Beam Tracking for Interference Alignment in Time-Varying MIMO Interference Channels: A Conjugate Gradient Based Approach

Junse Lee, Heejung Yu, and Youngchul Sung†

Abstract

In this correspondence, an adaptive beam tracking algorithm for interference alignment in time-varying multiple-input and multiple-output interference channels is presented. It is shown that obtaining a set of interference-aligning transmit beamforming matrices is equivalent to minimizing a certain Rayleigh quotient, and an approach based on the conjugate gradient method combined with metric projection is applied to this minimization problem to construct an adaptive algorithm for interference-aligning beam design. The convergence of the proposed algorithm in static channels is established and the steady-state behavior of the proposed algorithm in time-varying channels is investigated by numerical simulations. The performance of the proposed algorithm is evaluated numerically and numerical results show that the proposed algorithm performs well with low computational complexity.

Index Terms

Interference alignment, adaptive algorithm, conjugate gradient, metric projection, Rayleigh quotient

I. Introduction

Since Cadambe and Jafar showed that interference alignment (IA) achieves the maximum number of degrees of freedom (DoF) in multiuser interference channels [2], many practical and efficient beam design algorithms for IA have been developed for static multiple-input and multiple-output (MIMO) interference channels, e.g., [3], [4], [5], [6]. In this correspondence, we consider the beam design for IA in time-varying MIMO interference channels. In the time-varying channel case, designing a set of interference-aligning beamforming matrices at each time step requires high computational complexity if it is designed by applying one of the existing beam

†Corresponding author

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

J. Lee is with Samsung Electronics, H. Yu is with Dept. of Information and Communication Engineering, Yeungnam University, and Y. Sung is with Dept. of Electrical Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, South Korea. Phone: +82-42-350-3484, E-mail: junselee@kaist.ac.kr, heejung@yu.ac.kr, and ysung@ee.kaist.ac.kr.

A preliminary version of this work was presented in [1].

design algorithms devised for static channels to each time step afresh. To eliminate such heavy computational burden in the time-varying case, Yu *et al.* proposed an efficient beam tracking algorithm for IA [7] based on the eigenvector perturbation theory and their work of a least squares approach to IA [6]. In their method, the beam solution at one time step is obtained as the sum of that at the reference time step and a perturbation term derived from the channel difference between the two time steps. However, the tracking method is not a purely adaptive algorithm and requires a full eigen-decomposition periodically to provide a reference beam solution to which the perturbation term is added at each time step during the tracking interval, and shows performance degradation in the case of multiple data streams per user. In this correspondence, we propose a new purely *adaptive* beam design algorithm for IA that works in both static and slowly-fading MIMO interference channels and performs well even in the multiple stream case. The new algorithm is also based on the least squares approach to IA in [6], but here we modify the conjugate gradient (CG) descent method [8] by incorporating metric projection and apply the modified CG method to obtain an updated beam solution.

Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors. For a matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^\dagger$ indicate the transpose, Hermitian transpose, and pseudo-inverse of $\mathbf{A}$, respectively. $vec(\mathbf{A})$ denotes the vector composed of the columns of $\mathbf{A}$. $\mathcal{C}(\mathbf{A})$ and $\mathcal{C}^\perp(\mathbf{A})$ denote the column space of $\mathbf{A}$ and its orthogonal complement, respectively. We use $||\mathbf{a}||$ for 2-norm of vector $\mathbf{a}$. $\mathbf{I}$ and $\mathbf{0}$ stand for the identity and all-zero matrices, respectively. The notation $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means that $\mathbf{x}$ is complex Gaussian distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

II. System Model and Preliminaries

We consider a K-pair $N_r \times N_t$ MIMO interference channel in which transmitters and receivers have N_t and N_r antennas, respectively. In this interference network, the received signal at receiver k at time n is given by

$$\mathbf{y}_k[n] = \mathbf{H}_{kk}[n]\mathbf{V}_k[n]\mathbf{s}_k[n] + \sum_{l=1,\ l\neq k}^{K} \mathbf{H}_{kl}[n]\mathbf{V}_l[n]\mathbf{s}_l[n] + \mathbf{n}_k[n], \tag{1}$$

where $\mathbf{H}_{kl}[n]$ is the $N_r \times N_t$ MIMO channel matrix at time n from transmitter l to receiver k, $\mathbf{V}_l[n]$ and $\mathbf{s}_l[n]$ are the $N_t \times d_l$ beamforming matrix and the $d_l \times 1$ signal vector at transmitter l, respectively, and $\mathbf{n}_k[n] \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ is the zero-mean complex Gaussian noise vector at

$$\tilde{\mathbf{H}}[n] \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I}_d \otimes \mathbf{H}_{12}[n] & -\mathbf{A}_{13}[n] \otimes \mathbf{H}_{13}[n] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_d \otimes \mathbf{H}_{12}[n] & \mathbf{0} & -\mathbf{A}_{14}[n] \otimes \mathbf{H}_{14}[n] & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{I}_d \otimes \mathbf{H}_{12} & \mathbf{0} & \cdots & \cdots & -\mathbf{A}_{1K}[n] \otimes \mathbf{H}_{1K}[n] \\ \mathbf{I}_d \otimes \mathbf{H}_{21}[n] & \mathbf{0} & -\mathbf{A}_{23}[n] \otimes \mathbf{H}_{23}[n] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I}_d \otimes \mathbf{H}_{21}[n] & \mathbf{0} & \mathbf{0} & -\mathbf{A}_{24}[n] \otimes \mathbf{H}_{24}[n] & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{I}_d \otimes \mathbf{H}_{21}[n] & \mathbf{0} & \mathbf{0} & \cdots & \cdots & -\mathbf{A}_{2K}[n] \otimes \mathbf{H}_{2K}[n] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{I}_d \otimes \mathbf{H}_{K1}[n] & -\mathbf{A}_{K2} \otimes \mathbf{H}_{K2}[n] & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{I}_d \otimes \mathbf{H}_{K1}[n] & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \quad (4)$$

receiver k. We assume $d_1 = \cdots = d_K = d\ (\geq 1)$ and that the channel information is known to the transmitters and the receivers. For the time-varying channel model, we consider the widely-used Gauss-Markov channel model given by [9]

$$\mathbf{H}_{kl}[n+1] = \beta \mathbf{H}_{kl}[n] + \sqrt{1-\beta^2}\mathbf{W}_{kl}[n+1], \quad (2)$$

for each (k,l), where $\beta\ (\in [0,1])$ is the fading coefficient, $vec(\mathbf{W}_{kl}[n+1]) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, and $vec(\mathbf{H}_{kl}[0]) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ ($\mathbf{W}_{kl}[n+1]$ and $\mathbf{H}_{kl}[0]$ are both independent over (k,l)).

