

# Distributed Spatial Multiplexing with 1-Bit Feedback

Jatin Thukral and Helmut Bölcskei

**Abstract**—We analyze the feasibility of distributed spatial multiplexing with limited feedback in a slow-fading interference network with  $MN$  non-cooperating single-antenna sources and  $M$  non-cooperating single-antenna destinations. In particular, we assume that the sources are divided into  $M$  mutually exclusive groups of  $N$  sources each, every group is dedicated to transmit a common message to a unique destination, all transmissions occur concurrently and in the same frequency band and a dedicated 1-bit broadcast feedback channel from each destination to its corresponding group of sources exists. We provide a feedback-based iterative distributed (multi-user) beamforming algorithm, which “learns” the channels between each group of sources and its assigned destination. This algorithm is a straightforward generalization, to the multi-user case, of the feedback-based iterative distributed beamforming algorithm proposed recently by Mudumbai, Hespanha, Madhow and Barriac in [1] for networks with a single group of sources and a single destination. Putting the algorithm into a Markov chain context, we provide a simple convergence proof. We then show that, for  $M$  finite and  $N \rightarrow \infty$ , spatial multiplexing based on the beamforming weights produced by the algorithm achieves full spatial multiplexing gain of  $M$  and full per-stream array gain of  $N$ , provided the time spent “learning” the channels scales linearly in  $N$ . The network is furthermore shown to “crystallize” in the sense that, in the large- $N$  limit, the  $M$  individual fading links not only decouple (as reflected by full spatial multiplexing gain) but also converge to non-fading links. Finally, we quantify the impact of the performance of the iterative distributed beamforming algorithm on the crystallization rate, and we show that the multi-user nature of the network leads to a significant reduction in the crystallization rate, when compared to the  $M = 1$  case.

## I. INTRODUCTION

We consider a special class of interference networks, where  $MN$  non-cooperating sources are divided into  $M$  mutually exclusive groups  $\mathcal{G}_i, i = 1, 2, \dots, M$ , such that the  $N$  sources in the  $i_{th}$  group, denoted as  $\mathcal{S}_i^j, j = 1, 2, \dots, N$ , are dedicated to transmit, through slow-fading channels, a common message to their assigned single-antenna destination  $\mathcal{D}_i$ . All transmissions occur concurrently and in the same frequency band and the destinations  $\mathcal{D}_i$  do not cooperate. This network models the second hop of the coherent multi-user relaying protocol in [2] under the assumption that the first hop transmission is error-free. The results in [2] imply that, for the interference network considered in this paper, for  $M$  fixed and  $N \rightarrow \infty$ , full spatial multiplexing gain of  $M$  and a per-stream (distributed) array gain of  $N$  can be obtained, provided that each source knows the

channel to its assigned destination perfectly. In this paper, we analyze the case where the perfect channel state information assumption is relaxed to having a 1-bit broadcast feedback channel from each destination  $\mathcal{D}_i$  to its sources  $\mathcal{G}_i$ . These broadcast feedback channels are non-interfering. We provide a feedback-based iterative distributed (multi-user) beamforming algorithm, which “learns” the channels between each group of sources and its assigned destination. This algorithm is a straightforward generalization, to the multi-user case, of the feedback-based iterative distributed beamforming algorithm proposed recently by Mudumbai, Hespanha, Madhow and Barriac in [1] for networks with a single group of sources and a single destination. Making the simplifying assumption, compared to [1], of the fading coefficients as well as all the signals being real-valued allows us to put the iterative algorithm into a Markov chain context, thereby setting the stage for a simple convergence proof. We then show that, for  $M$  finite and  $N \rightarrow \infty$ , spatial multiplexing based on the beamforming weights produced by the iterative algorithm achieves full spatial multiplexing gain of  $M$  and full per-stream (distributed) array gain of  $N$ , provided the time spent “learning” the channels scales linearly in  $N$ . We furthermore demonstrate that the  $M$  effective links  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  in the network not only decouple (reflected by full spatial multiplexing gain) but also converge to non-fading links as  $N \rightarrow \infty$ , i.e., in the terminology of [3], the network “crystallizes”. Finally, we quantify the impact of the performance of the iterative algorithm on the crystallization rate, and we show that the multi-user nature of the network leads to a significant reduction in the crystallization rate, when compared to the  $M = 1$  case.

