{[j]} + \beta^{[j]}Z^{[j]})$

(7) Continue till the cost function f is sufficiently small.

3.3 The steepest descent algorithm on complex grassmann manifold for iA

Notice that our cost function $f(V)$ satisfies $f(VU) = f(V)$ for any unitary matrix U. Because

$$Q^{[k]}(VU) = \sum_{\substack{j=1 \\ j\neq k}}^{K} \frac{P^{[j]}}{d^{[j]}}H^{[kj]}V^{[j]}U^{[j]}U^{[j]\dagger}V^{[j]\dagger}H^{[kj]\dagger}$$

$$= \sum_{\substack{j=1 \\ j\neq k}}^{K} \frac{P^{[j]}}{d^{[j]}}H^{[kj]}V^{[j]}IV^{[j]\dagger}H^{[kj]\dagger}$$

$$= Q^{[k]}(V) \tag{23}$$

which means that multiplying unitary matrix U does not change the eigenvalues and their corresponding eigenvectors of the interference covariance matrix at each receiver. Thus, our cost function f should be minimized on the

Grassmann manifold rather than on the Stiefel manifold. This is because the Grassmann manifold treats V and VU as equivalent points, leading to a further reduction in the dimension of the optimization problem. Similar with the previous subsection, we firstly introduce the definition about the Grassmann manifold, then present the project operation and tangent space of Grassmann manifold for retraction and gradient, respectively.

The complex Grassmann manifold $Gr(n, p)$ is defined to be the set of all p-dimensional complex subspaces of $\mathbb{C}^{n\times p}$. The Grassmann manifold can be thought as a quotient space of the Stiefel manifold: $Gr(n, p) \simeq St(n, p)/St(p, p)$. Quotient space is more difficult to visualize, as it is not defined as a set of matrices; rather, each point of the quotient space is an equivalence class of $n \times p$ matrices.

However, we can understand quotient space in this way: assuming $X \in St(n, p)$ is a point on the Stiefel manifold, the columns of X span an orthonormal basis for a p-dimensional quotient subspace. That is to say, if $\lfloor X \rfloor$ denotes the subspace spanned by the columns of X, then $X \in St(n, p)$ implies $\lfloor X \rfloor \in Gr(n, p)$. Therefore, there is a one-to-one mapping between points on the Grassmann manifold $Gr(n, p)$ and equivalence classes of $St(n, p)$. From (14), it can be acquired that:

$$\dim(Gr(n, p)) = \dim(St(n, p)) - \dim(St(p, p))$$
$$= p(n - p) \tag{24}$$

Let $Y \in \mathbb{C}^{n\times p}$ be a rank-p matrix. The projection operator $\pi_{gr}(\cdot) : \mathbb{C}^{n\times p} \to Gr(n, p)$ onto the Grassmann manifold is defined to be

$$\pi_{gr}(Y) = \left\lfloor \arg\min_{X\in St(n,p)} \|Y - X\|^2 \right\rfloor \tag{25}$$

It also can be proved that there exists a unique solution if Y has full column rank [18]. From (25), it can be acquired that the projection of an arbitrary rank-p matrix Y onto the Grassmannn manifold is defined to be the subspace spanned by the point on the Stiefel manifold closest to Y in the Euclidean norm. Furthermore, if the QR decomposition of Y is $Y = QR$, the following equality holds:

$$\pi_{gr}(Y) = \lfloor QI_{n\times p} \rfloor \tag{26}$$

The proof of (26) also can be found in [16]. From (26), it is obvious that $\pi_{gr}(Y)$ is the subspace spanned by the first p columns of Q.

