System performance of interference alignment under TDD mode with limited backhaul capacity

Matha Deghel, Mohamad Assaad, Mérouane Debbah

System Performance of Interference Alignment under TDD Mode with Limited Backhaul Capacity

Matha Deghel,* Mohamad Assaad* and Mérouane Debbah*†
* Department of Telecommunications SUPELEC, Gif-sur-Yvette, France
† Mathematical and Algorithmic Sciences Lab, HUAWEI, France
E-mail: {matha.deghel, mohamad.assaad, merouane.debbah}@supelec.fr

Abstract—This paper considers a MIMO interference system where interference alignment (IA) technique is adopted to manage the problem of interference. We consider a time division duplex (TDD) system where each transmitter estimates its channel state information (CSI) by probing the receivers. In addition, the transmitters share their local CSI estimate between each other using a backhaul links of limited capacity. A quantization over the backhaul is therefore required to reduce the amount of information to exchange. We study in this paper the impact of this quantization on the system performance and determine the optimal number of transmitter-receiver pairs that maximizes the system throughput.

I. INTRODUCTION

Interference is one of the major drawbacks in wireless communication systems due to the large number of users communicating on the same channel. This problem has motivated the researchers to investigate transmitting schemes that can mitigate interference. Interference alignment (IA) was introduced in [1] as one of the most efficient interference management technique. It is based on the concept of designing precoding scheme that confine the interfering signals observed at each receiver into a low dimensional subspace, providing a larger subspace to decode the desired signal. In [2], IA has been shown to achieve maximum multiplexing gain in MIMO channels. One disadvantage of IA is that it requires global channel state information (CSI) at each of the transmitters, which is difficult to obtain in practical systems. Therefore, IA under limited feedback was studied and several quantization techniques were proposed, in order to aid the transmitters to acquire (probe) CSI knowledge from receivers and then to share it between each other. For instance, in [3] a compression scheme for the cloud radio access networks is proposed. In [4], the Grassmannian Manifold quantization technique was adopted to reduce the information exchange. Another scheme proposed in [5] is used in sending the channel conditions from users to transmitters.

An important factor to consider, which is related to the CSI acquisition process, is the CSI probing (acquisition) cost. We consider a TDD mode where receivers (users) send training sequences in the uplink so that the transmitter can estimate their channels. Since this scheme uses orthogonal sequences, their lengths are proportional to the number of active users in the system. In other words, after acquiring the CSI of L users, the rate is multiplied by $1 - L\theta$, where θ is the fraction (of

the slot) that takes the CSI acquisition of one user [6]. From the above, it can be seen that choosing the number of pairs L is important to investigate. If L is high, it means that we have more users to probe and then less time for data transmission. This problem was studied in [7] by optimizing the transmission rate which is function of the overhead caused by the probing process for an analog feedback strategy.

The context here is different from the aforementioned work. We assume prefect local CSI at the transmitters, but each of which needs to send its local knowledge to all other transmitters over finite capacity backhaul links. Then, the interference alignment solution is computed independently at each transmitting node.

The remainder of this paper is organized as follows. In Section II, the system model is described. Section III analyzes the performance of the system when the IA technique is used and when a quantization scheme is performed over the backhaul links of finite capacity. Therein, we derive the transmission rate of the system, then we define and solve an optimization problem which seeks to maximize this rate with respect to the number of pairs and bits. Finally, we give numerical results in Section IV and conclude in Section V.

Notation: Boldface uppercase symbols (i.e., $\mathbf{A}$) represent matrices whereas lowercases (i.e., $\mathbf{a}$) are used for vectors.; $(.)^*$ denotes the conjugate transpose of a vector or matrix; '$\mathbf{I}$' is used for square identity matrix; $|\,.\,|$ denotes the absolute value; $\|\,.\,\|$ represents the norm of second degree; $\mathcal{CN}(\mathbf{a}, \mathbf{A})$ is a complex Gaussian random vector with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$.

II. SYSTEM MODEL

In this section, we present the MIMO interference network under consideration, where we apply the IA technique over finite capacity backhaul links. We review the concept of IA and we propose a CSI sharing scheme to reduce the amount of information exchange required to achieve IA.

