Delay–rate tradeoff in
ergodic interference alignment

Oliver Johnson, Matthew Aldridge, and Robert Piechocki

Abstract

Ergodic interference alignment – as introduced by Nazer et al (NGJV) – is a technique that allows high performance in n-user interference networks with fast fading, which works by matching channel state matrices. However, it comes with the overhead of a long time delay until matchable channel states occur. (The delay is q^{n^2} for field size q – we call n^2 the delay exponent.)

In this paper, we outline new schemes, called JAP and JAP-B that reduce the expected delay, some of them at the cost of a reduction in rate from the NGJV scheme.

In particular, we give examples of good schemes for networks with few users, and show that in large networks, the delay exponent scales quadratically for a constant per-user rate and is constant for a constant sum-rate. We also show that half the single-user rate can be achieved reducing NGJV's delay exponent from n^2 to $(n-1)(n-2)$.

Index Terms

Interference alignment, interference network, multiuser networks, delay–rate tradeoff.

I. INTRODUCTION

Interference alignment [1], [2] describes a set of techniques that allows communication in n-user networks at significantly enhanced rates compared to standard 'resource division' schemes such as TDMA. Interference alignment schemes use the channel so that for each receiver all the interference lies in one subspace and the signal lies in another.

The schemes of Cadambe and Jafar [1] and of Maddah-Ali, Motahari and Khandani [2] apply a channel model divided into time or frequency sub-channels, and involve the inversion and multiplication of large matrices. (These schemes have been implemented by El Ayach, Peters and Heath [3] in the case of $n = 3$, and in both indoor and outdoor environments provide performance close to the degrees of freedom results predicted by theory.)

Given a fast-fading channel – that is a channel with independent and identically distributed (IID) fading coefficients in each time slot – Nazer, Gastpar, Jafar and Vishwanath [4] proposed an alternative scheme, which we call the

O. Johnson and M. Aldridge are with the Department of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, U.K. (e-mail: O.Johnson@bristol.ac.uk; maotj@bristol.ac.uk; m.aldridge@bristol.ac.uk).

R. Piechocki is with the Centre for Communications Research, University of Bristol, Bristol BS8 1UB, U.K. (e-mail: r.j.piechocki@bristol.ac.uk).

NGJV scheme. In brief, the NGJV scheme pairs communications using fading matrix H with those using fading matrix I − H, together providing a situation with no interference between receivers. Using previous results by Nazer and Gastpar concerning the reconstruction of linear combinations of signals [5], the authors provide a description of an achievable rate region. (We outline the NGJV scheme in more detail in Section III.)

Using ideas of 'bottleneck states', Jafar [6] showed that the rate region found by Nazer et al. [4] gives a sharp bound on sum-capacities in an idealized model where all SNRs are identical, and INRs are IID. The present authors [7], [8] showed that such bounds on sum-capacity also hold when nodes are positioned at random in a fixed region of space, in a model similar to that used by Özgur, Lévêque and Tse [9].

These arguments show that the performance of the NGJV scheme can be regarded as optimal, since in a n-user interference network, half the single-user rate is achievable for each user, no matter how large n is. That is 'each user can get half the cake'. However, this is achieved at the cost of a significant delay in communications.

For definiteness, we consider a model of communication over a finite field $\mathbb{F}_q$ of size q. Since the NGJV scheme [4] requires a particular $n \times n$ channel matrix with entries in $\mathbb{F}_q \setminus \{0\}$ to occur, the expected delay for a particular message is $(q - 1)^{n^2}$ (which is roughly q^{n^2} for large q). It is clear that even for n and q relatively small, this is not a practical delay. (For $n = 6$ and $q = 3$, for example, the delay is $2^{36} \approx 7 \times 10^{10}$.)

There are five questions we would like to try to answer:

1) Can we find a scheme that, like NGJV, achieves half the single-user rate, but at a lower time delay?

2) Can we find schemes that have lower time delays than NGJV, even at some cost to the rate achieved?

3) Specifically, which schemes from Question 2 perform well for situations where we have few users (n small)?

4) Specifically, which schemes from Question 2 perform well for situations where we have many users ($n \to \infty$)?

5) What is a lower bound on the best time delay possible for any scheme achieving a given rate for a given number of users?

In Sections IV and V, we define a new set of schemes, called JAP (Subsection IV-B), a beamforming extension JAP-B (Subsection IV-D), and child schemes derived from them (Section V) that have lower time delays than the NGJV scheme, for a variety of different rates, answering Question 2. As a special case, examined in Subsection IV-E, the JAP-B($[n]$) schemes achieve half the single-user rate, like NGJV, whilst reducing the time delay from q^{n^2} to $q^{(n-1)(n-2)}$, answering Question 1. In Section VI, we answer Questions 3 and 4, by finding and analysing the JAP schemes that perform the best for small and large n; Table 1 and Figure 1 illustrate the best schemes for small n, and Theorems 6 and 7 give the asymptotic behaviour of the schemes. Question 5 remains an open problem (although we do give a lower bound on the delay achievable for the schemes listed above).

Koo, Wu and Gill [10] have previously attempted to answer Questions 2 and 3. We briefly outline their work at the end of Section III.

II. MODEL: THE FINITE FIELD CHANNEL

Since ergodic interference alignment relies on matrices being exactly aligned, Nazer et al [4] give their main results in the context of the finite field channel, where there are only finitely-many possible fading matrices. (They

then use a quantisation argument to apply their results to the Gaussian case.) In order to allow comparison of our results, we use the same finite field model.

For a block length N, each source i independently produces a message $\mathbf{w}_i$ belonging to the set $\mathbb{F}_q^{m_i}$ for rate $R_i = (\log q^{m_i})/N = \frac{m_i}{n} \log q$. Each transmitter i separately (without knowledge of the other messages) encodes the message $\mathbf{w}_i$ as a codeword $(x_i[1], \ldots, x_i[N])$.