As discussed before, Grassmann manifold is a quotient space of the Stiefel manifold, thus its tangent space is a subspace of the Stiefel manifold's tangent space [18]. If $X \in St(n, p)$, the tangent space $T_{\lfloor X \rfloor}(n, p)$ at $\lfloor X \rfloor \in Gr(n, p)$ of Grassmann manifold is:

$$T_{\lfloor X \rfloor}(n, p) = \left\{ Z \in \mathbb{C}^{n\times p} : Z = X_\perp B, B \in \mathbb{C}^{(n-p)\times p} \right\} \tag{27}$$

Recall (18), the dimension of the tangent space $T_{\lfloor X \rfloor}(n,p)$ of the complex Grassmann manifold is:

$$\dim\left(T_{\lfloor X \rfloor}(n,p)\right) = \dim(T_X(n,p)) - \dim(T_X(p,p))$$
$$= p(2n - 2p) \tag{28}$$

Moreover, the inner product of $T_{\lfloor X \rfloor}(n,p)$ is given by:

$$\langle Z_1, Z_2 \rangle = \Re\{tr(Z_2^{\dagger} Z_1)\},\ Z_1, Z_2 \in T_{\lfloor X \rfloor}(n,p),$$
$$X \in St(n,p) \tag{29}$$

The derivation of (29) can be found in [16]. Therefore, under the defined inner product, the steepest descent direction [16] of the cost function $f(X)$ at the point $X \in Gr(n,p)$ is:

$$Z = -(I - XX^{\dagger})D_X \tag{30}$$

where D_X is the derivative of $f(X)$.

The proposed SD algorithm on the complex Grassmann manifold is presented in Algorithm 3. Similar with the previous proposed algorithms, the Armijo step rule is performed to find a proper convergence step length. From

---

Algorithm 3 The Steepest Descent Algorithm on Complex Grassmann Manifold for IA

Start with arbitrary precoding matrices $V^{[1]}, ..., V^{[K]}$,

for $j = 1...K$

(1) Compute the Jacobian matrix $D_R^{[j]}$ and $D_I^{[j]}$

(2) Then get the derivative of f:

$$D_V^{[j]} = (D_R^{[j]} + i D_I^{[j]})^T$$

(3) Get the steepest descent direction

$$Z^{[j]} = -(I - V^{[j]} V^{[j]H}) D_V^{[j]}$$

(4) Compute $B_1^{[j]} = \pi_{gr}(V^{[j]} + 2\beta^{[j]} Z^{[j]})$,

if $f(V^{[1]}, ..., V^K) - f(V^{[1]}, ..., V^{[j-1]}, B_1^{[j]}, V^{[j+1]}...V^{[K]}) \geq$

$\beta^{[j]} tr(Z^{[j]H} Z^{[j]})$, then set $\beta^{[j]} := 2\beta^{[j]}$, and repeat Step 4.

(5) Compute $B_2^{[j]} = \pi_{gr}(V^{[j]} + \beta^{[j]} Z^{[j]})$,

if $f(V^{[1]}, ..., V^{[K]}) - f(V^{[1]}, ..., V^{[j-1]}, B_2^{[j]}, V^{[j+1]}...V^{[K]}) <$

$\frac{1}{2}\beta^{[j]} tr(Z^{[j]H} Z^{[j]})$, then set $\beta^{[j]} := \frac{1}{2}\beta^{[j]}$, and repeat Step 5.

(6) $V^{[j]} = \pi_{gr}(V^{[j]} + \beta^{[j]} Z^{[j]})$

(7) Continue till the cost function f is sufficiently small.

---

(29) and (30), it can be easily concluded that the inner product needed for the Armijo step rule is

$$\left\langle Z^{[j]}, Z^{[j]} \right\rangle = tr\left(Z^{[j]\dagger} Z^{[j]}\right) \tag{31}$$

which is used in Step 4 and 5, and the steepest descent on Grassmann manifold of our cost function is

$$Z^{[j]} = -\left(I - V^{[j]} V^{[j]\dagger}\right) D_V^{[j]} \tag{32}$$

which is used in Step 3. And the project operation $\pi_{gr}(\cdot)$ in Step 6 (Step 4, 5) retracts the newly computed solution $V^{[j]}$ (or $B_1^{[j]}$, $B_2^{[j]}$) back onto Grassmann manifold $Gr(n,p)$. Using QR decomposition, we can easily compute the project operation.