A. MIMO Interference Channel Model

Consider the L-user MIMO interference network illustrated in Fig. 1 in which each transmitter (T_X) is equipped with N_t antennas and each receiver (R_X) has N_r antennas. Transmitter k has $d_k (\leq \min(N_t, N_r))$ data streams to send to its intended receiver (user) k.
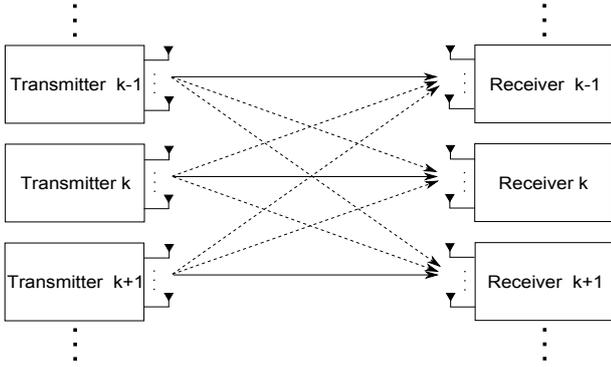
Figure 1: L-User MIMO interference network.

Under this system model, the received signal at user k can be expressed as follows:

$$\mathbf{y}_k = \sum_{i=1}^{L} \sqrt{\frac{\gamma_{ki}P}{d_i}} \mathbf{H}_{ki} \sum_{j=1}^{d_i} \mathbf{v}_i^j x_i^j + \mathbf{z}_k \,, \qquad (1)$$

in which $\mathbf{H}_{ki} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix between T_X i and R_X k with independent and identically distributed zero mean and unit variance Gaussian elements, γ_{ki} represents the path loss of channel $\mathbf{H}_{ki}$, P is the total power at each transmitter equally allocated among its streams, x_i^j denotes the jth data stream from transmitter i, $\mathbf{v}_i^j \in \mathbb{C}^{N_t \times 1}$ is the corresponding precoding vector of unit norm and $\mathbf{z}_k$ is the additive white Gaussian noise distributed according to $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_r})$. We denote by α_{ki} the fraction $\frac{\gamma_{ki}P}{d_i}$, for all k, i in $\{1, ..., L\}$.

We assume a TDD transmission where time is slotted and where a transmitter can acquire the channel state information of a receiver by probing, which consumes a fraction θ of a slot. Probing more users takes more time, and hence leaves a smaller fraction of resources for actual data transmission. As alluded earlier, if L users (i.e. it means we have L active transmitter-receiver pairs) are probed, then the actual rate becomes $(1 - L\theta)R$, where R is the transmission rate without considering the probing cost.

B. Application of Interference Alignment Technique

For the sake of simplicity, we restrict ourselves to a per-stream zero-forcing receiver. Specifically, Receiver k uses the *combiner* vector $\mathbf{u}_k^m \in \mathbb{C}^{N_r \times 1}$ of unit norm to detect the m-th stream from transmitter k, such as

$\hat{x}_k^m = (\mathbf{u}_k^m)^* \mathbf{y}_k$

$$= \overbrace{\sqrt{\alpha_{kk}} (\mathbf{u}_k^m)^* \mathbf{H}_{kk} \mathbf{v}_k^m x_k^m}^{\text{desired signal}} + \overbrace{\sqrt{\alpha_{kk}} \sum_{\substack{j=1 \\ j \neq m}}^{d_k} (\mathbf{u}_k^m)^* \mathbf{H}_{kk} \mathbf{v}_k^j x_k^j}^{\text{inter-stream interference (ISI)}}$$

$$+ \overbrace{\sum_{\substack{i=1 \\ i \neq k}}^{L} \sqrt{\alpha_{ki}} \sum_{j=1}^{d_i} (\mathbf{u}_k^m)^* \mathbf{H}_{ki} \mathbf{v}_i^j x_i^j}^{\text{inter-user interference (IUI)}} + \overbrace{(\mathbf{u}_k^m)^* \mathbf{z}_k}^{\text{noise}} \,. \qquad (2)$$

As can be seen in (2), two sources of interferences (ISI and IUI) affect the detection at each receiver. To manage this problem, we use the IA technique which consists on designing the set of precoder and combiner vectors such that [5]:

$$(\mathbf{u}_k^m)^* \mathbf{H}_{ki} \mathbf{v}_i^j = 0, \qquad \forall (k, m) \neq (i, j). \qquad (3)$$

We admit that each receiver obtains a perfect knowledge of the combiner vector designed at its paired transmitter.

In the ideal case where we have perfect global CSI at all the transmitters and for an achievable degree of freedom vector $\mathbf{d} = [d_1, ..., d_L]$, ILI and ISI can be canceled completely at the receivers since the conditions for *perfect alignment* are satisfied. The CSI sharing mechanism over the limited backhaul is detailed in the following.