At time t, receiver j sees channel output

$$Y_j[t] = \sum_{i=1}^{n} H_{ji}[t] x_i[t] + Z_j[t], \tag{1}$$

and needs to decode the message $\mathbf{w}_j$. We can rewrite this equation in matrix form as

$$\mathbf{Y}[t] = \mathsf{H}[t]\mathbf{x}[t] + \mathbf{Z}[t],$$

where we call H the *channel matrix* or *fading matrix*. Here, as in [4], the noise terms $Z_j[t]$ are IID sequences from a distribution on $\mathbb{F}_q$.

(Nazer et al [4] demand that the noise terms are is a mixture of a uniform distribution and a point mass at zero – that is with

$$Z_j = \begin{cases} 0 & \text{w.p. } 1 - \rho, \\ z & \text{w.p. } \rho/(q-1) \text{ for } z = \{1, \ldots, q-1\}, \end{cases} \tag{2}$$

but do not appear to use this hypothesis.)

Like Nazer et al [4], we use an 'ergodic' model, where the channel coefficients $H_{ji}[t]$ are drawn IID and uniformly from the field $\mathbb{F}_q \setminus \{0\}$ and are redrawn for each time slot.

We assume all transmitters and receivers have full causal channel state information for all transmitter–receiver pairs.

By a simple mutual information maximisation, it is easy to show that the the capacity of the single-user finite field channel $Y = hx + Z$ (for $h \neq 0$) is $\log q - H(Z)$. We refer to this quantity $\log q - H(Z)$ as $D(Z)$, since it is the relative entropy $D(Z \| U)$ for U uniform on $\mathbb{F}_q$.

It will turn out that schemes for the finite field interference channel often allow each user pair to achieve a fixed fraction of the single-user rate. We refer to the 'pre-$D(Z)$ term' as the *degrees of freedom*.

Definition 1: Given an achievable symmetric rate point $(R, \ldots, R)$, we define the *degrees of freedom* to be $\mathtt{DOF} = R/D(Z)$.

In particular, it's clear that a single user can achieve 1 degree of freedom.

III. THREE EXISTING SCHEMES:
NGJV, KWG AND TDMA

The description of the NGJV scheme [4] is based on two ideas.

The first idea, from the previous work of Nazer and Gastpar [5], involves the performance of finite-field multiple-access channels. Each receiver's problem takes the form

$$Y[t] = \sum_{i=1}^{n} h_i x_i[t] + Z[t] \tag{3}$$

giving a multiple-access channel with fixed coefficients H_i. Rather than reconstructing a single message $\mathbf{w}_j$ or all messages $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ – as the receiver would normally wish to do – it can be shown [5, Theorem 1], [4, Lemma 3] that the receiver can actually recover the 'pseudomessage' $\sum_{i=1}^{n} h_i \mathbf{w}_i$ at the rate $R = D(Z)$. (This is done by all transmitters using the same linear code, so for all i, $(x_i[1], \ldots, x_i[N]) = \mathsf{G} \mathbf{w}_i$, for an appropriate $m \times N$ generator matrix G.)

The second idea is that of a typical set of channel matrices: essentially, in a long enough sequence of channel matrices, each possible matrix $\mathsf{H} \in (\mathbb{F}_q \setminus \{0\})^{n \times n}$ should occur an approximately equal number of times. This means that transmissions using a particular matrix H can be paired with its complement $\mathsf{I} - \mathsf{H}$, since both occur a roughly equal number of times. (We do not discuss the typicality here, since we simply work in terms of the delay of a single transmission.)

The NGJV scheme works as follows. Each transmitter i sends two signals encoding the same message $\mathbf{w}_i$; first when the channel matrix is H and second when the channel matrix is $\mathsf{H}' := \mathsf{I} - \mathsf{H}$. In the first time period, each receiver j can reliably estimate the pseudomessage $\sum_{i=1}^{n} h_{ji} \mathbf{w}_i$, using the argument above. The receiver stores this pseudomessage in its memory. In the second period they can reliably estimate the pseudomessage $\sum_{i=1}^{n} h'_{ji} \mathbf{w}_i$, which it also stores in its memory. The receiver then adds together these two estimates of pseudomessages, recovering

$$
\sum_{i=1}^{n} h_{ji} \mathbf{w}_i + \sum_{i=1}^{n} h'_{ji} \mathbf{w}_i = \sum_{i=1}^{n} \left(h_{ji} + h'_{ji} \right) \mathbf{w}_i
$$
$$
= \sum_{i=1}^{n} \left(h_{ji} + (\delta_{ji} - h_{ji}) \right) \mathbf{w}_i
$$
$$
= \sum_{i=1}^{n} \delta_{ji} \mathbf{w}_i
$$
$$
= \mathbf{w}_j.
$$

Since both estimates can be made arbitrarily accurate by transmitting at a rate $R < D(Z)$, we conclude that each transmitter can send a message to its receiver at rate $D(Z)/2$ (since two channel uses are needed per message). This corresponds to $\mathtt{DOF} = 1/2$.

(Observe that the receiver needs to perform two separate estimates, and cannot simply add the channel outputs together. To do this would lead to a channel of the form $Y = x_i + (Z * Z)$, where $*$ denotes convolution. The overall rate in this case is $D(Z * Z)/2 \leq D(Z)/2$, with strict inequality unless $D(Z) = 0$ or Z is always 0, and the K-fold convolution has relative entropy $D(Z * Z * \cdots * Z)$ which 'usually' decreases exponentially in K [11].)

We define the expected time delay for the NGJV scheme to be the average number of time slots we must wait after seeing a channel matrix H until we see the corresponding matrix $\mathsf{I} - \mathsf{H}$. The time delay is geometrically distributed with parameter p, where p is the probability that the random channel matrix takes the value $\mathsf{I} - \mathsf{H}$. The mean of this random variable is $1/p$; hence the problem of finding the average time delay is reduced to a problem of finding the probability that a desired matrix appears in the next time slot. Since a channel matrix has n^2 entries,

each of which needs to take the correct one value of $q - 1$ possible values, the average time delay is

$$D = \frac{1}{\left(\frac{1}{q-1}\right)^{n^2}} = (q-1)^{n^2} \sim q^{n^2}.$$

(Here and elsewhere, we write $f(q) \sim g(q)$ if $f(q)/g(q) \to 1$ as $q \to \infty$.)

