

# Achieving Linear Scaling with Interference Alignment

Ayfer Özgür

Faculté Informatique et Communications  
Ecole Polytechnique Fédérale de Lausanne  
1015 Lausanne, Switzerland  
ayfer.ozgur@epfl.ch

David Tse

Department of EECS  
University of California at Berkeley  
Berkeley, CA 94720, USA  
dtse@eecs.berkeley.edu

**Abstract**—Recent results have shown that interference alignment can achieve  $K/2$  degrees of freedom in a  $K$ -user interference channel with time or frequency varying channel coefficients. For fixed number of users  $K$ , the number of degrees of freedom characterizes the asymptotic behavior of the performance in the high SNR limit but it does not answer the question of how the performance scales with  $K$  for any fixed SNR. In particular, it is unclear if a constant rate per user can be maintained as more users enter into the system. In this paper, we investigate the performance of the interference alignment scheme proposed in [5] for fixed SNR. We assume that the channel coefficients between the users are of the form  $r e^{j\theta}$  where  $r$  is fixed over the duration of communication and  $\theta$  is a fast fading phase. We show that for any value of the SNR and  $K$ , the aggregate rate achieved by the interference alignment scheme of [5] is lower bounded by  $c_1 K \log(1 + c_2 \text{SNR})$  where  $c_1$  and  $c_2$  are positive constants independent of both SNR and  $K$ . This result establishes the linear scaling of the interference alignment scheme for the considered random phase channel model.

## I. INTRODUCTION

There has been a lot of effort in developing efficient communication schemes for wireless networks in the last decade. The central phenomenon is dealing with interference when multiple users try to communicate over a shared medium. The state-of-the-art approaches of orthogonalizing the communication links or treating interference as noise become inefficient especially when the number of users that share the same spectrum is large. The scaling law formulation proposed in [1] and later studied by [2] highlights this aspect by seeking for network communication schemes that scale in system size. The key ideas that emerge from this line of research are spatial reuse, cooperation between users and distributed MIMO communication.

Apart from this line of research, the recent idea of interference alignment suggested by [3], [4], [5] introduces a new approach to cope with interference. Users design their signals so that they align at the receivers where they constitute interference while the desired signal at every receiver remains distinguishable from the interference. Cooperation between users is still the key but spatial reuse is not needed as in [1], [2]. In particular, [5] constructs an interference alignment scheme for the well-studied  $K$ -user interference channel that achieves  $K/2$  degrees of freedom when the channel is time-varying.  $K/2$  degrees of freedom imply linear scaling in the

number of users  $K$  (with pre-constant  $1/2$ ) in the high SNR limit. However it is not clear if this linear scaling kicks in at any finite SNR. The SNR required to enter the linear scaling regime may be increasing with  $K$ . (Imagine, for example, the case when the aggregate performance of a scheme is of the form  $\frac{K}{2} \log(1 + \text{SNR}/K^2)$  or  $\frac{K}{2} \log(1 + \text{SNR}/e^K)$ . The degrees of freedom are  $K/2$  in both cases but the performance quickly vanishes with increasing  $K$  for any fixed SNR.)

The interference alignment scheme in [5] is constructed over multiple time-realizations of the fast-fading interference channel. The signalling vectors for each transmitter (formed over multiple channel uses) are designed carefully so that they compact into small dimensional subspaces at the receivers where they are not desired. This requirement leads to a very particular structure for the signalling vectors as functions of the channel coefficients. In order to approach the theoretical upperbound of  $K/2$  degrees of freedom, the length of the signalling vectors should grow large, in other words, we need to code over many realizations of the interference channel. In particular, to achieve  $cK/2$  degrees of freedom per dimension for a constant  $0 < c < 1$ , the coding length should increase roughly as  $K^{2K^2}$  in the number of users  $K$ . Is there an associated cost in terms of SNR for increasing the block length? This question is out of the scope of a degrees-of-freedom discussion since degrees of freedom relate to the limit when SNR goes to infinity. In particular, it has been shown in [5] that for their construction of the signalling subspace, the interference subspace at each receiver remains distinct from the desired signal space with high probability for any block length. However, it is not clear if a constant angle between the two subspaces can be maintained for increasing block-length which otherwise would imply an associated cost in terms of SNR. This is the question that motivates this paper and that we are able to answer in the affirmative for the random phase channel model.