C. CSIT Sharing Over Finite Capacity Backhaul Links

As mentioned before, global CSI is required in order to solve the interference alignment problem given in (3). We assume that the transmitters have a perfect knowledge of their *local CSI*, meaning that the i-th transmitter estimates perfectly the channels $\mathbf{H}_{ki}$, for $k = 1, ..., L$. However, the local CSI (excluding the direct links) of other transmitters are obtained via backhaul links of limited capacity.
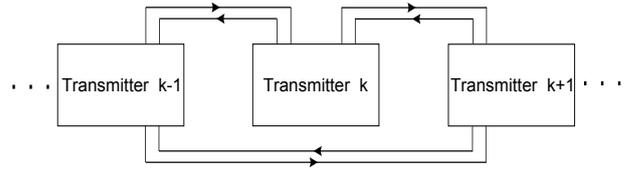

Figure 2: CSIT sharing over backhaul links.

We focus on the scenario shown in Fig. 2 where each T_X receives all the required CSI and independently designs the IA vectors [4]. But, since the backhaul links that connect transmitters to each other are of limited capacity, a codebook-based quantization is adopted to reduce the huge amount of information exchange.

In detail, let $\mathbf{h}_{ki}$ represents the vectorization of the channel matrix $\mathbf{H}_{ki}$ and suppose that all the transmitters share a predetermined codebook $\mathcal{CB}$ of size 2^B. Transmitter i selects the index n_o of the optimal codeword in $\mathcal{CB}$ according to: $n_o = \arg\max_{1 \leq n \leq 2^B} \left| \tilde{\mathbf{h}}_{ki}^* \hat{\mathbf{h}}_{ki}^n \right|^2$, where $\tilde{\mathbf{h}}_{ki} = \frac{\mathbf{h}_{ki}}{\|\mathbf{h}_{ki}\|}$ and B is the number of bits used to quantize $\mathbf{H}_{ki}$.

After quantizing all the matrices of its local CSI, transmitter i sends the corresponding optimal indexes to all other transmitters which share the same codebook, allowing these transmitters to reconstruct the quantized local knowledge of transmitter i. Using the model in [5], we define the quantization error as $e_{ki} = 1 - \frac{\left| \hat{\mathbf{h}}_{ki}^* \mathbf{h}_{ki} \right|^2}{\|\mathbf{h}_{ki}\|^2}$. The cumulative distribution function (CDF) of e_{ki} is then given by: $\Pr(e_{ki} \leq \varepsilon) = 2^B \varepsilon^Q$ for $0 \leq \varepsilon \leq 2^{-\frac{B}{Q}}$, where $Q = N_t N_r - 1$.

III. SYSTEM PERFORMANCE ANALYSIS

In this section, we analyze the performance for IA in a MIMO interference network with backhaul links of limited

capacity. We first derive the total transmission rate under the considered system. Then, we provide an optimization problem with its corresponding solution for maximizing this rate.

A. Transmission Rate Under Finite Capacity Backhaul Links

As shown in the previous section, the IA vectors are designed based on the available CSI knowledge which is obtained after the transmitting nodes share their perfect local knowledge between each other. It results that the IA technique is capable to completely cancel the ISI since local CSI is perfectly known, but not the ILI due to the quantization process which leads to imperfect global CSI at the transmitting nodes. Under such conditions, the signal to interference plus noise ratio (SINR) for stream m at receiver k is be given by:

$$\zeta_k^m = \frac{\alpha_{kk} \left| (\hat{\mathbf{u}}_k^m)^* \mathbf{H}_{kk} \hat{\mathbf{v}}_k^m \right|^2}{\sigma^2 + \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \sum_{j=1}^{d_i} \left| (\hat{\mathbf{u}}_k^m)^* \mathbf{H}_{ki} \hat{\mathbf{v}}_i^j \right|^2} \quad (4)$$

where $\hat{\mathbf{v}}_k^m$ and $\hat{\mathbf{u}}_k^m$ are the precoding and combining vectors, respectively, designed based on the available CSI described in the previous section. We denote by I_k^m the interference term that appears in the denominator of (4).

$$I_k^m = \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \sum_{j=1}^{d_i} \left| (\hat{\mathbf{u}}_k^m)^* \mathbf{H}_{ki} \hat{\mathbf{v}}_i^j \right|^2 = \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \sum_{j=1}^{d_i} \left| \mathbf{h}_{ki} \mathbf{s}_{k,i}^{m,j} \right|^2$$

$$= \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \left\| \mathbf{h}_{ki} \right\|^2 \sum_{j=1}^{d_i} \left| (\tilde{\mathbf{h}}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2, \quad (5)$$

where $\mathbf{w}_{ki}$ is a unit norm vector isotropically distributed in the null space of $\hat{\mathbf{h}}_{ki}$, $\mathbf{s}_{k,i}^{m,j} = \hat{\mathbf{v}}_i^j \otimes (\hat{\mathbf{u}}_k^m)^*$ ($\otimes$ is the Kronecker product) and $\tilde{\mathbf{h}}_{ki}$ is the normalized vector of channel $\mathbf{h}_{ki}$.