As we mentioned before, this expected delay will be quite large even for modest values of q and n. For this reason, we will concentrate on the delay exponent.

Definition 2: An interference alignment scheme with expected delay $D \sim Cq^T$ for some C and T has *delay exponent T* and *delay constant C*.

We regard reduction of the delay exponent as the key aim, with the delay coefficient playing a secondary role. In particular, the finite field model is in some sense an abstraction of the model where channel coefficients are Gaussians quantized into a set of size q, where q is chosen chosen large enough to reduce quantization error. When q is large, the delay exponent T dominates the delay constant C in determining size of the expected delay D.

To summarize, the NGJV schemes achieve $\mathtt{DOF} = 1/2$ for a delay exponent of n^2.

For comparison, time-division multiple access (TDMA), where each transmitter–receiver pair has sole access to the channel for an nth of the total time, achieves $\mathtt{DOF} = 1/n$ for a expected delay $D = n = nq^0$, and hence a delay exponent of $T = 0$. To an extent, our new schemes can be seen as 'interpolating' between the extremes of NGJV (high rate, high delay) and TDMA (low rate, low delay).

We also mention some new schemes outlined in a recent paper by Koo, Wu and Gill [10]. They attempted to answer our Questions 2 and 3, by finding schemes – we call them KWG schemes – with lower delay than the NGJV scheme. The KWG schemes suggest matching a larger class of matrices than simply H and $\mathsf{I} - \mathsf{H}$. By analysing the hitting probability of an associated Markov chain, they were able to reduce the expected delay, at the cost of a reduction in rate (and hence degrees of freedom). However, their schemes only affect the delay by a constant multiple, with the most successful scheme [10] only reducing the delay to $0.64(q-1)^{n^2} \sim 0.64q^{n^2}$. That is, the KWG schemes only reduce the delay constant C, leaving the delay exponent as $T = n^2$. For modest q and n (say $q = 3$, $n = 6$, again), we regard this delay as still impractical. Since the KWG schemes achieve a lower rate than the NGJV scheme for the same delay exponent, we shall only compare our results with the NGJV scheme.

IV. NEW ALIGNMENT SCHEMES: JAP AND JAP-B

A. Three important observations

In the NGJV scheme, all receivers were able to decode their message by summing their two pseudomessages

$$\sum_{i=1}^{n} h_{ji}[t_0]\mathbf{w}_i + \sum_{i=1}^{n} h_{ji}[t_1]\mathbf{w}_i = \mathbf{w}_j \quad \text{for } j = 1, \ldots, n.$$

In other words, the NGJV scheme relies on the linear dependence

$$\mathsf{H}[t_0] + \mathsf{H}[t_1] = \mathsf{I}.$$

This scheme has a large delay, because, given $\mathsf{H}[t_0]$, there is only one matrix, $\mathsf{H}[t_1] = \mathsf{I} - \mathsf{H}[t_0]$, that can complete the linear dependence. If there were a large collection of matrices that could complete the dependence, then the delay would be lower.

We make three observations to this end.

First, whilst NGJV matches two channel states $\mathsf{H}[t_0]$ and $\mathsf{H}[t_1]$ to form this linear dependence, we could use more than two. That is, if we have $K + 1$ channel matrices $\mathsf{H}[t_0], \mathsf{H}[t_1], \ldots, \mathsf{H}[t_K]$ such that

$$\mathsf{H}[t_0] + \mathsf{H}[t_1] + \cdots + \mathsf{H}[t_K] = \mathsf{I},$$

than receivers j can sum the $K + 1$ pseudomessages to recover their message,

$$\sum_{i=1}^{n} h_{ji}[t_0]\mathbf{w}_i + \sum_{i=1}^{n} h_{ji}[t_1]\mathbf{w}_i + \cdots + \sum_{i=1}^{n} h_{ji}[t_K]\mathbf{w}_i = \mathbf{w}_j.$$

Note that the transmission of a single message is now split among $K + 1$ channel states, rather than 2 as in NGJV. This means that the degrees of freedom of this scheme is reduced to $1/(K + 1)$ from NGJV's $1/2$.

Second, any linear combination of channel state matrices that sums to I is sufficient. That is, if there exist scalars $\lambda_0, \lambda_1 \in \mathbb{F}_q$ such that

$$\lambda_0 \mathsf{H}[t_0] + \lambda_1 \mathsf{H}[t_1] = \mathsf{I},$$

then all receivers can recover their message by forming the linear combination of pseudocodewords

$$\lambda_0 \sum_{i=1}^{n} h_{ji}[t_0]\mathbf{w}_i + \lambda_1 \sum_{i=1}^{n} h_{ji}[t_1]\mathbf{w}_i = \mathbf{w}_j \quad \text{for } j = 1, \ldots, n.$$

Third, NGJV requires all users to be able to decode their messages at the same time. However, receiver j can decode its message if

$$\sum_{i=1}^{n} h_{ji}[t_0]\mathbf{w}_i + \sum_{i=1}^{n} h_{ji}[t_1]\mathbf{w}_i = \mathbf{w}_j$$

regardless of whether this equality holds for other receivers as well. In other words, receiver j can decode its message if

$$h_{jj}[t_0] + h_{jj}[t_1] = 1$$

$$h_{ji}[t_0] + h_{ji}[t_1] = 0 \quad \text{for } i \neq j.$$

Putting these three observations together, we get the following: Let $\mathsf{H}[t_0], \mathsf{H}[t_1], \ldots, \mathsf{H}[t_K]$ be a sequence of $K + 1$ channel state matrices. If there exist scalars $\lambda_0, \lambda_1, \ldots, \lambda_K$ such that for some j

$$\lambda_0 h_{jj}[t_0] + \lambda_1 h_{jj}[t_1] + \cdots + \lambda_K h_{jj}[t_K] = 1 \tag{4}$$

$$\lambda_0 h_{ji}[t_0] + \lambda_1 h_{ji}[t_1] + \cdots + \lambda_K h_{ji}[t_K] = 0 \quad i \neq j, \tag{5}$$

then receiver j can recover its message by forming the linear combination of pseudocodewords

$$\lambda_0 \sum_{i=1}^{n} h_{ji}[t_0]\mathbf{w}_i + \cdots + \lambda_K \sum_{i=1}^{n} h_{ji}[t_K]\mathbf{w}_i = \mathbf{w}_j.$$

In fact, we only require

$$\lambda_0 h_{jj}[t_0] + \lambda_1 h_{jj}[t_1] + \cdots + \lambda_K h_{jj}[t_K] \neq 0$$

$$\lambda_0 h_{ji}[t_0] + \lambda_1 h_{ji}[t_1] + \cdots + \lambda_K h_{ji}[t_K] = 0 \quad \text{for } i \neq j.$$

since the coefficients λ_k can be rescaled to make the top equation (4) equal to 1 without breaking the bottom equation (5).