*Notation:* The superscripts  $T$  and  $^{-1}$  stand for transposition and inverse, respectively.  $|\mathcal{G}|$  denotes the cardinality of the set  $\mathcal{G}$ ,  $|x|$  is the absolute value of the scalar  $x$ , and  $\lfloor a \rfloor$  denotes the greatest integer that is smaller than or equal to the real number  $a$ .  $\mathcal{N}(\mu, \sigma^2)$  stands for the normal distribution with mean  $\mu$  and variance  $\sigma^2$ .  $\log(\cdot)$  denotes logarithm to the base 2.  $f_X(\cdot)$  stands for the probability density function (p.d.f.) of the random variable  $X$  and  $X \sim Y$  denotes equivalence in distribution.  $\mathcal{A} \equiv \mathcal{B}$  denotes that the sets (of terminals)  $\mathcal{A}$  and  $\mathcal{B}$  are equal.  $P(\omega)$  is the probability of event  $\omega$ ,  $\mathbb{E}[X]$  and  $\text{VAR}[X]$  are the expected value and the variance, respectively, of the random variable  $X$  and w.p. stands for *with probability*. Since the terminals in  $\mathcal{G}_i$  are assumed to have a common message for their assigned destination  $\mathcal{D}_i$ , we will be using the notation  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  to denote the corresponding single-input single-output link between the group  $\mathcal{G}_i$  and destination  $\mathcal{D}_i$ . Vectors and matrices are set in lower-case and upper-case bold-face letters, respectively.

This research was supported in part by the Swiss National Science Foundation (SNF) under grant No. 200020-109619 and by the Nokia Research Center Helsinki, Finland.

The authors are with ETH Zurich, Switzerland, (Email: {jatin,boelcskei}@nari.ee.ethz.ch)

## II. SYSTEM AND SIGNAL MODEL

We assume that  $x_i[n]$  is the common message of the sources in  $\mathcal{G}_i$  to be transmitted to  $\mathcal{D}_i$ . In the remainder of the paper, we distinguish between a training phase during which the  $N$  scalar channels between  $\mathcal{S}_i^j, j = 1, 2, \dots, N$ , and  $\mathcal{D}_i$  are “learned” for each  $i \in \{1, 2, \dots, M\}$ , and a data transmission phase following the training phase. During the training phase for  $\mathcal{G}_i$  the feedback broadcast channel  $\mathcal{D}_i \rightarrow \mathcal{G}_i$  is used once every frame of  $T_f$  time slots. The  $l$ th frame ( $l = 0, 1, \dots$ ), denoted by  $\mathcal{F}_l$ , consists of the time slots  $n = lT_f, lT_f + 1, \dots, (l+1)T_f - 1$ . In the  $l$ th frame, each source  $\mathcal{S}_i^j$  multiplies the sequence  $x_i[n]$  (which is a training sequence during the training phase) by a corresponding beamforming weight  $\alpha_i^j[l] \in \{1, -1\}$  before transmission; these beamforming weights are kept constant during the entire frame. We furthermore assume that all the channels in the network are flat-fading and remain constant throughout the entire time-interval of interest, i.e., during training and data transmission phases. Throughout the paper, we assume that  $M$  is finite. The corresponding input-output relations are now given by

$$y_i[n] = \left( \sum_{j=1}^N h_{i,i}^j \alpha_i^j[l] \right) x_i[n] + \underbrace{\sum_{\substack{r=1 \\ r \neq i}}^M \left( \sum_{j=1}^N h_{i,r}^j \alpha_r^j[l] \right) x_r[n]}_{\text{interference}} + w_i[n], \quad i = 1, 2, \dots, M, \quad (1)$$