Whereas the condition for IA is expressed as a set of bilinear equations in [10], the same condition can be expressed as *a system of linear equations with dummy variables*, given by [6]

$$\tilde{\mathbf{H}}[n]\mathbf{v}[n] = \mathbf{0}, \quad (3)$$

where $\mathbf{v}[n] \triangleq \left[vec(\mathbf{V}_1[n])^T, \cdots, vec(\mathbf{V}_K[n])^T\right]^T$ is the $KN_td \times 1$ aggregated beam vector, and $\tilde{\mathbf{H}}[n]$ is defined as (4) with size $K(K-2)N_rd \times KN_td$.

The key advantage of this formulation is that a set $\{\mathbf{V}_1[n], \cdots, \mathbf{V}_K[n]\}$ of beamforming matrices achieving IA in an IA-feasible case or achieving approximate IA in an IA-infeasible case can be obtained by minimizing $||\tilde{\mathbf{H}}[n]\mathbf{v}[n]||^2$ under a norm constraint on $\mathbf{v}[n]$. When $\{\mathbf{A}_{kl}[n]\}$ are given, the beam vectors are obtained by solving [6]

$$\min_{||\mathbf{v}[n]||=1} ||\tilde{\mathbf{H}}[n]\mathbf{v}[n]||^2 = \min_{||\mathbf{v}[n]||=1} \mathbf{v}^H[n]\mathbf{\Phi}[n]\mathbf{v}[n], \quad (5)$$

where $\mathbf{\Phi}[n] \triangleq \tilde{\mathbf{H}}^H[n]\tilde{\mathbf{H}}[n]$. When $\{\mathbf{V}_l[n]\}$ (which can be constructed from $\mathbf{v}[n]$) are given, on the other hand, the dummy variables $\{\mathbf{A}_{kl}[n]\}$ inside $\tilde{\mathbf{H}}[n]$ are given in closed form by [6]

$$\begin{cases} \mathbf{A}_{1l}[n] = (\mathbf{H}_{1l}[n]\mathbf{V}_l[n])^\dagger \mathbf{H}_{12}[n]\mathbf{V}_2[n], & l = 3, \cdots, K, \\ \mathbf{A}_{kl}[n] = (\mathbf{H}_{kl}[n]\mathbf{V}_l[n])^\dagger \mathbf{H}_{k1}[n]\mathbf{V}_1[n], & k, l = 2, \cdots, K, \ l \neq k. \end{cases} \tag{6}$$

Here, $\{\mathbf{A}_{kl}[n]\}$ is determined so that $\{\mathbf{A}_{kl}[n]\}$ is the least squares solution to minimize $\|\tilde{\mathbf{H}}[n]\mathbf{v}[n]\|^2$. Thus, a set of interference-aligning (or approximately interference-aligning in an infeasible case) beamforming matrices can be obtained by solving (5) and (6) iteratively with a proper initialization of $\mathbf{v}[n]$ for given n. It is shown in [7] that $\tilde{\mathbf{H}}[n]$ has nullity d with a properly chosen set $\{\mathbf{A}_{kl}[n]\}$ when IA is feasible and such a set can be found by solving (5) and (6) iteratively.

Note that $\mathbf{v}[n] = \left[vec(\mathbf{V}_1[n])^T, \cdots, vec(\mathbf{V}_K[n])^T\right]^T$. Thus, one might think that obtaining one null vector $\mathbf{v}_m[n]$ of $\tilde{\mathbf{H}}[n]$ (or $\mathbf{\Phi}[n]$) would yield all d beam vectors for the d streams of each user for interference alignment. However, this is not true. Due to the special structure of $\tilde{\mathbf{H}}[n]$, a null vector $\mathbf{v}_m[n]$ of $\tilde{\mathbf{H}}[n]$ has the structure of $\mathbf{v}_m[n] = [\mathbf{a}_{m1}^T \otimes \mathbf{q}_{m1}^T, \cdots, \mathbf{a}_{mK}^T \otimes \mathbf{q}_{mK}^T]^T$, where $\otimes$ is the Kroncker product, $\mathbf{a}_{mk}$ has size $d \times 1$ and $\mathbf{q}_{mk}$ has size $N_t \times 1$. Hence, the d subvectors for each user obtained from $\mathbf{v}_m[n]$ are identical after scaling. However, $\tilde{\mathbf{H}}[n]$ (or $\mathbf{\Phi}[n]$) has nullity d and hence it has d null vectors $\mathbf{v}_1[n], \cdots, \mathbf{v}_d[n]$. The d beam vectors for each user can be obtained from these d null vectors. (See [7] for detail.)