Or, writing $\mathbf{h}_j^{\text{int}}$ for the interference vector

$$\mathbf{h}_j^{\text{int}} = (h_{j1}, \ldots, h_{j\,j-1}, h_{j\,j+1}, \ldots, h_{jn}),$$

we can again rewrite the requirement as

$$\lambda_0 h_{jj}[t_0] + \lambda_1 h_{jj}[t_1] + \cdots + \lambda_K h_{jj}[t_K] \neq 0 \tag{6}$$

$$\lambda_0 \mathbf{h}_j^{\text{int}}[t_0] + \lambda_1 \mathbf{h}_j^{\text{int}}[t_1] + \cdots + \lambda_K \mathbf{h}_j^{\text{int}}[t_K] = \mathbf{0}. \tag{7}$$

If n equalities like the (6) and (7) above hold, we say that "receiver j can recover its message from $\mathsf{H}[t_0]$, $\mathsf{H}[t_1], \ldots, \mathsf{H}[t_K]$."

The time delay of this scheme is $t_K - t_0$.

Recall that the average delay is the reciprocal of the probability that a random matrix allows a receiver to recover its message. Thus it will be useful to note the following lemma.

Lemma 3: Conditional on the interference vectors $\mathbf{H}_j^{\text{int}}[t_0], \ldots, \mathbf{H}_j^{\text{int}}[t_K]$ being linearly dependent, the probability that receiver j can recover its message is $1 - O(q^{-1})$.

Proof: Since the interference vectors are linearly dependent, there exists a linear combination

$$\lambda_0 \mathbf{H}_j^{\text{int}}[t_0] + \lambda_1 \mathbf{H}_j^{\text{int}}[t_1] + \cdots + \lambda_K \mathbf{H}_j^{\text{int}}[t_K] = \mathbf{0}$$

where $L > 0$ of the λ_k are nonzero. Thus, receiver j can recover its message provided that the corresponding linear combination

$$\lambda_0 H_{jj}[t_0] + \lambda_1 H_{jj}[t_1] + \cdots + \lambda_K H_{jj}[t_K] \tag{8}$$

is nonzero; call the probability that this happens p.

When $\lambda_k \neq 0$, then $\lambda_k H_{jj}[t_k] =: V_k$ is uniform on $\mathbb{F}_q \setminus \{0\}$; when $\lambda_k = 0$, then $\lambda_k H_{jj}[t_k] = 0$ too. So (8) is the sum of L random variables V_k IID uniform on $\mathbb{F}_q \setminus \{0\}$. We can wrote the mass function of each V_k as $(1 + \rho)U - \rho\delta_0$, where U is uniform on $\mathbb{F}_q$, δ_0 is a point mass on 0 and $\rho = 1/(q - 1)$. Then the mass function of the L-fold convolution is

$$(1 - (-\rho)^L)U + (-\rho)^L \delta_0.$$

Hence, the probability that (8) is zero is

$$1 - p = \left(1 - (-\rho)^L\right)\frac{1}{q} + (-\rho)^L = \frac{1}{q} + \frac{1}{q(q-1)^{L-1}} = O(q^{-1}).$$

The result follows. ∎

B. The scheme JAP($\mathbf{a}$)

We now present our new scheme.

The idea behind the scheme is as follows: We start be seeing some channel state $\mathsf{H}[t_0]$. We then set t_1 to be the first time slot that allows receivers 1 to a_1 to recover their message (where a_1 is decided on in advance). Next, we set t_2 to be the first time slot that allows receivers the next a_2 receivers to recover their message. And so on, until all n receivers have recovered their message.

Specifically, fix $K \leq n$ and a sequence $[a_1, a_2, \ldots, a_K] =: \mathbf{a}$ of length K and weight n; that is, in the set

$$\mathcal{A}(n, K) := \left\{ \mathbf{a} \in \mathbb{Z}_+^K : \sum_{k=1}^K a_k = n \right\}.$$

We write A_k for the partial sums $A_k := a_1 + a_2 + \cdots + a_k$ (so in particular $A_1 = a_1$ and $A_K = n$).

Then we define the scheme JAP($\mathbf{a}$) as consisting of the following $K + 1$ steps:

- **Step 0:** Start with a matrix $\mathsf{H}[t_0]$.
- **Step 1:** Set t_1 to be the first time slot that allows the first a_1 receivers $1, 2, \ldots, A_1$ to recover their message from $\mathsf{H}[t_0], \mathsf{H}[t_1]$.
- ...
- **Step k:** Set t_k to be the first time slot that allows the next a_k receivers $A_{k-1}+1, A_{k-1}+2, \ldots, A_k$ to recover their message from $\mathsf{H}[t_0], \mathsf{H}[t_1], \ldots, \mathsf{H}[t_k]$.
- ...
- **Step K:** Set t_K to be the first time slot that allows the final a_K receivers $A_{K-1} + 1, A_{K-1} + 2, \ldots, A_K$ to recover their message from $\mathsf{H}[t_0], \mathsf{H}[t_1], \ldots, \mathsf{H}[t_K]$.

By the end of this process, all $n = A_K$ receivers have recovered their message.

Since the message was split over $K + 1$ time slots, the common rate of communication is $D(Z)/(K+1)$, which corresponds to $\mathtt{DOF} = 1/(K+1)$.