Following the model used in [5], the channel direction $\tilde{\mathbf{h}}_{ki}$ can be written as follows: $\tilde{\mathbf{h}}_{ki} = \sqrt{1 - e_{ki}} \hat{\mathbf{h}}_{ki} + \sqrt{e_{ki}} \mathbf{w}_{ki}$, where $\hat{\mathbf{h}}_{ki}$ is the channel quantization vector of $\mathbf{h}_{ki}$ and $\mathbf{w}_{ki}$ is independent of e_{ki}, with $\mathbf{w}_{ki}$ (of unit norm) isotropically distributed in the null space of $\hat{\mathbf{h}}_{ki}$. The product $\left| (\tilde{\mathbf{h}}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2$ can then be expressed as $\left| (\tilde{\mathbf{h}}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2 = \left| \sqrt{1 - e_{ki}} (\hat{\mathbf{h}}_{ki})^* \mathbf{s}_{k,i}^{m,j} + \sqrt{e_{ki}} (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2 = e_{ki} \left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2$. Therefore, I_k^m can be rewritten as:

$$I_k^m = \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \left\| \mathbf{h}_{ki} \right\|^2 e_{ki} \sum_{j=1}^{d_i} \left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2. \quad (6)$$

Transmission Rate : Based on [8], we define the transmission rate (throughput) achieved as the probability to get an SINR greater than a given threshold τ. In practice, this can be interpreted by the fact that if the SINR is lower than a certain value, then the transmitted signal can not be decoded correctly. Thus, we can write the throughput that corresponds to stream m of pair k as:

$$R_k^m = \Pr(\zeta_k^m \geq \tau) \quad (7)$$

Proposition 1. *The transmission rate corresponding to stream m at user k can be given by*

$$R_k^m = e^{-\frac{\sigma^2 \tau}{\alpha_{kk}}} MGF_{I_k^m}(-t), \quad (8)$$

where $t = \frac{\tau}{\alpha_{kk}}$ and $MGF_{I_k^m}$ is the moment generating function of the random variable I_k^m.

Proof. The proof is provided in Appendix A. $\square$

The above proposition provides a general formula for the transmission rate under the considered system, which depends on the MGF expression of I_k^m. This latter expression is given in the following proposition.

Proposition 2. *The expression $MGF_{I_k^m}(-t)$ is given by*

$$MGF_{I_k^m}(-t)$$
$$= \prod_{\substack{i=1 \\ i \neq k}}^{L} \left(\frac{\alpha_{ki} t d_i}{\delta} + 1 \right)^{-Q} {}_2F_1\left(b_i, Q; a_i + b_i; \frac{1}{1 + \frac{\delta}{\alpha_{ki} t d_i}}\right), \quad (9)$$

where ${}_2F_1$ is the hypergeometric function, $\delta = 2^{\frac{B}{Q}}$, $a_i = \frac{(Q+1)d_i}{Q} - \frac{1}{Q}$ and $b_i = (Q-1)a_i$.

Proof. Refer to Appendix B for the proof. $\square$

Let R denote the *total transmission rate* of the system which can be calculated by taking the sum of all transmission rates over all streams and pairs, such as:

$$R = \sum_{k=1}^{L} \sum_{m=1}^{d_k} R_k^m = \sum_{k=1}^{L} \sum_{m=1}^{d_k} e^{-\frac{\sigma^2 \tau}{\alpha_{kk}}} MGF_{I_k^m}(-t). \quad (10)$$

The above result does not consider the cost of probing that, as mentioned in Section II, will reduce the transmission rate by a factor $1 - L\theta$, where θ is the fraction of time to probe one user. Under all the above considerations, the total transmission rate can be given by the following proposition.

Proposition 3. *If we consider the probing cost, the total transmission rate R_p can be rewritten as:*

$$R_p = (1 - L\theta) \sum_{k=1}^{L} \sum_{m=1}^{d_k} \left[e^{-\frac{\sigma^2 \tau}{\alpha_{kk}}} \prod_{\substack{i=1 \\ i \neq k}}^{L} \left(\frac{\alpha_{ki} \tau d_i}{\alpha_{kk} 2^{\frac{B}{Q}}} + 1 \right)^{-Q} \right.$$
$$\left. \times {}_2F_1\left(b_i, Q; a_i + b_i; \frac{1}{1 + \frac{\alpha_{kk} 2^{\frac{B}{Q}}}{\alpha_{ki} \tau d_i}}\right) \right]. \quad (11)$$

Proof. Proposition 3 is proved by combining the probing cost, the expression in (10) and the result of Proposition 2 . $\square$