Discussion:

(i) Comparing $\dim(Gr(n,p)) = p(n - p)$ in (24) with $\dim(St(n,p)) = np - \frac{1}{2}p(p + 1)$ in (14), a further dimension reduction can be observed. Similarly, from (28), we can see another advantage of using the Grassmann manifold rather than the Stiefel manifold which is that $T_{\lfloor X \rfloor}(n,p)$ has only $p(2n - 2p)$ dimensions, whereas tangent space of $St(n,p)$ has $p(2n - p)$ dimensions. And from [20], it can be obtained that in our system model, if each transceiver is equipped with same amount of antenna $(M = N)$, then

$$\sum_{k=1}^{K} d^{[k]} = K \cdot d = \frac{K \cdot M}{2} \tag{33}$$

and

$$d = \frac{M}{2} \tag{34}$$

(ii) Recall (12), (14), and (24), we can get that if M is large enough (M not only can represent the number of antennas each transceiver equipped, but also can refer to the number of time extension slots [1, 20]), hence

$$\frac{\dim(St(M,d))}{\dim\left(\mathbb{C}^{M\times d}\right)} \approx \frac{3}{4} \tag{35}$$

which is a clear evidence for dimension-descension. And

$$\frac{\dim(Gr(M,d))}{\dim\left(\mathbb{C}^{M\times d}\right)} = \frac{1}{2} \tag{36}$$

holds for any integer M. (36) means that optimization on Grassmann manifold would reduce dimension further. The trend of dimension-descension can be roughly illustrated in Fig. 3.

4 Numerical results and discussion

Without symbol extension, the feasible condition of k-user interference alignment [20] is given by:
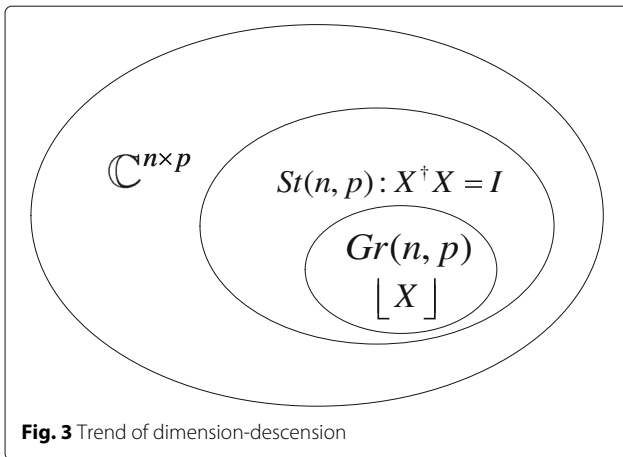
$$M^{[j]} + N^{[j]} \geq (K + 1)d^{[j]} \tag{37}$$

Fig. 3 Trend of dimension-descension

For satisfying feasibility and simple computation, we consider a 3-user 2×2 MIMO interference channel where the desired DoF per user $d^{[j]}$ is 1. All the algorithms are executed under the same scenario including randomly generated channel coefficients, initial precoding matrices, and convergence step length. We simulated the proposed three SD algorithms through 100 simulation realizations.

In order to compare the convergence performance, the average values of 100 realizations results are illustrated in Fig. 4. It can be observed that the algorithms on manifolds have better convergence performance comparing with the classical optimization method as our expectation.

Meanwhile, since there are two interference signals at each receiver, as shown in Fig. 5, the angles

between the spaces spanned by each interference signals asymptotically converge to zero within one simulation realization, which is another evidence for achieving the perfect interference alignment.

Finally, we compare the system sum-rate of the proposed algorithms. Figure 6 shows that the SD algorithm on Stiefel manifold and the SD algorithm on Grassmann manifold almost have the same performance and outperform the other classical optimization algorithms (Distributed IA in [6, 7]). More importantly, at high SNR, the DoF of the three proposed algorithms nearly achieve 3, which is the maximum theoretical value ($KM/2 = 3$) . Therefore, the perfect interference alignment is successfully achieved.