A similar result is presented in a concurrent work [6]. The result in [6] is stronger than the result of the current paper in the sense that it establishes linear scaling for a larger class of channel fading distributions with a tighter lowerbound while here we restrict attention to the random phase channel model. The main difference between [6] and the current paper is that [6] constructs a new interference alignment scheme that relies

on the ergodicity of the channel. In this paper we consider the original interference alignment scheme proposed in [5]. We believe linear scaling for the scheme in [5] is of interest in its own right since this scheme can offer an advantage in terms of the number of dimensions required to achieve the same performance. In particular, to achieve an aggregate rate  $\frac{cK}{2} \log \text{SNR}$  for a constant  $0 < c < 1$ , the ergodic scheme of [6] requires an order of  $(K^2 \text{SNR}^c)^{K^2}$  independent realizations of the fading interference channel while the modified version of the interference alignment scheme of [5], that we discuss in Section IV, requires an order of  $2^{K^2}$  independent channel realizations.

In the next section, we introduce the random phase channel model considered in this paper. Section III contains our main result and Section V is devoted to its proof. Section IV contains an overview of a modified version of the interference alignment scheme presented in [5]. Section VI provides a discussion on the extension of our result to other channel models. Notationwise, we use lower case for scalars, bold font lower case to denote vectors and upper case to denote matrices. Vectors are column vectors by default.  $H^\dagger$  denotes the complex conjugate transpose of the matrix  $H$ ,  $H(p, t)$  denotes its entry on the  $p$ 'th row and  $t$ 'th column,  $H(p, t)^*$  is the complex conjugate of  $H(p, t)$ . The trace of the matrix  $H$  is denoted by  $\text{tr}(H)$ .  $I_d$  is the  $d \times d$  identity matrix.  $C(n, k)$  denotes the choose function of  $n$  and  $k$ .

## II. MODEL

We consider a  $K$ -user interference network composed of  $2K$  nodes paired up into  $K$  pairs. Transmitter  $k$  wants to communicate to receiver  $k$ ,  $k \in \{1, \dots, K\}$ , in the presence of interference from the other users. The signal received by receiver  $k$  is given by

$$y_k = h_{k1}x_1 + h_{k2}x_2 + \dots + h_{kK}x_K + z_k$$

where  $x_1, \dots, x_K$  are the corresponding signals transmitted by transmitters  $1, \dots, K$ . Each of these signals is constrained to an average power of  $P$  Watts.  $z_k$  is the circularly symmetric white Gaussian noise at receiver  $k$  of unit variance. The channel coefficient  $h_{kl}$  between transmitter  $l$  and receiver  $k$  is of the form

$$h_{kl}^{(b)} = r_{kl} e^{j\theta_{kl}^{(b)}} \quad (1)$$

where the magnitude  $r_{kl}$  is an arbitrary positive and finite number, fixed over the duration of communication. We assume a block fading model for the phase. The phase  $\theta_{kl}^{(b)}$  remains constant during a given block  $b$  of certain duration and fades independently across blocks according to a uniform distribution over  $[0, 2\pi]$ . Thus,  $\theta_{kl}^{(b)}$  is i.i.d across  $k, l$  and  $b$ .

## III. MAIN RESULT

The following theorem is the main result of this paper.

*Theorem 1:* On the random phase  $K$ -user interference channel described in Section II, interference alignment can achieve a per user rate

$$R_k \geq \frac{2}{5} \left( \frac{1}{2} - \mu \right)^2 \log(1 + 2\mu \text{SNR}_k).$$

for any  $0 < \mu < 1/2$ , where  $\text{SNR}_k$  is the received SNR of user  $k$  defined as  $\text{SNR}_k = r_{kk}^2 P$ .

For any  $0 < \mu < 1/2$ , Theorem 1 provides a lower bound on the per user rate that is independent of the number of users in the system. This allows to conclude that on the random phase fast fading channel, interference alignment can achieve linear scaling in the number of users.

## IV. OVERVIEW OF THE INTERFERENCE ALIGNMENT SCHEME

In this section, we give a brief overview of the interference alignment scheme. We consider a slightly modified version of the original scheme in [5], so that each user uses the same signalling vectors that are simpler functions of the channel coefficients. As the blocklength increases to infinity, this modified scheme achieves  $K/2$  degrees-of-freedom as the original scheme in [5] but the blocklength of the scheme scales slightly better with  $K$ .