B. Throughput Maximization

The total transmission rate in (11) is a function of several parameters. Among these parameters, we focus on the number of bits B and the number of pairs L. We analyze the system performance by maximizing the expression of the transmission rate in (16), function of L and B, under the constraint of finite (total) capacity C of backhaul links. But, as it can be seen

in (11), solving this problem for the general case is of high complexity. Therefore, before proceeding in the analysis, we make the following assumptions: (i) all the transmitters have the same number of streams d and (ii) all the direct links and all the cross links have equal path loss γ_1 and γ_2, respectively. Under these assumptions, we can rewrite (11) as

$$
R_p = (1 - L\theta) \sum_{k=1}^{L} \sum_{m=1}^{d} \left[e^{-\frac{\sigma^2 \tau}{\alpha_1}} \prod_{\substack{i=1 \\ i \neq k}}^{L} (\frac{\alpha_2 \tau d}{\alpha_1 2^{\frac{B}{Q}}} + 1)^{-Q} \right.
$$

$$
\times \left. {}_2F_1(b, Q; a+b; \frac{1}{1 + \frac{\alpha_1 2^{\frac{B}{Q}}}{\alpha_2 \tau d}}) \right] = (1 - L\theta) L d e^{-\frac{\sigma^2 \tau}{\alpha_1}}
$$

$$
\times \left[(\frac{\alpha_2 \tau d}{\alpha_1 2^{\frac{B}{Q}}} + 1)^{-Q} {}_2F_1(b, Q; a+b; \frac{1}{1 + \frac{\alpha_1 2^{\frac{B}{Q}}}{\alpha_2 \tau d}}) \right]^{L-1}.
$$

$$(12)$$

where $\alpha_1 = \frac{\gamma_1 P}{d}$ and $\alpha_2 = \frac{\gamma_2 P}{d}$.

Now, we can define the optimization problem as follows:

$$\underset{B,L}{\text{maximize}} \qquad R_p(B,L) \qquad (13)$$

$$\text{subject to} \qquad L(L-1)^2 B \leq C, \qquad (14)$$

where $L(L-1)^2 B$ is the total number of bits exchanged on the backhaul links of limited capacity C. This expression is obtained from the fact that we have L transmitters, each of which shares $L-1$ channels to $L-1$ other transmitters.

Remark 1. *To ensure the feasibility of the interference alignment problem, one additional condition (given in [9]) to consider is that $N_t + N_r \geq d(L+1)$, which puts a limitation on the maximum number of pairs.*

We propose the following algorithm to solve the maximization problem defined by (13) and (14).

Optimization algorithm

Fix the capacity C of backhaul links.
for $B = 1$ to N_B **do**
  for $L = 1$ to N_L **do**
    if condition (14) is not statisfied **then**
      put $R_p(B,L) = 0$.
    else
      compute the transmission rate $R_p(B,L)$.
    end if
  end for
end for
Choose B and L which correspond to the max value of R_p.

Note that the value of N_L should be chosen based on the condition given in the remark before.

IV. NUMERICAL RESULTS

In this section we present the numerical results. We assume a homogeneous system where the number of antennas $N_t = N_r = 15$, $d = 2$, $\tau = 0.8$ and $\gamma_1 = 1$. We take $N_B = 30$
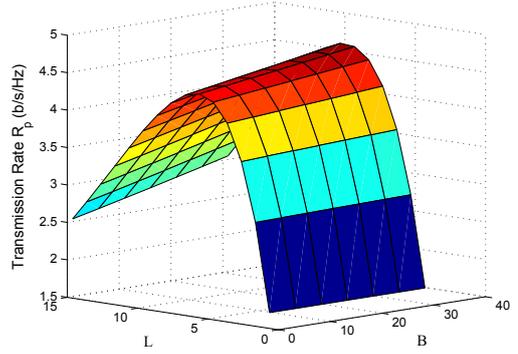

Figure 3: The transmission rate R_p for different combinations of number of bits and number of pairs. the factor $\theta = 0.01$.
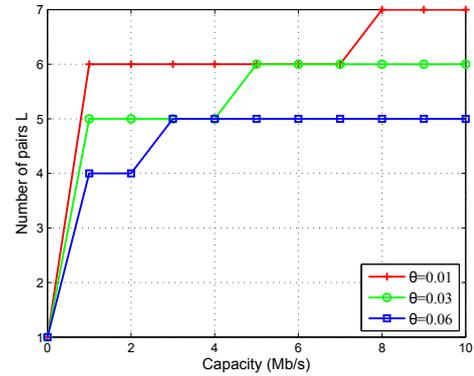

Figure 4: Optimal number of pairs L for different values of the capacity. The factor $\theta = \{0.01, 0.03, 0.06\}$.

which is the maximum number of bits that we can use for the quantization process, and $N_L = 14$ which satisfies the condition given in the previous section. In addition, we use $10 \log_{10}(\frac{P}{\sigma^2})$ to represent the SNR in dB.