C. Delay exponent of JAP schemes

We now examine the delay exponent for our new schemes.

Theorem 4: Consider the n-user finite field interference network. Fix K and $\mathbf{a} \in \mathcal{A}(n, K)$. We use the scheme JAP($\mathbf{a}$) as outlined above. Then

1) the expected time for the kth round to take place is $D \sim q^{T_k(\mathbf{a})}$, where $T_k(\mathbf{a}) = a_k(n - k - 1)$;
2) the delay exponent for the whole scheme is

$$T(\mathbf{a}) := \max_k T_k(\mathbf{a}) = \max_k a_k(n - k - 1).$$

Proof: Recall that the expected delay is the reciprocal of the probability the desired match can be made.

Suppose we are about to begin stage k of a scheme JAP($\mathbf{a}$). By Lemma 3, the probability we can complete the stage is $1 - O(q^{-1})$ multiplied by the probability that the interference vectors for the next a_k receivers $\mathbf{H}_j^{\mathrm{int}}[t_0], \ldots, \mathbf{H}_j^{\mathrm{int}}[t_k]$ are linearly dependent.

If the first $k-1$ interference vectors are already linearly dependent, then we are done (with high probability, by Lemma 3). Assume they are not.

Write $\mathcal{S}$ for the span of the first $k-1$ interference vectors for one of the desired a_k receivers j,

$$\mathcal{S} := \operatorname{span}\{\mathbf{H}_j^{\text{int}}[t_0], \ldots, \mathbf{H}_j^{\text{int}}[t_{k-1}]\}.$$

Since all possible interference vectors in $(\mathbb{F}_q \setminus \{0\})^{n-1}$ are equally likely, the probability that the next matrix completes a linear dependence is

$$\frac{|\mathcal{S} \cap (\mathbb{F}_q \setminus \{0\})^k|}{|(\mathbb{F}_q \setminus \{0\})^{n-1}|} = \frac{q^k s}{(q-1)^{n-1}},$$

where s is the proportion of vectors in $\mathcal{S}$ with no zero entries. By counting the possible coefficients in $\mathbb{F}_q$ used in the span, the inclusion–exclusion formula gives us

$$s = 1 - (K-1)\frac{1}{q} + O\left(\frac{1}{q^2}\right) = 1 - O(q^{-1}).$$

Hence, the desired probability is

$$\frac{q^k s}{(q-1)^{n-1}}\left(1 - O(q^{-1})\right) = \frac{q^k\left(1 - O(q^{-1})\right)}{(q-1)^{n-1}}\left(1 - O(q^{-1})\right)$$
$$\sim q^{-(n-k-1)}$$

(where the $1 - O(q^{-1})$ term comes from Lemma 3).

This must hold for all a_k receivers, which happens with probability $q^{-a_k(n-k-1)}$, hence the first result.

For the second result, note that, as $q \to \infty$, the delay is dominated by the delay for the slowest round. ∎

D. Improving delay with beamforming: JAP-B

Beamforming slightly improves the performance of JAP($\mathbf{a}$) schemes, combining ideas from the original Cadambe–Jafar interference alignment [1] with the JAP scheme.

In round k we can guarantee that the interference matches up for receiver $l := A_{k-1} + 1$. Each transmitter i, instead of repeating their message w_i, rather encodes $(h_{li}[t_k]])^{-1}h_{li}[t_0]w_i$. (Since the coefficient h_{li} cannot be 0, the inverse term certainly exists.) The total received interferences at receiver l at times t_0 and t_k are both equal to $\sum_{i \neq l} h_{li}[t_0]w_i$, so can be estimated and cancelled.

We refer to such schemes that take advantage of beamforming as JAP-B($\mathbf{a}$) schemes.

Theorem 5: The delay exponent of a JAP-B($\mathbf{a}$) scheme indexed by sequence $\mathbf{a}$ is

$$T_B(\mathbf{a}) := \max_k (a_k - 1)(n - k - 1).$$

Proof: At each round, receiver $A_{k-1} + 1$ will automatically be able to recover its message, leaving the JAP scheme to align interference for the other $a_k - 1$ users. (Independence of the coefficients h, ensures that the scheme still has the same problem to solve.) ∎

In particular, the JAP-B scheme will always outperform the JAP scheme with the same sequence $\mathbf{a}$.

E. An interesting special case: JAP-B($[n]$)

An interesting special case of the JAP-B schemes is the case when $K = 1$ and $a_1 = n$; we call this scheme JAP-B($[n]$).

In this case we have $1/(K + 1) = 1/2$ degrees of freedom for a rate of $D(Z)/2$. From Theorem 5, we see that the delay exponent is

$$T_B([n]) = (a_1 - 1)(n - 1 - 1) = (n - 1)(n - 2).$$

Effectively, the JAP-B($[n]$) scheme works by using beamforming to automatically cancel user 1's interference, then for users $2, 3, \ldots, n$ requiring the existence of diagonal matrices $\mathsf{D}_0, \mathsf{D}_1$ such that $\mathsf{D}_0\mathsf{H}[t_0] + \mathsf{D}_1\mathsf{H}[t_1] = \mathsf{I}$.

Note that this is the same rate as is achieved by the original NGJV scheme, but that the delay exponent has been reduced from NGJV's n^2 to $(n - 1)(n - 2) = n^2 - (3n - 2)$. For small n in particular, this is a worthwhile improvement (see Figure 1).

V. Child schemes: using time-sharing

Another way to generate new alignment schemes is by time-sharing schemes designed for a smaller number of users.

Call the NGJV, KWG, JAP and JAP-B schemes 'parent schemes'. Given a parent scheme for the m-user network, we can modify for the any n-user network with $n > m$, giving what we call a 'child scheme'.

Specifically, we use TDMA to split the network into $\binom{n}{m}$ subnetworks, each of which contains a unique collection of just $m < n$ of the users. Within each of these m-user subnetworks, a parent scheme is used, whilst the other $n - m$ transmitters remain silent.

Such a child scheme clearly has the same delay exponent as the parent scheme, with the rate – and thus the degrees of freedom – reduced by a factor of m/n. So an m-user JAP-B scheme shared between n users gives $\mathsf{DOF} = m/n(K + 1)$.