Communication is performed by jointly coding over  $M$  independent realizations (blocks) of the interference channel. The  $M$  blocks can be regarded as  $M$  parallel interference channels. Following [5], communication over these  $M$  parallel channels is denoted as

$$\mathbf{y}_k = \sum_{l=1}^K H_{kl} \tilde{\mathbf{x}}_l + \mathbf{z}_k \quad (2)$$

where  $H_{kl}$ 's are  $M \times M$  diagonal matrices with diagonal entries corresponding to different blocks,  $H_{kl}(b, b) = h_{kl}^{(b)}$ ,  $1 \leq b \leq M$  given in (1),  $\tilde{\mathbf{x}}_l$  is an  $M \times 1$  vector containing the  $M$  symbols transmitted by transmitter  $l$  over the  $M$  parallel channels and  $\mathbf{y}_k$  is an  $M \times 1$  vector of the observed signals at receiver  $k$ .

The message of transmitter  $l$  is encoded into  $d_s$  independent streams  $x_l(1), \dots, x_l(d_s)$  and transmitted along vectors  $\tilde{\mathbf{s}}_1, \dots, \tilde{\mathbf{s}}_{d_s}$  so that  $\tilde{\mathbf{x}}_l$  in (2) is given by

$$\tilde{\mathbf{x}}_l = \sum_{i=1}^{d_s} \tilde{\mathbf{s}}_i x_l(i) = \tilde{S} \mathbf{x}_l$$

where  $\tilde{S}$  is an  $M \times d_s$  matrix with columns  $\tilde{\mathbf{s}}_1, \dots, \tilde{\mathbf{s}}_{d_s}$  and  $\mathbf{x}_l$  is an  $d_s \times 1$  vector. The signaling vectors  $\tilde{\mathbf{s}}_1, \dots, \tilde{\mathbf{s}}_{d_s}$  are same for each transmitter and are chosen as the following particular function of the channel coefficient matrices,

$$\tilde{\mathbf{s}}_i = \left( \prod_{k \neq l} (H_{kl})^{\alpha_{kl}[i]} \right) \mathbf{w} \quad (3)$$

where  $\mathbf{w}$  is the  $M \times 1$  all-1 vector and  $\alpha_{kl}[i]$ , for  $k, l \in \{1, \dots, K\}$ ,  $k \neq l$  take values in  $\{0, 1, 2, \dots, n\}$  subject to the constraint

$$\sum_{\substack{k, l=1 \\ k \neq l}}^K \alpha_{kl}[i] = n,$$

where  $n$  is a positive integer. The  $K(K-1)$ -length sequence  $\{\alpha_{12}[i], \alpha_{13}[i], \dots, \alpha_{K(K-1)}[i]\}$  of  $\alpha$ 's is different for different  $i$ 's, so that no two signaling vectors are identical. The

number of distinct signalling vectors that can be constructed in this way is given by

$$d_s = C\left(n + K(K - 1) + 1, K(K - 1) + 1\right).$$

The signal received by receiver  $k$  under this strategy is given by

$$\mathbf{y}_k = H_{kk}\tilde{S}\mathbf{x}_k + \sum_{l \neq k} H_{kl}\tilde{S}\mathbf{x}_l + \mathbf{z}_k. \quad (4)$$

The interference experienced by each receiver is aligned to the columns of the  $M \times d_I$  matrix  $B$  where the  $i$ 'th column  $\mathbf{b}_i$  of  $B$  is given by

$$\mathbf{b}_i = \left( \prod_{k \neq l} (H^{[kl]})^{\beta_{kl}[i]} \right) \mathbf{w}$$

where  $\beta_{kl}[i]$ 's take values in  $\{0, 1, 2, \dots, n+1\}$  subject to the constraint  $\sum_{k,l=1, k \neq l}^K \beta_{kl}[i] = n+1$ . Note that

$$d_I = C\left(n + K(K - 1) + 2, K(K - 1) + 1\right).$$

We choose  $M = d_s + d_I$ .