III. Adaptive Beam Tracking for Interference Alignment

In this section, we propose an adaptive algorithm for obtaining a set of interference-aligning beamforming matrices based on (5) and (6). Since the d beam vectors for each user achieving (approximate) IA are given by the d eigenvectors of $\mathbf{\Phi}[n]$ corresponding to the d smallest eigenvalues,[1] we need to find these d eigenvectors in an adaptiver manner. First, the smallest eigenvalue and the corresponding eigenvector of $\mathbf{\Phi}[n]$ under a unit-norm constraint on the eigenvector are obtained by simply solving (5). On the other hand, finding the second smallest and following eigenvalues and eigenvectors needs more elaboration. Note that $\mathbf{\Phi}[n]$ is a Hermitian matrix and thus its eigenvectors are orthonormal by the spectral theorem. Hence, the eigenvector

[1]The same eigenvalue is counted according to its geometric multiplicity.

corresponding to the j-th smallest eigenvalue is obtained by solving

$$\min_{\mathbf{v}[n]:\mathbf{v}[n]\perp\{\check{\mathbf{v}}_1[n],\cdots,\check{\mathbf{v}}_{j-1}[n]\},\|\mathbf{v}[n]\|=1} \mathbf{v}^H[n]\mathbf{\Phi}[n]\mathbf{v}[n], \tag{7}$$

where $\check{\mathbf{v}}_1[n],\cdots,\check{\mathbf{v}}_{j-1}[n]$ are the $j-1$ eigenvectors of $\mathbf{\Phi}[n]$ corresponding to the $j-1$ smallest eigenvalues. Thus, (5) and (7) should be solved in an efficient adaptive way. To this end, first note that finding the smallest eigenvalue and the corresponding eigenvector of $\mathbf{\Phi}[n]$ in (5) under the unit-norm constraint on the eigenvector is equivalent to obtaining the minimum value of the Rayleigh quotient

$$R(\mathbf{\Phi}[n],\mathbf{v}[n]) \triangleq \frac{\mathbf{v}[n]^H\mathbf{\Phi}[n]\mathbf{v}[n]}{\mathbf{v}^H[n]\mathbf{v}[n]}. \tag{8}$$

This Rayleigh quotient minimization problem can be solved adaptively by applying a gradient descent method. Among various descent methods, we adopt the *conjugate gradient (CG) descent* which is suitable for Hermitian $\mathbf{\Phi}[n]$, does not require matrix inversion as the Newton method, and shows fast convergence [8]. On the other hand, in the problem (7), the minimization under the unit-norm constraint part is equivalent to (8) but we have an additional constraint that $\mathbf{v}[n]$ is constained in the orthogonal complement of the span of $\check{\mathbf{v}}_1[n],\cdots,\check{\mathbf{v}}_{j-1}[n]$. To solve this problem adaptively, we apply projection to the gradient descent [11] and propose a projected conjugate gradient method that consists of two steps: the first step is a CG descent step for cost reduction and the second step is projection of the CG step output onto the orthogonal complement of the span of $\check{\mathbf{v}}_1[n],\cdots,\check{\mathbf{v}}_{j-1}[n]$. Here, tracking of each of the d smallest eigenvectors is performed sequentially. That is, the $j-1$ eigenvectors $\check{\mathbf{v}}_1[n],\cdots,\check{\mathbf{v}}_{j-1}[n]$ for the adaptive tracking of the j-th eigenvector come from the adaptive tracking of the first $j-1$ eigenvectors. Together with the dummy variable update (6), the CG descent applied to the minimization of (8) combined with projection provides an efficient adaptive beam design algorithm for IA in static and time-varying channels, which is described in Algorithm 1. The CG subroutine from [8] is modified to include the projection step and described below:

CG subroutine ($\mathbf{v}$, $\mathbf{\Phi}$, $\mathbf{\Pi}_{\mathbf{S}}^{\perp}$, N_1)

Initialization: $\mathbf{x}(0)=\mathbf{v}$, $b(0)=0$, and $\lambda(0)=\frac{\mathbf{x}(0)^H\mathbf{\Phi}\mathbf{x}(0)}{||\mathbf{x}(0)||^2}$

for $k=0,1,\cdots,N_1-1$

Step 1. If $k=0$, then $\mathbf{r}(0)=\mathbf{p}(0)=\frac{\lambda(0)\mathbf{x}(0)-\mathbf{\Phi}\mathbf{x}(0)}{||\mathbf{x}(0)||^2}$.

Step 2. Compute $t(k)=\frac{(-B+\sqrt{B^2-4CD})}{2D}$, where $B=\frac{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{p}(k)}{||\mathbf{x}(k)||^2}-\lambda(k)\frac{\mathbf{p}(k)^H\mathbf{p}(k)}{||\mathbf{x}(k)||^2}$, $C=\frac{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{x}(k)}{||\mathbf{x}(k)||^2}-\lambda(k)\frac{\mathbf{p}(k)^H\mathbf{x}(k)}{||\mathbf{x}(k)||^2}$, and $D=\frac{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{p}(k)}{||\mathbf{x}(k)||^2}\frac{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{x}(k)}{||\mathbf{x}(k)||^2}-\frac{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{x}(k)}{||\mathbf{x}(k)||^2}\frac{\mathbf{p}(k)^H\mathbf{p}(k)}{||\mathbf{x}(k)||^2}$.

Algorithm 1 The Conjugate Gradient Algorithm for Interference Alignment (CGIA)

Require: Initialize $\{\mathbf{A}_{kl}[0]\}$ and $\hat{\mathbf{v}}_1[0], \cdots, \hat{\mathbf{v}}_d[0]$.

while $n = 0, 1, \cdots$ **do**

Construct $\mathbf{\Phi}[n] = \tilde{\mathbf{H}}^H[n]\tilde{\mathbf{H}}[n]$ with $\{\mathbf{A}_{kl}[n-1]\}$ and $\{\mathbf{H}_{kl}[n]\}$.

Update $\hat{\mathbf{v}}_1[n], \cdots, \hat{\mathbf{v}}_d[n]$ as follows.

$\mathbf{S} = [\ ]$

for $m = 1$ to d **do**

$\mathbf{\Pi}_{\mathbf{S}}^{\perp} = (\mathbf{I} - \mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H)$

$\hat{\mathbf{v}}_m[n] = \text{CG subroutine}(\hat{\mathbf{v}}_m[n-1], \mathbf{\Phi}[n], \mathbf{\Pi}_{\mathbf{S}}^{\perp}, N_1)$

$\mathbf{S} = [\mathbf{S}, \hat{\mathbf{v}}_m[n]]$

end for

Obtain $\{\mathbf{V}_k[n]\}$ from $\hat{\mathbf{v}}_1[n], \cdots, \hat{\mathbf{v}}_d[n]$. (Step *) (See [7] for this step.)