Three reasons leading to the fact that the proposed algorithms on manifolds have better performance are presented below:

(i) The advantage of the proposed algorithms is attributed to the reason that we reformulate the constrained optimization problem to an unconstrained one on manifolds with lower complexity and better numerical properties; then, locally parameterize the manifolds by a Euclidean projection of the tangent space on to the manifolds instead of moving along the geodesic, as stated in the previous sections. Moreover, the convergence performance curve of SD algorithm on the Stiefel manifold and the curve of SD algorithm on the Grassmann manifold almost overlap. Recall that optimization on the Grassmann manifold would reduce dimension further; therefore, the SD
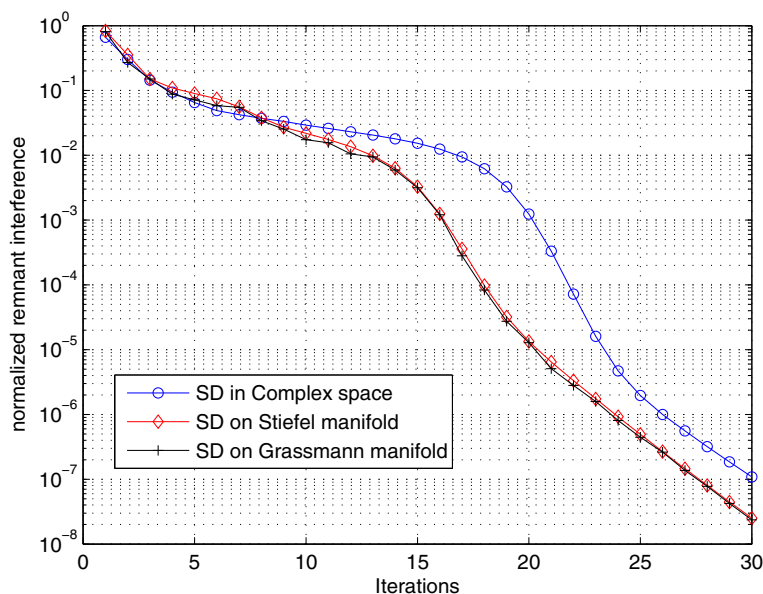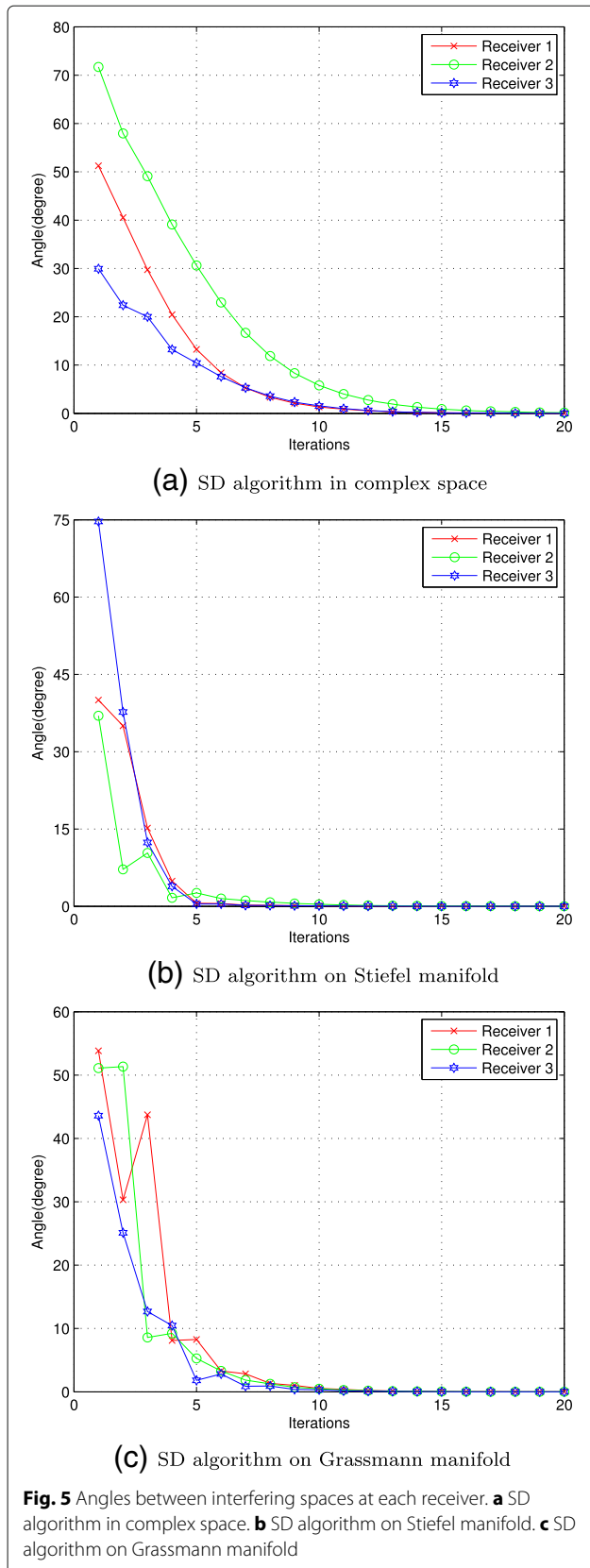