Fig. 3 plots the transmission rate of the system in (12) for different combinations of the number of bits B and the number of pairs L, when $\theta = 0.01$, $\gamma_2 = 0.1$ and SNR = 10 dB. As can be seen from this figure, for a fixed B, the transmission rate R_p is very sensitive to the variation of the number of pairs L. However, for a fixed L, the function R_p is less sensitive to the variation of the number of bits B.

Now, we consider the same path loss γ_2 as before and we use the algorithm given in Section III to obtain the optimal values of B and L that maximize the rate R_p for different values of the total capacity C and the fraction θ.

Fig. 4 displays the variation of the optimal number of pairs L as a function of the capacity C, for different values of θ. The more we increase the capacity, the more we relax the constraint in (19), then the possibility that the optimal value of L increases is higher. Moreover, if the fraction θ is low (i.e. low probing cost), the optimal number of pairs can reach higher values which will raise the system rate. For instance, for $\theta = 0.01$, the number of pairs L reaches 7 pairs at its
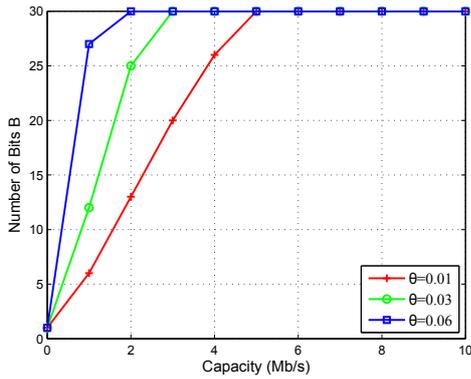
Figure 5: Optimal number of pairs B for different values of the capacity. The factor $\theta = \{0.01, 0.03, 0.06\}$.

Figure 7: The variation of the optimal number of transmitter-receiver pairs in function of the path loss coefficient γ_2.
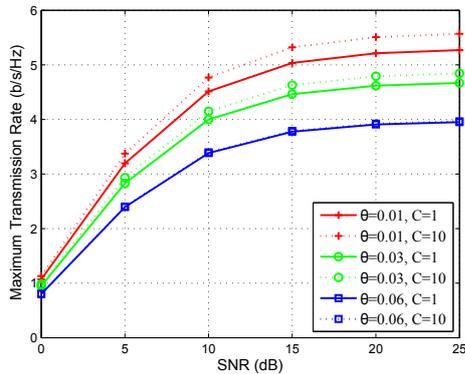

Figure 6: The variation of the maximum transmission rate in function of the SNR. We consider different values of the factor $\theta = \{0.01, 0.03, 0.06\}$ and $C = \{1, 10\}$ Mb/s.

Figure 8: The variation of the maximum transmission rate in function of the path loss coefficient γ_2.

maximum instead of 5 for $\theta = 0.06$.

Fig. 5 shows the variation of the optimal number of bits B as a function of the capacity C. For the same θ, B increases until it reaches the maximal value 30. If we fix the value of C, we can see that B takes larger values for higher θ. This can be explained by the fact that the transmission of the system is more sensitive to L than to B.

Fig. 6 plots the variation of the maximum transmission rate as function of the SNR for different values of θ and the capacity C. For large values of θ, L is low which explains the decrease of the rate. For the same θ, if we increase C we get larger margins for L and B, and then better rates.

Fig. 7 and 8 represent the variation of the optimal number of pairs and the maximum transmission rate, respectively, in function of the path loss coefficient γ_2. While increasing γ_2, the optimal number of pairs and the maximum rate decrease, because we pass from a low interfering scenario to a highly interfering one.

V. CONCLUSION

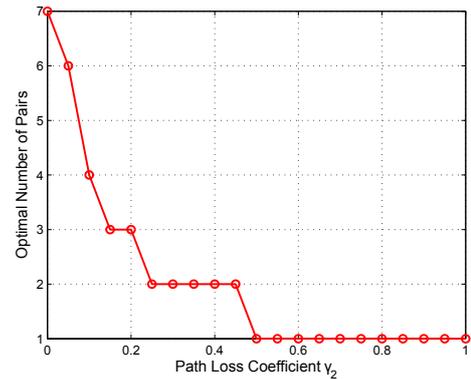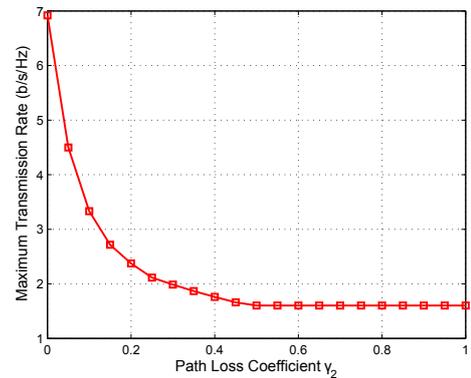In this paper, we consider a TDD system with L-user interference channel under the IA technique. The TDD strategy constrains the maximal number of pairs we can choose. Given the requirement of IA to share CSI between transmitters over backhaul links of finite capacity, we use a quantization scheme to reduce the amount of information to exchange. The throughput of the system is derived and it depends on the number of quantization bits and the number L of pairs. We have investigated the effect of each of these parameters on the variation of the throughput of the system. We have also proposed an algorithm to maximize the transmission rate function of L and B, for a given capacity C. It is shown that this rate is more sensitive to the variation of L and less to B, and that this sensitivity depends on the fraction θ of the slot reserved for probing.