In [5], it has been shown that if the channel coefficients  $h_{kl}^{(b)}$  are drawn from a continuous distribution and are i.i.d across  $k, l$  and  $b$ , then with high probability the signalling vectors constructed in (3) are linearly independent. Moreover, at each receiver  $k$ , the desired signal space spanned by the columns of the matrix  $H_{kk}\tilde{S}$ , and the interference space spanned by the columns of the matrix  $B$  are disjoint. (We say two subspaces are disjoint if there is no vector other than the zero vector that lies in both of these subspaces.) Thus, performing zero-forcing at the receivers that nulls out the interference space, the degrees of freedom achieved by this scheme are

$$\lim_{n \rightarrow \infty} \frac{K d_s}{M} = \frac{K}{2}. \quad (5)$$

Note that although the signal subspace and the interference subspace remain disjoint for any  $n$  (and  $M$ ) and this is what matters in the high SNR limit, the angle between the two subspaces may be decreasing with increasing  $n$  (and  $M$ ) which may be a limitation at finite SNR.

## V. PROOF OF MAIN RESULT

In this section, we lowerbound the performance of the scheme presented in the previous section for any SNR and  $K$  assuming the channel model in (1). Let us start by normalizing the signaling vectors defined in (3) to unit length. Let

$$d_i = \prod_{k \neq l} r_{kl}^{\alpha_{kl}[i]} \quad \mathbf{s}_i = \frac{1}{d_i \sqrt{M}} \tilde{\mathbf{s}}_i,$$

and  $S$  be the corresponding signalling matrix with columns  $\mathbf{s}_1, \dots, \mathbf{s}_{d_s}$ . Recall that the randomness in the channel in (1) comes from the phases and  $r_{kl}$ 's are given finite numbers. For the entries of the normalized signalling matrix  $S$ , we have

$$|S(p, m)|^2 = \frac{1}{M} \quad \text{for} \quad 1 \leq p \leq M, 1 \leq m \leq d_s, \quad (6)$$

and

$$\mathbb{E}[S(p, m)S(t, m)^*] = 0 \quad \text{for} \quad p \neq t \quad (7)$$

$$\mathbb{E}[S(p, m)S(p, t)^*] = 0 \quad \text{for} \quad m \neq t \quad (8)$$

where the first line follows from the fact that if  $p \neq t$ ,  $S(p, m)$  and  $S(t, m)^*$  relate to different rows of  $S$ , i.e, independent sets of  $\theta$ 's. This yields  $\mathbb{E}[S(p, m)S(t, m)^*] = \mathbb{E}[S(p, m)]\mathbb{E}[S(t, m)^*] = 0$  since if  $\theta$  is uniformly distributed over  $[0, 2\pi]$ , the distribution of  $e^{j\theta}$  is identical to  $e^{j\alpha\theta}$  for any integer  $\alpha \neq 0$ . The second line follows from the fact that when  $m \neq t$ ,  $S(p, m) \neq S(p, t)$  and there will remain at least one random variable  $H_{kl}(p, p)$  in (3) that is not multiplied with its complex-conjugate.

Using the signalling matrix  $S$  at the transmitters, the received vector at user  $k$  is given by

$$\mathbf{y}_k = H_{kk}S\mathbf{x}_k + \sum_{l \neq k} H_{kl}S\mathbf{x}_l + \mathbf{z}_k. \quad (9)$$

In general, we have the freedom to choose an arbitrary covariance matrix for the signals  $\mathbf{x}_k$  that can depend on  $S, H_{kk}$  and  $B$  as long as the power constraint  $\mathbb{E}_{S, \mathbf{x}_k}[\|S\mathbf{x}_k\|^2] \leq MP$  is satisfied. Since we claim achievability, we arbitrarily assume that we transmit independent streams encoded using a Gaussian codebook of power  $\frac{MP}{d_s}$ , that is

$$\mathbb{E}_{\mathbf{x}_k}[\mathbf{x}_k \mathbf{x}_k^\dagger] = \frac{MP}{d_s} I_{d_s}.$$

Note that this specific choice satisfies the power constraint  $\mathbb{E}[\|S\mathbf{x}_k\|^2] \leq MP$ .