If $\text{mod}(n, N_2) = 0$, then update $\{\mathbf{A}_{kl}[n]\}$ by (6). Otherwise, $\{\mathbf{A}_{kl}[n] = \mathbf{A}_{kl}[n-1]\}$.

end while

Step 3. Update the desired vector: $\mathbf{x}(k+1) = \mathbf{x}(k) + t(k)\mathbf{p}(k)$

Step 4. Projection: $\mathbf{x}(k+1) = \mathbf{\Pi}_{\mathbf{S}}^{\perp}\mathbf{x}(k+1)$

Step 5. Compute $\lambda(k+1) = \frac{\mathbf{x}(k+1)^H\mathbf{\Phi}\mathbf{x}(k+1)}{||\mathbf{x}(k+1)||^2}$

Step 6. Obtain the residual: $\mathbf{r}(k+1) = \frac{\lambda(k+1)\mathbf{x}(k+1) - \mathbf{\Phi}\mathbf{x}(k+1)}{||\mathbf{x}(k+1)||^2}$

Step 7. Update the direction: $\mathbf{p}(k+1) = \mathbf{r}(k+1) - \frac{\mathbf{r}(k+1)^H\mathbf{\Phi}\mathbf{p}(k)}{\mathbf{p}(k)^H\mathbf{\Phi}\mathbf{p}(k)}\mathbf{p}(k)$

Output: $\mathbf{v}' = \mathbf{x}(N_1)$

In Algorithm 1, a subvector normalization step can be added to Step * without disturbing the solution structure when IA is feasible. This is easy to see in the case of $d = 1$. (In this case, $\mathbf{A}_{kl} = a_{kl}$ simply.) Suppose that $\underline{\mathbf{v}} = [\underline{\mathbf{v}}_1^T, \cdots, \underline{\mathbf{v}}_K^T]^T$ and $\{\underline{a}_{kl}\}$ are a solution to (3). Then, $\underline{\mathbf{v}}' = [\eta_1\underline{\mathbf{v}}_1^T, \cdots, \eta_K\underline{\mathbf{v}}_K^T]^T$ and $\{\eta_l\underline{a}_{kl}/\eta_j, j = 2 \text{ if } k = 1,\ j = 1 \text{ if } k \neq 1\}$ are also a solution to (3). This is also valid in the case of $d > 1$. In CGIA, we have freedom to design (N_1, N_2), where N_1 is the number[2] of CG updates per time step, and N_2 is the period of dummy variable updates.

[2]When $N_1 = 1$, the CG step is simple gradient descent. Thus, when $N_1 = 1$, the proposed method for each of the d smallest eigenvectors reduces to the projected gradient method of Goldstein [11]. The proposed algorithm here can be used to general multiple extreme eigenvector tracking for a Rayleigh quotient beyond the considered problem here.

N_1 and N_2 should be properly designed to yield a desired trade-off between performance and complexity.

IV. Analysis of CGIA

In this section, we investigate the properties of the proposed CGIA algorithm. (For notational simplicity, the time index is omitted if unnecessary.) First, note that CGIA updates the beam vectors so that the Rayleigh quotient (8) (or, equivalently, $||\tilde{\mathbf{H}}[n]\mathbf{v}[n]||^2$ under the unit norm constraint on $\mathbf{v}[n]$) is minimized. However, the interference metric of interest is the interference leakage γ_k at receiver k defined as the portion of the total interference power leaking into the signal space [10], i.e., $\gamma_k \triangleq \sum_{i=d+1}^{N_r} \lambda_i(\mathbf{\Gamma}_k)/\sum_{i=1}^{N_r} \lambda_i(\mathbf{\Gamma}_k)$, where $\lambda_1(\mathbf{\Gamma}_k) \geq \cdots \geq \lambda_{N_r}(\mathbf{\Gamma}_k)$ are the ordered eigenvalues of the $N_r \times N_r$ interference covariance matrix $\mathbf{\Gamma}_k = \sum_{l\neq k} \mathbf{H}_{kl}\mathbf{V}_l\mathbf{V}_l^H\mathbf{H}_{kl}^H$ at receiver k. Here, the subspace spanned by the eigenvectors corresponding to $\lambda_1(\mathbf{\Gamma}_k), \cdots, \lambda_d(\mathbf{\Gamma}_k)$ is assumed to be the interference subspace, and the remaining subspace corresponding to $\lambda_{d+1}(\mathbf{\Gamma}_k), \cdots, \lambda_{N_r}(\mathbf{\Gamma}_k)$ is assumed to be the subspace intended for the desired signal. As the Rayleigh quotient given by (8) decreases, it is desirable for the interference leakage also to decrease. This desired property is shown in the following proposition.

Proposition 1: For general K and d, if the Rayleigh quotient R in (8) goes to zero, then the interference leakage γ_k at receiver k goes to zero for all $k = 1, \cdots, K$.

Proof: See the appendix.

Thus, by making the Rayleigh quotient (8) small we can make the interference leakage at each receiver small. With the desired property assured, we now investigate the convergence property of CGIA. The convergence of CGIA in static channels is established in the following proposition.

Proposition 2: The CGIA algorithm converges for any initial condition and (N_1, N_2) for time-invariant channels.

Proof: See the appendix.