In particular, time-sharing the NGJV schemes for smaller networks gives a collection of schemes with a lower delay exponent $m^2 < n^2$ than the main NGJV scheme for a given number of users, reducing the degrees of freedom from $1/2$ to $m/2n$.

(We are not aware that the idea of time-sharing NGJV schemes has previously appeared in the literature. However, the idea seems simple enough that we regard this as the 'current benchmark' against which we should compare our new schemes.)

Interestingly, it seems that child schemes derived from time-sharing an NGJV-like JAP-B($[n]$) parent scheme are particularly effective, and very often performs better than other JAP-B schemes. We discuss this point further in the next section.

VI. BEST SCHEMES

A. General case

Given a number of users n and a desired number of degrees of freedom $\mathrm{DOF} = 1/(K+1)$, we wish to find a scheme with the lowest delay exponent.

For $K = n-1$ or n, when $\mathrm{DOF} = 1/n$ or $1/(n+1)$, the best JAP-B schemes have delay exponent $T_B([1, \ldots, 1, 2]) = T_B([1, \ldots, 1, 1, 1]) = 0$. This is the same delay exponent as TDMA, which has $\mathrm{DOF} = 1/n$ also. Thus we need not consider schemes with $K = n-1$ or n.

For $K \leq n-2$ the best parent scheme will be a JAP-B scheme with parameter vector $\mathbf{a} \in \mathcal{A}(n, K)$. We write $T(n, K)$ for this best delay exponent, that is

$$T(n, K) := \min_{\mathbf{a} \in \mathcal{A}(n,K)} T_B(\mathbf{a})$$

$$= \min_{\mathbf{a} \in \mathcal{A}(n,K)} \max_{k \in \{1,2,\ldots,K\}} (a_k - 1)(n - k - 1).$$

We can bound $T(n, K)$ as follows.

Theorem 6: Fix n and $K \leq n - 2$. For $T(n, K)$ as defined above, we have the following bounds:

$$\frac{n}{K}(n-2) - (2n - K - 2) \leq T(n, K) \leq \frac{n}{K}(n-2)$$

The gap between the upper and lower bounds grows linearly with n.

The following lemma on partial harmonic sums will be useful.

Lemma 7: Let $S(n, K)$ be the partial harmonic sum

$$S(n, K) := \sum_{k=1}^{K} \frac{1}{n-k-1} = \frac{1}{n-2} + \cdots + \frac{1}{n-K-1}.$$

Then we have the bounds

$$\frac{K}{n-2} \leq S(n, K) \leq \frac{K}{n-K-2}.$$

Proof: The proof of the lemma is a standard argument, noting that $1/(n-k-1) \geq 1/(n-2)$, and that for $x \in [m, m+1]$, we have $1/(m+1) \leq 1/x$. ∎

We can now prove Theorem 6.

Proof of Theorem 6: The value of $T(n, K)$ is lower-bounded by the value of the same minimisation problem relaxed to allow the a_k to be real. That is,

$$T(n, K) = \min_{\mathbf{a} \in \mathbb{Z}_+^K : \sum_k a_k = n} \max_{k \in \{1,2,\ldots,K\}} (a_k - 1)(n - k - 1)$$

$$\geq \min_{\mathbf{a} \in \mathbb{R}_+^K : \sum_k a_k = n} \max_{k \in \{1,2,\ldots,K\}} (a_k - 1)(n - k - 1).$$

The relaxed problem is solved by waterfilling, setting $a_k - 1 = c/(n-k-1)$, where requiring the weight of $\mathbf{a}$ to be n forces

$$c = \frac{n-K}{S(n,K)} \geq \frac{(n-K)(n-K-2)}{K},$$

TABLE I

BEST JAP-B($\mathbf{a}$) SCHEMES FOR SMALL VALUES OF n AND K, AND THEIR DELAY EXPONENTS. (ASTERISKS MEAN THAT THE CHOICE OF $\mathbf{a}$ ACHIEVING THIS DELAY EXPONENT IS NON-UNIQUE.)

		$n=3$		$n=4$		$n=5$		$n=6$		$n=7$		$n=8$	
$K=1$	DOF $=1/2$	[3]	2	[4]	6	[5]	12	[6]	20	[7]	30	[8]	42
$K=2$	DOF $=1/3$	TDMA	0	$[1,3]$	2	$[2,3]$	4	$[3,3]$	8	$[3,4]$	12	$[4,4]$	18
$K=3$	DOF $=1/4$			TDMA	0	$[1,1,3]^*$	2	$[1,2,3]^*$	4	$[2,2,3]$	6	$[2,3,3]$	8
$K=4$	DOF $=1/5$					TDMA	0	$[1,1,1,3]^*$	2	$[1,1,2,3]^*$	4	$[1,2,2,3]^*$	6
$K=5$	DOF $=1/6$							TDMA	0	$[1,1,1,1,3]^*$	2	$[1,1,1,2,3]^*$	4
$K=6$	DOF $=1/7$									TDMA	0	$[1,1,1,1,1,3]^*$	2
$K=7$	DOF $=1/8$											TDMA	0

where we have used Lemma 7. Rearrangement gives the lower bound.

An upper bound is obtained by using the same c and taking

$$a_k - 1 = \left\lceil \frac{c}{n-k-1} \right\rceil \leq \frac{c}{n-k-1} + 1.$$

This gives

$$
\begin{aligned}
T_B(\mathbf{a}) &\leq c + \max_k (n-k-1) \\
&= \frac{n-K}{S(n,K)} + (n-2) \\
&\leq \frac{(n-K)(n-2)}{K} + (n-2),
\end{aligned}
$$

where we have used Lemma 7 Rearrangement gives the upper bound. ∎

B. Few users: Small n

For small values of n, we can find the best parent JAP-B schemes by hand. (The task is simplified by noting that the optimal a_k will be nonzero and increasing in k.) Table I gives the delay exponents of the best JAP-B schemes for $n = 3, \ldots, 8$ and $K \leq n-2$.