Let  $Q$  be an  $M \times d_s$  matrix whose columns form an orthonormal basis for the orthogonal complement of the interference space, that is

$$\begin{aligned} Q^\dagger Q &= I_{d_s} \\ Q^\dagger B &= 0. \end{aligned} \quad (10)$$

Recall that the dimension of the interference space is  $d_I$  so its orthogonal complement has dimension  $M - d_I = d_s$ . Given (9), receiver  $k$  nulls out the interference by multiplying its observation by  $Q^\dagger$ ,

$$\hat{\mathbf{y}}_k = Q^\dagger \mathbf{y}_k = Q^\dagger H_{kk}S\mathbf{x}_k + Q^\dagger \mathbf{z}_k.$$

We can view  $\hat{\mathbf{y}}_k$  as the output of a MIMO channel with input  $\mathbf{x}_k$  and channel transfer matrix  $Q^\dagger H_{kk}S$ . Note that the additive Gaussian noise has covariance  $Q^\dagger I_M Q = I_{d_s}$ . Normalizing by the number of blocks  $M$ , user  $k$  can achieve the rate (see [7]),

$$R_k = \frac{1}{M} \log \det(I_{d_s} + (MP/d_s)Q^\dagger H_{kk}S S^\dagger H_{kk}^\dagger Q). \quad (11)$$

Note that the rate  $R_k$  is a random variable since  $S, H_{kk}$  and  $Q$  in (11) are functions of the random channel realizations. In the sequel, we will prove that when  $M$  is large

$$R_k \geq \frac{\frac{d_s}{M} (\frac{d_s}{M} - \mu)^2}{1 + \frac{d_s^2}{M^2}} \log \left( 1 + r_{kk}^2 P \frac{M\mu}{d_s} \right). \quad (12)$$

for any  $0 < \mu < \frac{d_s}{M}$  with high probability. Recall from (5) that  $\lim_{n \rightarrow \infty} d_s/M = 1/2$ . Thus, taking  $n \rightarrow \infty$  yields the result in Theorem 1.

The lower bound (12) will be proven in two steps: We will first prove a lower bound on the expected value of  $R_k$  over the random channel realizations,

$$\mathbb{E}_\theta[R_k] \geq \frac{\frac{d_s}{M}(\frac{d_s}{M} - \mu)^2}{1 + \frac{d_s^2}{M^2}} \log \left( 1 + r_{kk}^2 P \frac{M\mu}{d_s} \right) \quad (13)$$

and then show that  $R_k$  concentrates around its expected value when  $M$  is large,

$$\mathbb{P}(|R_k - \mathbb{E}_\theta[R_k]| \geq t) \leq 2e^{-t^2 M/2(\eta \log M)^2}. \quad (14)$$

for any  $t > 0$  and a constant  $\eta > 0$  independent of  $M$  and  $t$ .

Let us start by proving (13). We have,

$$\begin{aligned} \mathbb{E}_\theta[R_k] &= \frac{1}{M} \mathbb{E}_{H_{kk}, S, Q} [\log \det(I_{d_s} + (MP/d_s)Q^\dagger H_{kk} S S^\dagger H_{kk}^\dagger Q)] \\ &= \frac{d_s}{M} \mathbb{E}_{H_{kk}, S, Q} [\log(1 + \frac{MP}{d_s} \lambda)] \\ &\geq \frac{d_s}{M} \log(1 + \frac{MP}{d_s} \nu) \mathbb{P}(\lambda > \nu) \end{aligned}$$

for any  $\nu \geq 0$ , where  $\lambda$  is a random eigenvalue of the positive semi-definite matrix  $Q^\dagger H_{kk} S S^\dagger H_{kk}^\dagger Q$ . We can further lower bound  $\mathbb{E}_\theta[R_k]$  by using the Paley-Zygmund inequality (see [2, App. I]),

$$\mathbb{P}(\lambda > \nu) \geq \frac{(\mathbb{E}[\lambda] - \nu)^2}{\mathbb{E}[\lambda^2]},$$

Below, we compute  $\mathbb{E}[\lambda]$  and  $\mathbb{E}[\lambda^2]$  and show that

$$\mathbb{E}[\lambda] = \frac{d_s}{M} r_{kk}^2 \quad (15)$$

$$\mathbb{E}[\lambda^2] \leq r_{kk}^4 + \frac{d_s^2}{M^2} r_{kk}^4. \quad (16)$$

Choosing  $0 < \mu < \frac{d_s}{M}$  and  $\nu = r_{kk}^2 \mu$  yields the lower bound (13). Equations (15) and (16) are proven below.