Since CGIA converges for any initialization in static channels, CGIA is stable in static channels. Next, we consider the stability and steady-state behavior of the algorithm in time-varying channels. In the case of standard CG methods, the stability was analyzed in [12]. However, the existing analysis approaches cannot be applied to the proposed CGIA algorithm since it includes not only the CG step but also the dummy variable update step. A rigorous proof of stability in time-varying channels is difficult. However, in [13], under the first order AR channel model (2), the stability of CGIA is analyzed in the case of $d = 1$ ($\mathbf{A}_{kl}[n] = a_{kl}[n]$) under several strong assumptions by showing that the Rayleigh quotient does not increase as time elapses. Here, we briefly explain the idea. At the end of time step n, we have $\{\mathbf{v}[n], a_{kl}[n]\}$, and the corresponding minimum Rayleigh quotient is determined by $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n],a_{kl}[n])} \triangleq \tilde{\mathbf{H}}^H_{(\mathbf{H}_{kl}[n],a_{kl}[n])}\tilde{\mathbf{H}}_{(\mathbf{H}_{kl}[n],a_{kl}[n])}$, where $\tilde{\mathbf{H}}_{(\mathbf{H}_{kl}[n_1],a_{kl}[n_2])}$ denotes the matrix $\tilde{\mathbf{H}}$ in (4) constructed with $\mathbf{H}_{kl}[n_1]$ and $a_{kl}[n_2]$. At time step $n+1$, first the matrix $\tilde{\mathbf{H}}$ is perturbed to become $\tilde{\mathbf{H}}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$, and then CGIA updates the beam vector as $\mathbf{v}[n+1]$ by finding the minimum eigenvalue and the corresponding eigenvector of $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$ with the CG step. After this CG step, the minimum Rayleigh quotient or equivalently the minimum eigenvalue of $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$ may increase from that of $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n],a_{kl}[n])}$. However, the following dummy variable update step always reduces the minimum Rayleigh quotient by updating $\{a_{kl}[n+1]\}$ optimally. At the end of time step $n+1$, the minimum Rayleigh quotient is given by that determined by $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n+1])}$. Thus, if the increase in the minimum Rayleigh quotient caused by the change from $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n],a_{kl}[n])}$ to $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$ after the channel variation/CG step is compensated for by the decrease in the minimum Rayleigh quotient caused by the change from $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$ to $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n+1])}$ after the dummy variable update step, the algorithm is stable and shows the steady-state behavior when these two quantities are equal. The stability condition can be summarized as

$$c_1\sqrt{1-\beta} \le c_2 R(\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}, \mathbf{v}[n+1])) + \delta \tag{9}$$

for some $\delta \ge 0$, where the left-hand side (LHS) term in (9) denotes an upper bound on the Rayleigh quotient increase in the channel variation/CG step obtained by applying perturbation theory to the eigenvalues of $\mathbf{\Phi}_{(\mathbf{H}_{kl}[n+1],a_{kl}[n])}$ and the right-hand side (RHS) term of (9) denotes the reduction in the Rayleigh quotient by the dummy variable update step obtained by a geometrical interpretation of the update of the dummy variables under several assumptions. With the expression, it can be seen that the faster fading rate forms the higher level of steady state Rayleigh quotient (or equivalently the interference leakage). Although some strong assumption in [13] for obtaining (9) may not be justified, simulations show that CGIA is indeed stable and shows a good steady-state behavior in most time-varying channels and the steady-state interference leakage increases with the mobile speed, as shown in Section V.

Finally, the complexity of CGIA is analyzed in terms of the number of complex multiplications. The main advantage of CGIA over the existing approaches is computational efficiency. The computational complexity of CGIA and other algorithms including the perturbation approach [7] and the interative interference alighment (IIA) algorithm [10] is shown in Table I. For CGIA, we can make a trade-off between complexity and performance by adjusting parameters (N_1, N_2). For example, in the case of low operating SNR and mobile speeds, it is not necessary to use large N_1. Even with the small number of CG steps, i.e., small N_1, the algorithm can achieve a residual interference level lower than the thermal noise for a normal range of operating SNR, when IA is feasible.

V. Numerical results

In this section, we provide some numerical results to evaluate the performance of CGIA. Throughout the simulations, we generated a first-order Gauss-Markov channel process described in (2) for each user independently with 1GHz carrier frequency and 66.7μs symbol duration (the symbol duration of 3GPP LTE) and evaluated the performance of CGIA.

First, to see the convergence speed of CGIA, we ran the algorithm in two cases: (a) a single stream case of $K = 4$, $N_t = 3$, $N_r = 2$, $d = 1$ for which IA is feasible but does not have a closed-form solution and (b) a two stream case of $K = 3$, $N_t = N_r = 4$, $d = 2$. Figs. 1 (a) and (b) show the interference leakage obtained by CGIA as time elapses in the two cases. It is seen that CGIA converges and then reaches the steady state fast in both the single-stream and two-stream cases. As expected, the steady-state leakage level is formed at a higher level for a higher mobile speed. The attained leakage level also depended on (N_1, N_2). The (N_1, N_2) values shown in the figure were chosen to yield a sufficiently low leakage level around $10^{-4} \sim 10^{-5}$ at low mobile speeds.

Fig. 2 shows an example of the complexity of several beam design methods for IA including the IIA algorithm in [10], the iterative LS (ILS) algorithm in [6], and the tracking algorithm based on a perturbations approach in [7] with an eigen-decomposition every 10 symbols, in the same setup as that in Fig. 1 (a). (The slope of IIA corresponds to the case of 100 iterations per time step. 100 iterations per time step showed reasonable convergence in the considered case.) As seen in Fig. 2, the perturbations approach has the smallest slope; the slopes of the non-recursive methods are not comparable to the methods exploiting the channel coherence. However, the advantage of CGIA over the perturbations approach is shown next.