REFERENCES

[1] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug 2008.

[2] T. Gou and S. Jafar, "Degrees of freedom of the K user M×N MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec 2010.

[3] S. Park, O. Simeone, O. Sahin, and S. Shamai, "Performance evaluation of multiterminal backhaul compression for cloud radio access networks," *CoRR*, vol. abs/1311.6492, 2013. [Online]. Available: http://arxiv.org/abs/1311.6492

[4] M. Rezaee, M. Guillaud, and F. Lindqvist, "CSIT sharing over finite capacity backhaul for spatial interference alignment," *CoRR*, vol. abs/1302.1008, 2013.

[5] X. Chen and C. Yuen, "Performance analysis and optimization for interference alignment over MIMO interference channels with limited feedback," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1785–1795, April 2014.

[6] P. Chaporkar and A. Proutiere, "Optimal joint probing and transmission strategy for maximizing throughput in wireless systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1546–1555, Oct. 2008. [Online]. Available: http://dx.doi.org/10.1109/JSAC.2008.081020

[7] O. E. Ayach, A. E. Lozano, and R. W. Heath, "On the overhead of interference alignment: Training, feedback, and cooperation," *CoRR*, vol. abs/1204.6100, 2012. [Online]. Available: http://arxiv.org/abs/1204.6100

[8] K. Huang and V. Lau, "Stability and delay of zero-forcing SDMA with limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6499–6514, Oct 2012.

[9] C. Yetis, T. Gou, S. Jafar, and A. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sept 2010.

[10] O. Ayach and R. Heath, "Interference alignment with analog channel state feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 626–636, February 2012.

[11] B. Jhannesson and N. Giri, "On approximations involving the beta distribution," *Communications in Statistics - Simulation and Computation*, vol. 24, no. 2, pp. 489–503, 1995.

[12] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, September 2007.

[13] S. Nadarajah and S. Kotz*, "On the product and ratio of gamma and beta random variables," *Allgemeines Statistisches Archiv*, vol. 89, no. 4, pp. 435–449, 2005.

[14] H. Bateman, *Tables of Integral Transforms*. New York: McGraw-Hill Book Company, 1954.

APPENDIX

A. Transmission Rate Calculation.

Using (7), the transmission rate that corresponds to stream m of pair k is given by $R_k^m = \Pr(\zeta_k^m \geq \tau)$. In the expression of ζ_k^m given in (4), we denote g as $g = \alpha_{kk} \left| (\hat{\mathbf{u}}_k^m)^* \mathbf{H}_{kk} \hat{\mathbf{v}}_k^m \right|^2$. The random variable $\left| (\hat{\mathbf{u}}_k^m)^* \mathbf{H}_{kk} \hat{\mathbf{v}}_k^m \right|^2$ has an exponential distribution with parameter 1 (see [10]), then g has an exponential distribution with parameter $\frac{1}{\alpha_{kk}}$. Thus, the SINR expression ζ_k^m can be represented as $\zeta_k^m = \frac{g}{I_k^m + \sigma^2}$.

The transmission rate in (7) can be re-expressed as:

$$R_k^m = \Pr(\zeta_k^m \geq \tau)$$
$$= \Pr(\frac{g}{I_k^m + \sigma^2} \geq \tau) = \Pr(g \geq \tau I_k^m + \tau \sigma^2)$$
$$= \int \mathrm{CCDF}_g(\tau I_k^m + \tau \sigma^2) f(I_k^m) dI_k^m, \quad (15)$$

where $\mathrm{CCDF}_g(x) = e^{-\frac{x}{\alpha_{kk}}}$ is the complementary cumulative distribution function of random variable g, and $f(I_k^m)$ is the probability density function of I_k^m. Thus, we get

$$R_k^m = \int \mathrm{CCDF}_g(\tau I_k^m + \tau \sigma^2) f(I_k^m) dI_k^m$$
$$= \int e^{-\frac{(\tau I_k^m + \tau \sigma^2)}{\alpha_{kk}}} f(I_k^m) dI_k^m$$
$$= e^{-\frac{\sigma^2 \tau}{\alpha_{kk}}} \mathrm{MGF}_{I_k^m}(-\frac{\tau}{\alpha_{kk}}). \quad (16)$$

B. Derivation of the Moment generating function of I_k^m.

From (5), $I_k^m = \sum_{\substack{i=1 \\ i \neq k}}^{L} \alpha_{ki} \|\mathbf{h}_{ki}\|^2 e_{ki} \sum_{j=1}^{d_i} \left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2$.