We have

$$\begin{aligned} \mathbb{E}[\lambda] &= \frac{1}{d_s} \mathbb{E}_{H_{kk}, S, Q} [\text{tr}(Q^\dagger H_{kk} S S^\dagger H_{kk}^\dagger Q)] \\ &= \frac{1}{d_s} \mathbb{E}_{H_{kk}, S, Q} \left[ \sum_{m,r=1}^M H_{kk}(m, m) (S S^\dagger)(m, r) \right. \\ &\quad \left. H_{kk}^\dagger(r, r) (Q Q^\dagger)(r, m) \right] \\ &= \frac{1}{d_s} r_{kk}^2 \mathbb{E}_{S, Q} \left[ \sum_{r=1}^M (S S^\dagger)(r, r) (Q Q^\dagger)(r, r) \right] \quad (17) \end{aligned}$$

$$= \frac{1}{d_s} r_{kk}^2 \mathbb{E}_Q \left[ \sum_{r=1}^M \frac{d_s}{M} (Q Q^\dagger)(r, r) \right] \quad (18)$$

$$= \frac{d_s}{M} r_{kk}^2, \quad (19)$$

where (17) follows by taking the expectation with respect to  $H_{kk}$  that yields  $\mathbb{E}[H_{kk}(m, m)H_{kk}(r, r)] = \delta_{mr}$ . Note that  $H_{kk}$  is independent of  $S$  and  $Q$ . Step (18) follows from the fact that  $(S S^\dagger)(r, r) = d_s/M$  which follows from (6). Step (19) follows from the fact that  $\text{tr}(Q Q^\dagger) = \text{tr}(Q^\dagger Q) = d_s$ .

We can upper bound the second moment of  $\lambda$  as follows,

$$\begin{aligned} \mathbb{E}[\lambda^2] &= \frac{1}{d_s} \mathbb{E}_{H_{kk}, S, Q} [\text{tr}(Q^\dagger H_{kk} S S^\dagger H_{kk}^\dagger Q Q^\dagger H_{kk} S S^\dagger H_{kk}^\dagger Q)] \\ &= \frac{1}{d_s} \mathbb{E}_{H_{kk}, S, Q} \left[ \sum_{m,r,p,t=1}^M H_{kk}(m, m) (S S^\dagger)(m, r) H_{kk}^\dagger(r, r) \right. \\ &\quad \left. (Q Q^\dagger)(r, p) H_{kk}(p, p) (S S^\dagger)(p, t) H_{kk}^\dagger(t, t) (Q Q^\dagger)(t, m) \right] \\ &= \frac{1}{d_s} r_{kk}^4 \mathbb{E}_{S, Q} \left[ \sum_{m,p=1}^M (S S^\dagger)(m, m) (S S^\dagger)(p, p) |(Q Q^\dagger)(p, m)|^2 \right] \\ &\quad + \frac{1}{d_s} r_{kk}^4 \mathbb{E}_{S, Q} \left[ \sum_{\substack{m,r=1 \\ m \neq r}}^M |(S S^\dagger)(m, r)|^2 (Q Q^\dagger)(r, r) (Q Q^\dagger)(m, m) \right] \quad (20) \end{aligned}$$

where the last equality follows by identifying the non-zero terms after taking the expectation with respect to  $H_{kk}$ . The first term in (20) yields,

$$\begin{aligned} &\frac{1}{d_s} r_{kk}^4 \mathbb{E}_{S, Q} \left[ \sum_{m,p=1}^M (S S^\dagger)(m, m) (S S^\dagger)(p, p) |(Q Q^\dagger)(p, m)|^2 \right] \\ &= \frac{1}{d_s} r_{kk}^4 \mathbb{E}_S \left[ \sum_{m,p=1}^M \frac{d_s^2}{M^2} |(Q Q^\dagger)(p, m)|^2 \right] \\ &= \frac{d_s^2}{M^2} r_{kk}^4 \end{aligned}$$

since  $\sum_{m,p=1}^M |(Q Q^\dagger)(p, m)|^2 = \text{tr}(Q Q^\dagger Q Q^\dagger) = \text{tr}(Q Q^\dagger) = d_s$ . The second term in (20) can be upper bounded as