Fig. 3 shows the sum rate performance of CGIA with respect to SNR in three cases: $K = 3$, $N_t = N_r = 2$, $d = 1$; $K = 4$, $N_t = 3$, $N_r = 2$, $d = 1$; and $K = 3$, $N_t = N_r = 4$, $d = 2$. In all the cases, the non-recursive IIA algorithm in [3] run at the static channel was used as a performance reference. In this method, we ran the IIA algorithm with 1000 iterations allowing sufficient convergence. First, the result for the two single-stream cases is shown in Fig. 3 (a). It is seen in Fig. 3 (a) that CGIA yields almost the same performance as the non-recursive IIA algorithm in the state channel case (i.e., 0 km/h). (In the static channel case, the two methods used the same channel.) It is also seen that the performance degradation of CGIA due to the mobile speed is not significant, when the mobile speed is low, and the performance degradation is noticeable at high SNR and high mobile speed. Next, the result for the two-stream case is shown in Fig. 3 (b). Here, we considered the non-recursive IIA algorithm again as a performance reference, the tracking algorithm based on eigenvector perturbation in [7] (denoted as the least squares with iteration and tracking (LSINT) algorithm), and CGIA. For LSINT, an eigen-decomposition is applied every 10 symbols. In the static case, the three algorithms used the same channel. It is seen in the static-channel two-stream case that CGIA performs almost the same as IIA whereas LSINT performs a bit worse than the other two algorithms. It is seen that on the contray to the single-stream case, the performance degradation due to mobile speed increase is negligible for CGIA even at high SNR in the two-stream case. However, LSINT shows severe performance degradation as the mobile speed increases. Althought it is not shown here due to space limitations, LSINT performs similarly well in the single-stream case. This means that multi-dimensional subspace tracking based on eigen-space perturbation used for LSINT is sensitive and error accumulates quickly as time elapses. Thus, CGIA is advantageous for IA beam tracking in multi-stream cases.

VI. Conclusions

In this correspondence, we have proposed CGIA for transmit beam tracking for interference alignment in time-varing MIMO interference channels. We have established the convergence of CGIA for static channels and have investigated its steady-state behavior numerically in the time-varying channel case. Numerical results show that CGIA converges fast and performs well for time-varying MIMO interference channels with significantly reduced computational complexity. CGIA provides an alternative adaptive algorithm for interference alignment with significant complexity reduction and comparable sum rate performance in time-varying MIMO interference channels.

APPENDIX

Proof of Proposition 1: The Rayleigh quotient R is equivalent to $||\tilde{\mathbf{H}}\mathbf{v}||^2$ under the constraint $||\mathbf{v}|| = 1$. Exploiting the structure of $\tilde{\mathbf{H}}$ in (4), we can rewrite $||\tilde{\mathbf{H}}\mathbf{v}||^2$ as

$$R = ||\tilde{\mathbf{H}}\mathbf{v}||^2 = \sum_{l=3}^{K} ||\mathbf{H}_{12}\mathbf{V}_2 - \mathbf{H}_{1l}\mathbf{V}_l\mathbf{A}_{1l}^T||^2 + \sum_{k=2}^{K} \sum_{l=1,l\neq 1,l\neq k}^{K} ||\mathbf{H}_{k1}\mathbf{V}_1 - \mathbf{H}_{kl}\mathbf{V}_l\mathbf{A}_{kl}^T||^2. \quad (10)$$

(Basically, the IA condition (3) is obtained so as to align the interference from an unwanted transmitter to receiver 1 to the reference subspace $\mathbf{H}_{12}\mathbf{V}_2$ and to align the interference from an unwanted transmitter to receiver $k(\neq 1)$ to the reference subspace $\mathbf{H}_{k1}\mathbf{V}_1$. Note that the matrix $\mathbf{A}_{kl}$ is for subspace equivalence. [6]) Now, as $R \downarrow 0$, the interference aligns since each term in the RHS of (10) is non-negative and hence, the rank of the $N_r \times N_r$ interference covariance matrix $\mathbf{\Gamma}_k$ at receiver k, given by

$$\mathbf{\Gamma}_k = \sum_{i\neq k}^{K} \mathbf{H}_{ki}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ki}^H,$$

becomes d eventually. Therefore, the smallest $N_r - d$ eigenvalues of $\mathbf{\Gamma}_k$ goes to zero as $R \downarrow 0$. So does the interference leakage γ_k at receiver k for all k by the definition of γ_k. ■

Proof of Proposition 2: First, consider the tracking of the smallest eigenvalue and the corresponding eigenvector. For given $\mathbf{H}_{kl}[n]$ and $\mathbf{A}_{kl}[n]$, the CG update, which computes the new beam vector $\hat{\mathbf{v}}_1[n]$ minimizing the Rayleigh quotient, does not increase the Rayleigh quotient, i.e.,

$$R\left(\mathbf{\Phi}(\mathbf{H}_{kl}[n], \mathbf{A}_{kl}[n-1]), \hat{\mathbf{v}}_1[n-1]\right) \geq R\left(\mathbf{\Phi}(\mathbf{H}_{kl}[n], \mathbf{A}_{kl}[n-1]), \hat{\mathbf{v}}_1[n]\right).$$

Furthermore, given $\{\mathbf{V}_l[n]\}$ constructed from $\hat{\mathbf{v}}_1[n], \cdots, \hat{\mathbf{v}}_d[n]$, it was proved in [6] that the dummy variable update (6) does not increase the Rayleigh quotient since (6) itself is the least squares solution to minimize the Rayleigh quotient as a function of $\mathbf{A}_{kl}$, i.e.,

$$R(\mathbf{\Phi}(\mathbf{H}_{kl}[n], \mathbf{A}_{kl}[n-1]), \hat{\mathbf{v}}_1[n]) \geq R(\mathbf{\Phi}(\mathbf{H}_{kl}[n], \mathbf{A}_{kl}[n]), \hat{\mathbf{v}}_1[n]).$$