Fig. 4 Convergence performance

(a) SD algorithm in complex space

(b) SD algorithm on Stiefel manifold

(c) SD algorithm on Grassmann manifold

Fig. 5 Angles between interfering spaces at each receiver. **a** SD algorithm in complex space. **b** SD algorithm on Stiefel manifold. **c** SD algorithm on Grassmann manifold

algorithm on the Grassmann manifold will guarantee performance and reduce the computation complexity at the same time.

(ii) It is noticed that our cost function actually is the interference power spilled from the interference space to the desired signal space. With the better convergence performance, the SD algorithms on manifolds will have less remnant interference in the desired signal space within same iteration times. Therefore, it's guaranteed to achieve higher SINR [21]:

$$SINR = \frac{signal\ power}{noise + remnant\ interference} \qquad (38)$$

which leads to high capacity.

(iii) At each receiver, the zero-forcing filter is adopted. It will project the desired signal power and the remnant interference onto the subspace which is orthogonal with the subspace spanned by the interference. After performing the SD algorithms on manifolds, it is observed that in the Euclidean norm distance, the subspace spanned by desired signal is more close to the orthogonal complement of the interference subspace. Therefore, even the proposed algorithms on manifolds finally get the same remnant interference as the classical optimization methods results. The algorithms on manifolds will suffer from less power lose during the projection operated by zero-forcing filter; hence, higher system capacity can be achieved.

Above all, we didn't just correct the defects of our previous results. More importantly, many innovations and improvements are made in this paper. Firstly, through the in-depth study of matrix manifold, many novel manifold conceptions (such as Stiefel manifold, Grassmann manifold, quotient space, dimension decrease, and projection) are introduced as the theoretical foundation to support reformulation of interference alignment objective function on manifolds and also promote the proposed algorithms lower complexity, better convergence performance and higher system capacity. Secondly, from a self-contained system perspective, we cite part of our previous result in Section 3 Algorithm 1 and propose novel algorithms on three different topologies (complex space, Stiefel manifold, and Grassmann manifold) for interference alignment. More importantly, we uniform the flow of the three proposed algorithms to make the comparison of algorithms' results become more meaningful and logicality.

We notice that better throughputs may be attained by using non-unitary precoding, or by applying power waterfilling in the equivalent non-interfering MIMO channels.
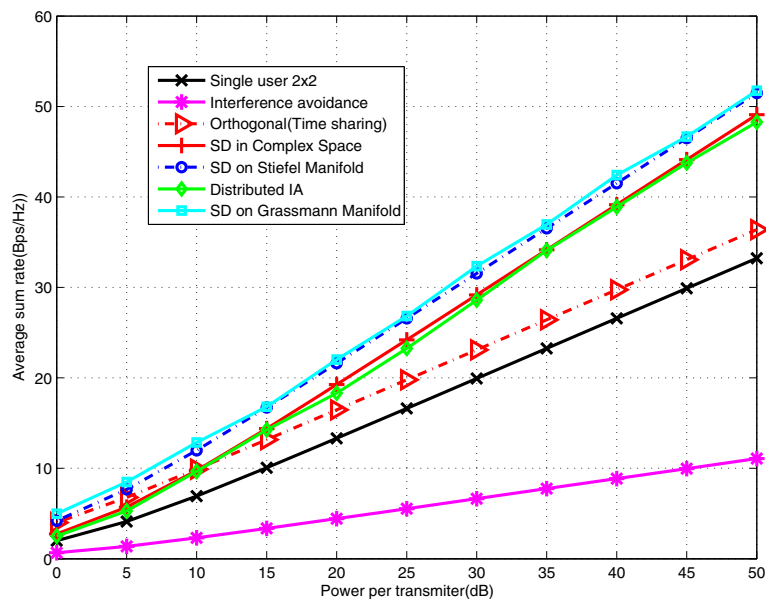
Fig. 6 Sum-rate capacity