$$\begin{aligned} &\frac{1}{d_s} r_{kk}^4 \mathbb{E}_{S, Q} \left[ \sum_{\substack{m,r=1 \\ m \neq r}}^M |(S S^\dagger)(m, r)|^2 (Q Q^\dagger)(r, r) (Q Q^\dagger)(m, m) \right] \\ &\leq \frac{1}{d_s} r_{kk}^4 \mathbb{E}_S \left[ \sum_{\substack{m,r=1 \\ m \neq r}}^M |(S S^\dagger)(m, r)|^2 \right] \quad (21) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{d_s} r_{kk}^4 \sum_{\substack{m,r=1 \\ m \neq r}}^M \sum_{p,t=1}^{d_s} \mathbb{E}_S [S(m, p) S^*(r, p) S^*(m, t) S(r, t)] \\ &= \frac{1}{d_s} r_{kk}^4 \sum_{\substack{m,r=1 \\ m \neq r}}^M \sum_{p=1}^{d_s} |S(m, p)|^2 |S(r, p)|^2 \quad (22) \end{aligned}$$

$$\leq r_{kk}^4 \quad (23)$$

where inequality (21) follows from the fact that  $0 \leq (Q Q^\dagger)(r, r) \leq 1$  and (22) follows from (7) and (8).

In order to prove (14), we make use of the following concentration inequality due to McDiarmid [8].

*Theorem 2:* Let  $(L_1, L_2, \dots, L_n)$  be a family of independent random variables with  $L_k$  taking values in a set  $A_k$  for each  $k$ . Suppose that the real-valued function  $f$  defined on  $\Pi A_k$  satisfies

$$\sup_{l_1, \dots, l_n, l'_k} |f(l_1, \dots, l_k, \dots, l_n) - f(l_1, \dots, l'_k, \dots, l_n)| \leq c_k.$$

Then, for any  $t \geq 0$ ,

$$\mathbb{P}(|f(L_1, \dots, L_n) - E[f(L_1, \dots, L_n)]| \geq t) \leq 2e^{-\frac{2t^2}{\sum c_k^2}}.$$

Let  $\boldsymbol{\theta}_1$  denote the vector of  $K^2$  channel state realizations for the first block, i.e.,  $\boldsymbol{\theta}_1 = [\theta_{kl}^{(1)}, k, l = 1, \dots, K]$ . Note that  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M$  are independent random variables and they determine the corresponding rows of  $S, B$  and  $H_{kk}$ 's. To apply Theorem 2, we will first show that

$$|R_k(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) - R_k(\boldsymbol{\theta}'_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)| \leq \frac{2\eta}{M} \log M$$

for a constant  $\eta > 0$ . To show this, first note that the choice of the  $Q$  matrix in (10) is not unique but any matrix that satisfies the conditions in (10) will yield the same rate in (11). Let us choose  $Q$  in the following particular way: Let  $\mathbf{b}_1$  denote the first row of the matrix  $B$  and  $\hat{B}$  denote the  $(M-1) \times d_I$  matrix containing the remaining  $M-1$  rows. Let  $\hat{Q}$  be an  $(M-1) \times (d_s-1)$  matrix such that

$$\begin{aligned} \hat{Q}^\dagger \hat{Q} &= I_{d_s-1} \\ \hat{Q}^\dagger \hat{B} &= 0. \end{aligned}$$

There exists a matrix  $Q$  satisfying (10) and of the form

$$Q = \left[ \begin{array}{c|ccc} \mathbf{v} & 0 & \cdots & 0 \\ \hline & & \hat{Q} & \end{array} \right] := [ \mathbf{v} \mid \tilde{Q} ],$$

where  $Q$  depends on  $\mathbf{b}_1$  only through  $\mathbf{v}$ . Observe that the last  $d_s-1$  columns of  $Q$  are orthonormal to each other and at the same time orthogonal to the columns of  $B$ . They are independent of  $\mathbf{b}_1$  by construction. The vector  $\mathbf{v}$  is chosen such that it completes these  $d_s-1$  column vectors to an orthonormal basis for the orthogonal complement of the range space of  $B$ .