By the two inequalities, the Rayleigh quotient for the smallest eigenvector monotonically decreases for CGIA, regardless of the value of (N_1, N_2), but the Rayleigh quotient is lower bounded by zero because $\mathbf{\Phi}$ is a semi-positive definite matrix. Thus, the smallest eigenvector and the corresponding eigenvector of CGIA converges by the monotone convergence theorem. Next, consider the second smallest eigenvalue and the corresponding eigenvector $\hat{\mathbf{v}}_2[n]$. Since the first eigvenvector $\hat{\mathbf{v}}_1[n]$ converges, the subspace $\mathcal{C}^{\perp}(\hat{\mathbf{v}}_1[n])$ converges too. Then, we can apply the same monotone convergence argument to the tracking of the second smallest eigenvalue and the corresponding eigenvector used in the proof of the convergence of the first eigenvalue and eigenvector, since the tracking of the second eigenvalue and eigenvector is the same as that of the first eigenvalue and eigenvector except that the space is confined in $\mathcal{C}^{\perp}(\hat{\mathbf{v}}_1[n])$. We can apply the same argument sequentially[3] to the tracking of the m-th smallest eigenvalue and eigenvector for $m \leq d < \infty$. ■

REFERENCES

[1] J. Lee, H. Yu, Y. Sung, and Y. H. Lee, "Adaptive beam tracking for interference alignment for time-varying MIMO interference channels: Conjugate gradient approach," in *Proc. ICASSP*, Prague, Czech Republic, May 2011.

[2] 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, pp. 3425 – 3441, Aug. 2008.

[3] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3309 – 3322, June 2011.

[4] D. A. Schmidt, C. Shi, R. Berry, M. L. Honig, and W. Utschick, "Minimum Mean Squared Error Interference Alignment," in *Proc. Asilomar*, Pacific Grove, CA, Nov. 2009.

[5] S. W. Peters and R. W. Heath, "Interference alignment via alternating minimization," in *Proc. IEEE ICASSP*, Taipei, Taiwan, Apr. 2009.

[6] H. Yu and Y. Sung, "Least squares approach to joint beam design for interference alignment in multiuser multi-input multi-output interference channels," *IEEE Trans. Signal Process.*, vol. 58, pp. 4960 – 4966, Sep. 2010.

[7] H. Yu, Y. Sung, H. Kim, and Y. H. Lee, "Beam tracking for interference alignment in slowly-fading MIMO interference channels: Perturbations approach under a linear framework," *IEEE Trans. Signal Process.*, vol. 40, pp. 1910 – 1926, Apr. 2012.

[8] X. Yang, T. K. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problems of a symmetric matrix," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 37, pp. 1550 – 1556, Oct. 1989.

[9] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, pp. 1065– 1076, May 2002.

[10] K. Gomadam, V. Cadambe, and S. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pp. 1–6, 30 2008-Dec. 4.

[11] A. A. Goldstein, "Convex programming in Hilbert space," *Bull. Amer. Math. Soc.*, vol. 70, pp. 709–710, 1964.

[12] P. S. Chang and J. A. N. Willson, "Analysis of conjugate gradient algorithms for adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, pp. 409 – 418, Feb. 2000.

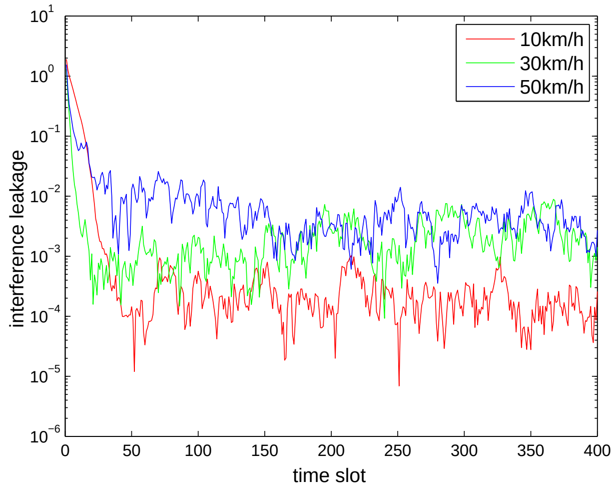
[13] J. Lee, H. Yu, and Y. Sung, "Interference alignment based conjugate gradient approach in time-varying channels," *WISRL Technical Report*, Mar. 2012.

[3] Rigorous ϵ-δ statements are omitted for simplicity since the idea of the proof is clear.

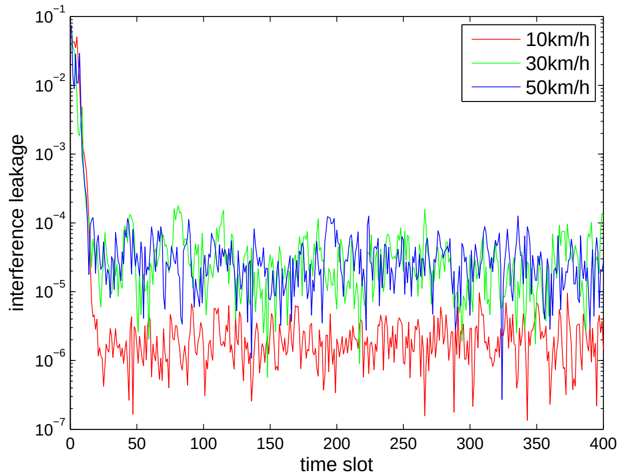
TABLE I

COMPUTATIONAL COMPLEXITY FOR THE K-PAIR (N_r, N_t) MIMO IC WITH d DATA STREAMS PER USER (J IS THE NUMBER OF ITERATIONS TO OBTAIN INITIAL BEAM VECTORS IN THE EIGENDECOMPOSITION PHASE OF THE PERTURBATION APPROACH AND L_{QR} IS THE NUMBER OF ITERATIONS FOR THE ITERATIVE QR ALGORITHM.)