We can also consider child schemes based on parent JAP-B schemes. Figure 1 plots the performance of NGJV and all JAP-B schemes, as well as child schemes derived from them, for $n = 3, \ldots, 7$. Note that for many values of n and DOF, the scheme with the lowest delay exponent is JAP-B($[n]$) or one of the child schemes derived from it. (Note however, that the the parent schemes with $n = 5, K = 2$ and $n = 8, K = 2$, as well as child schemes derived from them, outperform JAP-B($[n]$) for some degrees of freedom.)
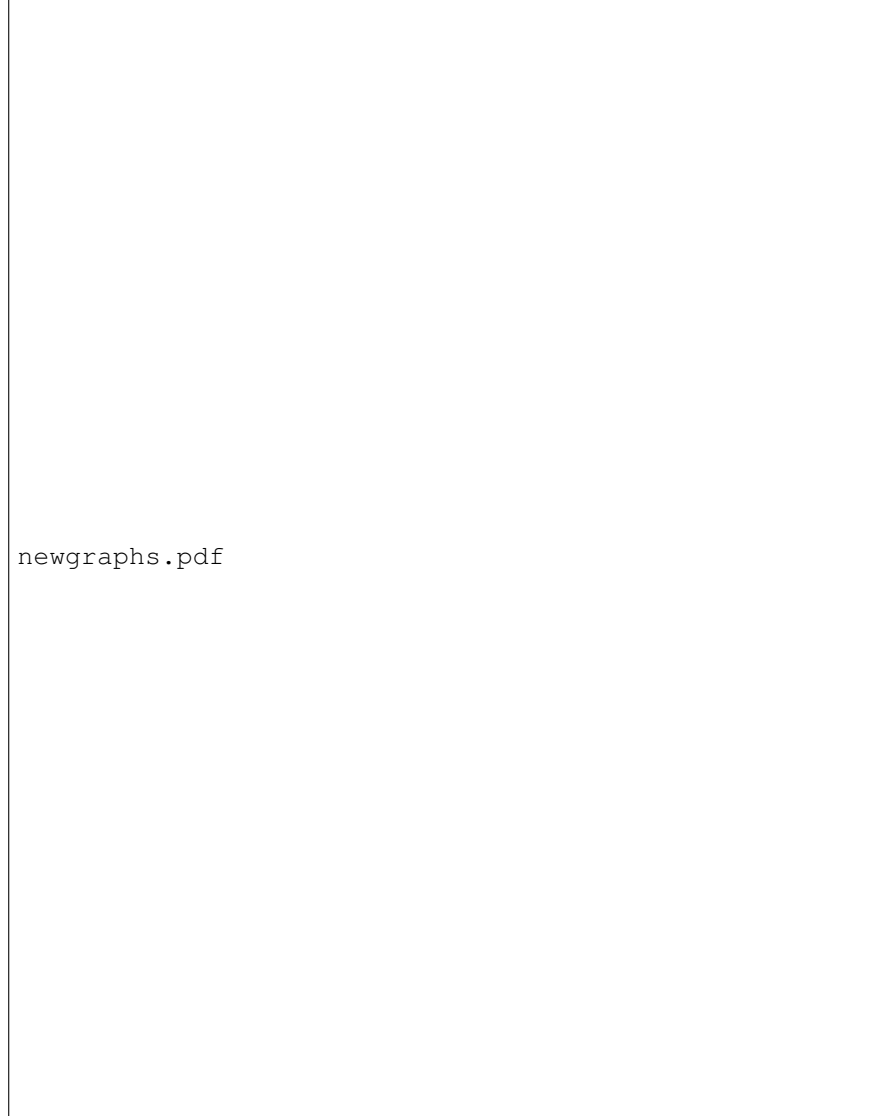
newgraphs.pdf

Fig. 1. Graphs of delay exponent against degrees of freedom NGJV and best JAP-B parent schemes and child schemes derived from them and TDMA.

C. Many users: $n \to \infty$

We now consider the performance of schemes in the many-user limit $n \to \infty$.

In particular, we are interested two limiting regimes, specifying how the degrees of freedom $\text{DOF}(n)$ should scale with the number of users n. In regime I the per-user rate is held constant; in regime II the sum-rate is kept constant, so each user's individual rate falls like $1/n$.

- **Regime I**, where we hold the degrees of freedom constant as $n \to \infty$. That is, we want to communicate at fixed fraction of the single-user rate, as in the NGJV scheme. In this regime I, we take $\text{DOF}(n) = \alpha$ for some $\alpha \in (0, 1/2]$. (The NGJV scheme corresponds to $\alpha = 1/2$.)

- **Regime II**, where we allow the degrees of freedom to fall as the number of users increases, scaling like $1/n$. That is, we want to communicate at a fixed multiple of the rate allowed by resource division schemes like TDMA. In regime II, we take $\text{DOF}(n) = \beta/n$ for some $\beta \geq 1$. (TDMA corresponds to $\beta = 1$.)

First, we consider how parent JAP-B schemes perform in the many-user limit.

Theorem 8: For regimes I and II, as above, and as $n \to \infty$, we have the following results for the delay exponent $T(n)$ of parent JAP-B schemes:

- **Regime I:** Fix $\alpha \in (0, 1/2]$. Then the delay exponent for $\text{DOF}(n) = \alpha$ scales quadratically like

$$T(n) \sim \frac{1}{\lfloor 1/\alpha \rfloor - 1} n^2.$$

- **Regime II:** Fix $\beta > 1$. Then the delay exponent for $\text{DOF}(n) = \beta/n$ scales linearly, in that for any $\epsilon > 0$ and all n sufficiently large,

$$(\beta - 2 - \epsilon)n \leq T(n) \leq (\beta + \epsilon)n.$$

Proof: For regime I, note that $1/(K+1) = \text{DOF} = \alpha$, so we need to take

$$K = \left\lfloor \frac{1}{\alpha} - 1 \right\rfloor = \left\lfloor \frac{1}{\alpha} \right\rfloor - 1.$$

But the general bound on delay exponents from Theorem 6, tells us that for fixed K we have $T(n, K) \sim \frac{1}{K} n^2$. The result follows.