For this choice of  $Q$ , Fischer-Hadamard's inequality yields

$$\begin{aligned} R_k &= \frac{1}{M} \log \det(I_{d_s} + (MP/d_s)Q^\dagger H_{kk}SS^\dagger H_{kk}^\dagger Q) \\ &\leq \frac{1}{M} \log \det(1 + (MP/d_s)\mathbf{v}^\dagger H_{kk}SS^\dagger H_{kk}^\dagger \mathbf{v}) \\ &\quad + \frac{1}{M} \log \det(I_{d_s-1} + (MP/d_s)\tilde{Q}^\dagger H_{kk}SS^\dagger H_{kk}^\dagger \tilde{Q}). \end{aligned} \quad (24)$$

The first term in this expression can be upper bounded by  $\eta \log M/M$  for a constant  $\eta$  while the second term is independent of the channel states  $\boldsymbol{\theta}_1$  at first block, due to the all-zero first row of  $\tilde{Q}$ . Let us denote the second term in (24) by  $\psi(\boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)$ . It can be shown that

$$R_k(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) - \psi(\boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) \geq 0,$$

by considering the interlacing property for eigenvalues of Hermitian matrices and the fact all eigenvalues of the matrix  $I_{d_s} + (MP/d_s)Q^\dagger H_{kk}SS^\dagger H_{kk}^\dagger Q$  are larger than or equal to 1. This fact together with the triangle inequality yields

$$\begin{aligned} &|R_k(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) - R_k(\boldsymbol{\theta}'_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)| \\ &\leq |R_k(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) - \psi(\boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)| \\ &\quad + |R_k(\boldsymbol{\theta}'_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M) - \psi(\boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)| \leq \frac{2\eta}{M} \log M. \end{aligned}$$

Noting that this bound applies uniformly for all  $\boldsymbol{\theta}$ 's and using Theorem 2 yields the concentration inequality in (14). This completes the proof of the main result of this paper.

## VI. DISCUSSION

The degrees of freedom result in [5] is established for any fast fading channel model with channel coefficients drawn from an i.i.d continuous distribution. In this paper, we establish a linear scaling result for the channel model where the fading is only through the phases and the magnitudes of the channel coefficients remain constant during communication. The extension of our result to channel distributions with fading magnitudes is non-trivial. Two problems arise with the construction of signalling vectors in Section IV when the magnitudes are fading. First, the signalling vectors constructed in (3) become correlated and secondly as the blocklength increases most of them start to resemble standard basis vectors which are eigenvectors for the diagonal channel matrices in (4). This is not the case for the random phase model we consider in this paper. The first problem can be simply resolved by using an orthogonal basis for signalling, instead of using directly the columns of  $S$  as we did in here. The question that remains is whether these orthogonal basis vectors also converge to standard basis vectors or not. If they do, since standard basis vectors are eigenvectors for diagonal matrices, the direct channel matrices may not be able to rotate the signalling vectors and provide enough separation between the desired signal and interference subspaces at the receivers. This may, in turn, imply low SNR for the signal component orthogonal to interference.

## REFERENCES

- [1] P. Gupta and P. R. Kumar, *The Capacity of Wireless Networks*, IEEE Trans. on Information Theory 42 (2), pp.388-404, 2000.
- [2] A. Özgür, O. Lévêque, D. Tse, *Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad-Hoc Networks*, IEEE Trans. on Information Theory 53 (10), pp.3549-3572, 2007.
- [3] M. Maddah-Ali, A. Motahari, A. Khandani, *Signaling over MIMO Multi-base Systems -Combination of Multiple-Access and Broadcast Schemes*, Proc. of ISIT 2006, 2104-2108.
- [4] S. A. Jafar, S. Shamai, *Degrees of Freedom Region for the MIMO X Channel*, IEEE Trans. on Information Theory 54 (8), pp.3425-3441, 2008.
- [5] V. Cadambe, S. Jafar, *Interference Alignment and Degrees of Freedom for the K-user Interference Channel*, IEEE Trans. on Information Theory 54 (8), 2008, 3425-34412.
- [6] B. Nazer, M. Gastpar, S. A. Jafar and S. Vishwanath, *Ergodic Interference Alignment*, arXiv:0901.4379, e-print.
- [7] E. Telatar, *Capacity of Multi-Antenna Gaussian Channels*, European Trans. on Telecommunications 10 (6), 585-596, November 1999.
- [8] M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed, *Probabilistic Methods for Algorithmic Discrete Mathematics*, Springer, 1998.