Nevertheless, these methods to increase throughputs can be performed as the second step after the interference alignment is achieved. Thus in this paper, we only need to concentrate on the first step to find the perfect solutions of interference alignment.

More importantly, we notice that the limited CSI will give a disturbance to our algorithm [22–24]. However, its negative influence on our algorithm is tolerable due to three main reasons. First, the proof for the convergence of our proposed algorithm is solid: our cost function is non-negative with the low bound zero. It monotonically decreases within each iteration; therefore, it must converge to a solution. Second, due to the property of steepest decent method, our algorithm will converge to a local optimum point even with a disturbance. Third, the numerical property of manifolds will guarantee the residual interference (the value of local optimum point) is very small which leads to high capacity.

5 Conclusion and future work
In this paper, we focus on the interference alignment schemes by employing manifold optimization theory. By restricting the optimization only at the transmitters' side, the overhead induced by alternation between the forward and reverse links will be alleviated significantly. A classical SD algorithm in multi-dimensional complex space is proposed first. Then, we reform the optimization problem on Stiefel manifold and propose a novel SD algorithm on this manifold with lower dimensions. Moreover, aiming at further reducing the complexity, the Grassmann manifold

is introduced to derive the corresponding algorithm for reaching the perfect interference alignment. Numerical simulations show that comparing with previous methods, the proposed algorithms on manifolds have better convergence performance, higher system capacity, and also achieve the maximum DoF.

In our future work, we will employ Newton-type method on manifolds to achieve quadratic convergence and try to find global optimum results. On the other hand, we already begin the research on Grassmannian differential quantization theory and deep-learning method [25] to offer a efficient feedback strategy for limited CSI interference alignment. This complicated work is still in progress.

Authors' contributions
CZ conceived the methods and wrote the paper. CZ and ZL analyzed the simulation data. GZ and TH gave valuable suggestions on the structure of the paper. ZL and TH revised the original manuscript. All authors read and approved the manuscript.

Availability of data and materials
All data are fully available without restriction.

Zhang *et al. EURASIP Journal on Wireless Communications and Networking* (2019) 2019:196