For regime II, $1/(K+1) = \text{DOF} = \beta/n$, so we need

$$K = \left\lfloor \frac{n}{\beta} - 1 \right\rfloor \sim \frac{n}{\beta} - 1 = \frac{1 - \beta/n}{\beta/n},$$

giving

$$\frac{\beta}{1 - \beta/n}(n - 2) - (2n - K - 2) \leq T(n) \leq \frac{\beta}{1 - \beta/n}(n - 2).$$

Since

$$\frac{\beta}{1 - \beta/n} = \beta \left(1 + \frac{\beta}{n} + \left(\frac{\beta}{n} \right)^2 + \cdots \right) = \beta + o(1)$$

and

$$K \sim \frac{1 - \beta/n}{\beta/n} \to 0,$$

the result follows. ∎

Note that in regime I with $\alpha = 1/2$, we get $T(n) \sim n^2$, the same as NGJV.

We noted previously child schemes produced by sharing the parent scheme JAP-B($[m]$) were particularly effective. The following theorem shows this.

Theorem 9: For regimes I and II, as above, and as $n \to \infty$, we have the following results for the delay exponent $T(n)$ of child schemes based on JAP-B($[m]$) parent schemes:

- **Regime I:** Fix $\alpha \in (0, 1/2]$. Then the delay exponent for $\text{DOF}(n) = \alpha$ scales quadratically, in that

$$T(n) = 4\alpha^2 n^2 - 6\alpha n + 2 + o(1) \sim 4\alpha^2 n^2.$$

- **Regime II:** Fix $\beta > 1$. Then the delay exponent for $\mathrm{DOF}(n) = \beta/n$ is constant, in that

$$T(n) = (\lfloor 2\beta \rfloor - 1)(\lfloor 2\beta \rfloor - 2).$$

Proof: Recall from Section V that sharing the scheme JAP-B($[m]$) amongst n users gives $\mathrm{DOF} = m/2n$ for delay exponent $T = (m-1)(m-2)$.

For regime I, note that $m/2n = \mathrm{DOF}(n) = \alpha$, so we need to take $m = \lfloor 2\alpha n \rfloor$, giving $T(n) = (\lfloor 2\alpha n \rfloor - 1)(\lfloor 2\alpha n \rfloor - 2)$. The result follows.

For regime II, note that $m/2n = \mathrm{DOF}(n) = \beta/n$, so we need to take $m = \lfloor 2\beta \rfloor$, giving $T(n) = (\lfloor 2\beta \rfloor - 1)(\lfloor 2\beta \rfloor - 2)$. ∎

Note that asymptotically, this means that in both regimes child schemes from JAP-B($[m]$) parent schemes are asymptotically more effective than any other parent scheme. This is because

$$4\alpha^2 n^2 \leq \frac{1}{\lfloor 1/\alpha \rfloor - 1} n^2$$

(with inequality unless $\alpha = 1/2$, when no child scheme will achieve the desired degrees of freedom) and any constant is less than $(\beta - 2)n$ for n sufficiently large.

Note also that by the same argument as the above proof, sharing the NGJV parent scheme gives $T(n) = 4\alpha^2 n^2$ in regime I – less good than sharing JAP-B($[m]$), but the same to first-order terms.

VII. CONCLUSION

In the introduction to this paper, the questions we attempted to answer were:

1) Can we find a scheme that, like NGJV, achieves half the single-user rate, but at a lower time delay?
2) Can we find schemes that have lower time delays than NGJV, even at some cost to the rate achieved?
3) Specifically, which schemes from Question 2 perform well for situations where we have few users (n small)?
4) Specifically, which schemes from Question 2 perform well for situations where we have many users ($n \to \infty$)?
5) What is a lower bound on the best time delay possible for any scheme achieving a given rate for a given number of users?

In answer to question 2, we defined the new sets of parent schemes JAP and the even more effective JAP-B, and also derived child schemes from them. We noted that these had lower time delays – and sometimes significantly lower – at the costs of some loss in rate (or equivalently degrees of freedom). We saw that the child schemes from JAP-B($[n]$) schemes were often particularly effective.

In answer to question 1, we noted that the JAP-B($[n]$) schemes keep the degrees of freedom to $1/2$ while reducing the delay exponent from n^2 to $(n-1)(n-2) = n^2 - (3n-2)$.

In answer to Questions 3 and 4, we explicitly found the best schemes JAP-B schemes for $n \leq 8$, and analysed the asymptotic behaviour of our schemes as $n \to \infty$.

Question 5 remains an open problem.

ACKNOWLEDGMENTS

REFERENCES

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

[2] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3457 –3470, 2008.

[3] O. El Ayach, S. W. Peters, and R. W. Heath Jr., "The feasibility of interference alignment over measured MIMO-OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4309 –4321 , 2010.

[4] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," in *Proceedings of the IEEE International Symposium on Information Theory, Seoul, Korea, June 2009*, 2009, pp. 1769 –1773. 1769–1773.

[5] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3498 –3516, Oct. 2007.

[6] S. A. Jafar, "The ergodic capacity of interference networks," 2009, in submission, see: `arXiv:0902.0838v1`.

[7] O. T. Johnson, M. P. Aldridge, and R. Piechocki, "Interference alignment-based sum capacity bounds for random dense Gaussian interference networks," to appear in *IEEE Trans. Inform. Theory*, vol. 57, no. 1, 2011, see `arXiv:0909.0641`.

[8] M. P. Aldridge, O. T. Johnson, and R. Piechocki, "Asymptotic sum-capacity of random Gaussian interference networks using interference alignment," in *Proceedings of the IEEE International Symposium on Information Theory, Austin, TX, USA, June 2010*, 2009, pp. 410 –414 .

[9] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3549–3572, 2007.

[10] J. C. Koo, W. Wu, and J. Gill, "Delay-rate tradeoff for ergodic interference alignment in the Gaussian case," 48th Allerton Conference on Communication Control and Computing, 2010, see `arxiv:1001.2582`.

[11] O. T. Johnson and Y. M. Suhov, "Entropy and convergence on compact groups," *J. Theoret. Probab.*, vol. 13, no. 3, pp. 843–857, 2000.