Since $\mathbf{w}_{ki}$ and $\mathbf{s}_{k,i}^{m,j}$ are independent and identically distributed (i.i.d.) isotropic vectors in the null space of $\hat{\mathbf{h}}_{ki}$, $\left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2$ is i.i.d. $\beta(1, Q-1)$ distributed for all i, where $Q = N_t N_r - 1$. $\sum_{j=1}^{d_i} \left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2$ is the sum of i.i.d. Beta variables, which can be approximated to another Beta distribution [11]. Thus, we have $\sum_{j=1}^{d_i} \left| (\mathbf{w}_{ki})^* \mathbf{s}_{k,i}^{m,j} \right|^2 \sim d_i \beta(a_i, b_i)$, in which $a_i = \frac{(Q+1)d_i}{Q} - \frac{1}{Q}$ and $b_i = (Q-1)a_i$.

According to the theory adopted in [12], $e_{ki} \|\mathbf{h}_{ki}\|^2$ has $\Gamma(Q, 2^{\frac{B}{Q}})$ as distribution, where $2^{\frac{B}{Q}}$ is the inverse scale parameter. Let $\delta = 2^{\frac{B}{Q}}$. It follows that $I_k^m = \sum_{\substack{i=1 \\ i \neq k}}^{L} \rho_{ki} X_i Y_i$, where $\rho_{ki} = \alpha_{ki} d_i$, $X_i \sim \Gamma(Q, \delta)$ and $Y_i \sim \beta(a_i, b_i)$. Note that Q and δ are the shape and rate parameters, respectively. $X_i Y_i$ is the product of a Gamma and Beta random variables. Then, the pdf of $Z_i = X_i Y_i$ (for $z_i > 0$) is given by [13] $f_{Z_i}(z_i) = \frac{\delta^Q \Gamma(b_i)}{\Gamma(Q) B(a_i, b_i)} z_i^{Q-1} e^{-\delta z_i} \Psi(b_i, 1 + Q - a_i; \delta z_i)$, where Ψ is the Kummer function defined in [13].

Now, we calculate the MGF of Z_i at $-t$:

$$MGF_{Z_i}(-t) = \int_{-\infty}^{+\infty} e^{-t z_i} f_{Z_i}(z_i) dz_i$$

$$= \kappa \int_0^{+\infty} z_i^{Q-1} e^{-t z_i - \delta z_i} \Psi(b_i, 1 + Q - a_i; \delta z_i) dz_i$$

$$\overset{(i)}{=} \kappa \frac{\Gamma(Q) \Gamma(a_i)}{\delta^Q \Gamma(a_i + b_i)} \left(\frac{t}{\delta} + 1 \right)^{-Q} {}_2F_1(b_i, Q; a_i + b_i; \frac{1}{1 + \frac{\delta}{t}})$$

$$\overset{(ii)}{=} \left(\frac{t}{\delta} + 1 \right)^{-Q} {}_2F_1(b_i, Q; a_i + b_i; \frac{1}{1 + \frac{\delta}{t}}), \quad (17)$$

where $\kappa = \frac{\delta^Q \Gamma(b_i)}{\Gamma(Q) B(a_i, b_i)}$.

The equality (i) is obtained using the relation from [14]. The equality (ii) holds since the Beta function $B(a, b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)}$.

We can write $I_k^m = \sum_{\substack{i=1, i \neq k}}^{L} \rho_{ki} Z_i$, which is the sum of weighed random variables Z_i with ρ_{ki} as weights.

The MGF of I_k^m at $-t$ is then given by:

$$\mathrm{MGF}_{I_k^m}(-t) = \prod_{\substack{i=1 \\ i \neq k}}^{L} MGF_{Z_i}(-t \rho_{ki}) =$$

$$\prod_{\substack{i=1 \\ i \neq k}}^{L} (\frac{\alpha_{ki} d_i t}{\delta} + 1)^{-Q} {}_2F_1(b_i, K; a_i + b_i; \frac{1}{1 + \frac{\delta}{\alpha_{ki} d_i t}}). \quad (18)$$

The desired result follows from (18).