Page 11 of 11

References
1. H. Maleki, V. R. Cadambe, S. A. Jafar, Index coding an interference alignment perspective. IEEE Trans. Inf. Theory. **60**(9), 5402–5432 (2014)
2. A. G. Davoodi, S. A. Jafar, Generalized degrees of freedom of the symmetric k-user interference channel under finite precision CSIT. IEEE Trans. Inf. Theory. **63**(10), 6561–6572 (2017)
3. B. Yuan, H. Sun, S. A. Jafar, Replication-based outer bounds and the optimality of half the cake for rank-deficient MIMO interference networks. IEEE Trans. Inf. Theory. **63**(10), 6607–6621 (2017)
4. M. Morales-Cespedes, J. Plata-Chaves, D. Toumpakaris, S. A. Jafar, A. G. Armada, Cognitive blind interference alignment for macro-femto networks. IEEE Trans. Signal Process. **65**(19), 5121–5136 (2017)
5. A. G. Davoodi, S. A. Jafar, Transmitter cooperation under finite precision CSIT: a GDoF perspective. IEEE Trans. Inf. Theory. **63**(9), 6020–6030 (2017)
6. K. S. Gomadam, V. R. Cadambe, S. A. Jafar, A distributed numerical approach to interference alignment and applications to wireless interference networks. IEEE Trans. Inf. Theory. **57**(6), 3309–3322 (2011)
7. H. G. Ghauch, C. B. Papadias, in *IEEE Global Telecommunications Conference (GLOBECOM)*. Interference alignment: a one-sided approach (IEEE, Texas, 2011), pp. 1–5
8. J. Fanjul, O. Gonzalez, I. Santamaria, C. Beltran, Homotropy continuation for spatial interference alignment in arbitrary MIMO X networks. IEEE Trans. Signal Proc. **65**(7), 1752–1764 (2017)
9. S. M. Razavi, Unitary Beamformer designs for MIMO interference broadcast channels. IEEE Trans. Signal Process. **64**(8), 2090–2102 (2016)
10. H. Al-Shatri, X. Li, R. S. Ganesan, A. Klein, T. Weber, Maximizing the sum rate in cellular networks using multiconvex optimization. IEEE Trans. Wirel. Commun. **15**(5), 3199–3211 (2016)
11. S. Said, L. Bombrun, Y. Berthoumieu, J. H. Manton, Riemannian Gaussian distributions on the space of symmetric positive definite matrices. IEEE Trans. Inf. Theory. **63**(04), 2153–2170 (2017)
12. S. Said, H. Hajri, L. Bombrun, B. Vemuri, Gaussian distributions on Riemannian symmetric spaces statistical learning with structured covariance matrices. IEEE Trans. Inf. Theory. **64**(02), 757–772 (2018)
13. A. Hjorungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. (Cambridge University Press, Cambridge, 2011), pp. 58–65
14. S. Boyd, L. Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, 2004), pp. 437–489
15. E. Polak, *Optimization: Algorithms and Consistent Approximations*. (Springer-Verlag, Berlin, 1997), pp. 223–275
16. X. Zhang, *Matrix Analysis and Applications*. (Tsinghua University Press, Beijing, 2004), pp. 320–365
17. C. Zhang, H. Yin, G. Wei, in *IEEE Vehicular Technology Conference (VTC Fall)*. One-sided precoder designs for interference alignment (IEEE, Quebec City, 2012), pp. 1–5
18. P. A. Absil, R. Mahony, R. Sepulchre. Optimization algorithms on matrix manifolds (Princeton University Press, New Jersey, 2008), pp. 53–69
19. K. Benidis, Y. Sun, P. Babu, D. P. Palomar, Orthogonal sparse PCA and covariance estimation via procrustes reformulation. IEEE Trans. Signal Process. **64**(23), 6211–6226 (2016)
20. C. Wang, H. Sun, S. A. Jafar, Genie chains: exploring outer bounds on the degrees of freedom of MIMO interference networks. IEEE Trans. Inf. Theory. **62**(10), 5573–5602 (2016)
21. Y. Cao, N. Zhao, F. R. Yu, M. Jin, Y. Chen, J. Tang, V. C. M. Leung, Optimization or alignment: secure primary transmission assisted by secondary networks. IEEE J. Sel. Areas Commun. **36**(4), 905–917 (2018)
22. C. Hao, B. Clerckx, Achievable sum DoF of the k-user MIMO interference channel with delayed CSIT. IEEE Trans. Commun. **64**(10), 4165–4180 (2016)
23. M. Torrellas, A. Agustin, J. Vidal, Achievable DoF-delay trade-offs for the k-user MIMO interference channel with delayed CSIT. IEEE Trans. Inf. Theory. **62**(12), 7030–7055 (2016)
24. S. Y. Yeh, I. H. Wang, Degrees of freedom of the bursty MIMO X channel without feedback. IEEE Trans. Inf. Theory. **64**(4), 2298–2320 (2018)
25. H. Huang, Y. Song, J. Yang, G. Gui, F. Adachi, Deep-learning-based millimeter-wave massive MIMO for hybrid precoding. IEEE Trans. Veh. Technol. **68**(3), 3027–3032 (2019)

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.