Algorithm	Major Computation	Complex multiplications
CGIA	Compute $\mathbf{\Phi} = \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$	$\frac{3}{2}\left(K(K-1)(K-2)\frac{N_t d(N_t d+1)}{2} N_r d\right) + K(K-2)N_t N_r d^2$
	Compute B,C,D in CG subroutine	$N_1 \times (2KN_t d + 2)$
	Compute $t(k)$ in CG subroutine	$N_1 \times (4KN_t d + K^2 N_t^2 d^2)$
	Projection step in CG subroutine	$N_1 \times ((d+1)(2d^2 - 5d + 6)/6 + KN_t d(2d^2 + 6d - 11)/6 + \Sigma_{k=2}^{d} k!)$
	Compute λ in CG subroutine	$N_1 \times (2KN_t d + K^2 N_t^2 d^2)$
	Compute $\mathbf{r}$ in CG subroutine	$N_1 \times (KN_t d)$
	Determine $\mathbf{A}_{kl}$ (every N_2 time steps)	$N_t N_r d + 3N_r d^2 + (d+1)! + d^2$
Perturbations approach [7]	- Eigen-decomposition phase -	
	Compute $\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$	$J\left\{\frac{3}{2}\left(K(K-1)(K-2)\frac{N_t d(N_t d+1)}{2} N_r d\right) + K(K-2)N_t N_r d^2\right\}$
	Iterative QR method	$(J-1)\left\{L_{QR}(K^2 N_t^2 d^2)\right\}$
	Determine $\{\mathbf{A}_{kl}^{(i)}\}$	$(J-1)\left\{N_t N_r d + 3N_r d^2 + (d+1)! + d^2\right\}$
	Eigen-decomposition	$\frac{13}{3}(KN_t d)^3$
	- Tracking phase -	
	Construct $\mathbf{G}_m[n]$	$\frac{3}{2}\left(K(K-1)(K-2)\frac{N_t d(N_t d+1)}{2} N_r d\right) + K(K-2)N_t N_r d^2$
	Update	$2K(KN_t d - d)N_t d + (KN_t d)^2$
IIA [3]	All computation	$K((2N_r N_t d + (N_r^2 + N_t^2)d)(K-1)$ $+L_{QR}(N_t d(N_t - d + 1) + N_r d(N_r - d + 1) + \frac{2d^3 - 3d^2 + d}{3}))$



(a)



(b)

Fig. 1. Interference leakage of CGIA for several mobile speeds: (a) $K = 4$, $N_t = 3$, $N_r = 2$, $d = 1$ $(N_1 = 30,\ N_2 = 1)$ and (b) $K = 3$, $N_t = N_r = 4$, $d = 2$, $(N_1 = 100, N_2 = 1)$

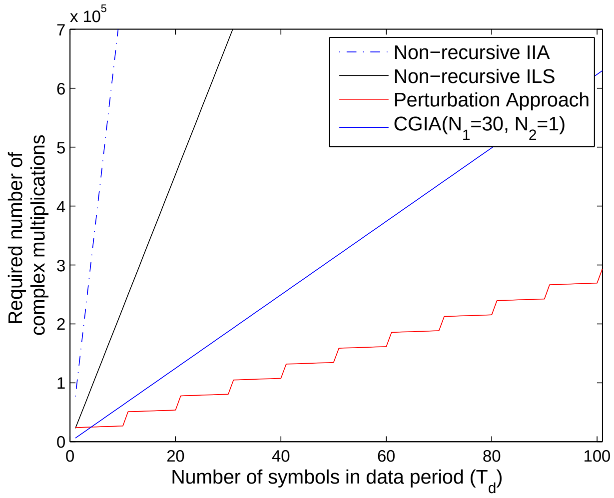
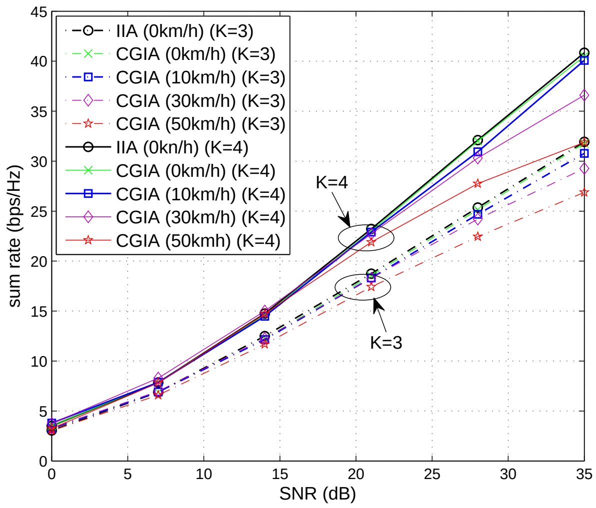
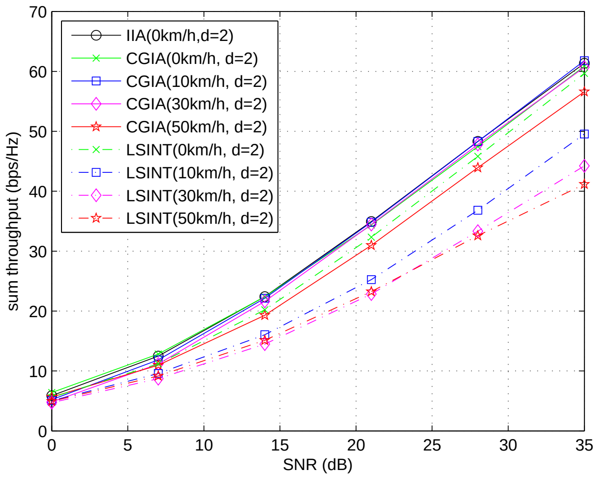


Fig. 2. Computational complexity example: $K = 4$, $N_t = 3$, $N_r = 2$, $d = 1$ ($N_1 = 30$, $N_2 = 1$)



(a)



(b)

Fig. 3. Sum rate performance: CGIA versus the IIA algorithm in [10]: (a) $K = 3$, $N_t = N_r = 2$, $d = 1$ ($N_1 = 6,\ N_2 = 1$) and $K = 4$, $N_t = 3$, $N_r = 2$, $d = 1$ ($N_1 = 30,\ N_2 = 1$) and (b) $K = 3,\ N_t = N_r = 4,\ d = 2$ ($N_1 = 100,\ N_2 = 1$)