where  $l = \lfloor \frac{n}{T_f} \rfloor$ ,  $y_i[n] \in \mathbb{R}$  denotes the symbol received at  $\mathcal{D}_i$  in the  $n$ th time slot,  $h_{i,r}^j \in \mathbb{R}$  stands for the fading coefficient between  $\mathcal{S}_r^j$  and  $\mathcal{D}_i$  and  $w_i[n]$  denotes the  $\mathcal{N}(0, N_o)$  i.i.d. noise sequence at  $\mathcal{D}_i$ . The fading coefficients  $h_{i,r}^j$ , for all  $i, r, j$ , are assumed i.i.d.  $\mathcal{N}(0, 1)$ . The signals  $x_i[n]$  obey the average power constraint

$$\mathbb{E}[|x_i[n]|^2] \leq \frac{P}{N}, \quad i = 1, 2, \dots, M,$$

so that the average power transmitted by each group  $\mathcal{G}_i$  is limited by  $P$ . During the training phase, the frame-rate sequences  $\alpha_i^j[l]$  are updated based on the 1-bit feedback received at the end of each frame. The goal of this process is to find the beamforming weights  $\alpha_i^j = \text{sign}(h_{i,i}^j)$ . While it is in general difficult to put the individual groups into perfect beamforming configuration, so that

$$\sum_{j=1}^N h_{i,i}^j \alpha_i^j = \sum_{j=1}^N |h_{i,i}^j|, \quad \text{for group } \mathcal{G}_i,$$

we will show that if in each group a sufficient number of sources is in beamforming configuration, full spatial multiplexing gain can be achieved. In what follows, we shall say that in a given group  $\mathcal{G}_i$ , sources satisfying  $h_{i,i}^j \alpha_i^j > 0$  are aligned, whereas sources with  $h_{i,i}^j \alpha_i^j < 0$  are reverse-aligned. We denote the final beamforming weights  $\alpha_i^j$  produced during the training phase as  $\tilde{\alpha}_i^j$ . In the data transmission phase, the groups  $\mathcal{G}_i, i = 1, 2, \dots, M$ , perform co-channel data transmission (i.e., spatial multiplexing) based on  $\tilde{\alpha}_i^j$  so

that the corresponding input-output relation is given by (1) with  $\alpha_i^j[l] = \tilde{\alpha}_i^j$ .

## III. ITERATIVE DISTRIBUTED BEAMFORMING

We shall next describe the algorithm carried out during the training phase. The overall training phase is assumed to consist of  $T_{tr} = k_o MN$  frames (each of which contains  $T_f$  time slots) divided into  $M$  blocks of  $k_o N$  frames each. The role of the parameter  $k_o$  will become clear later. During each of these  $M$  blocks, precisely one of the groups  $\mathcal{G}_i$  follows the three-step iterative distributed beamforming algorithm, described below, while all the other groups of sources  $\mathcal{G}_r, r \neq i$ , remain silent. At the end of the training phase of  $k_o MN$  frames, each of the groups  $\mathcal{G}_i$  is in (close-to) beamforming configuration with respect to (w.r.t.) its assigned destination. The order in which the groups follow the three-step procedure below can be decided offline and communicated to all the nodes in the network. Without loss of generality (w.l.o.g.), we assume that the group  $\mathcal{G}_i$  is being processed during the  $i$ th block defined as the set of frames  $l = (i-1)k_o N, (i-1)k_o N + 1, \dots, ik_o N - 1$ . Since much of the analysis in the current and the next section deals with a single group  $\mathcal{G}_i$  only, we shall consider, w.l.o.g., the group  $\mathcal{G}_1$ , drop the index  $i = 1$  and, wherever appropriate, use the convention  $\mathcal{G} \equiv \mathcal{G}_1, \mathcal{S}^j \equiv \mathcal{S}_1^j, \mathcal{D} \equiv \mathcal{D}_1, \alpha^j = \alpha_1^j, \hat{\alpha}^j = \hat{\alpha}_1^j$  and  $h^j = h_{1,1}^j$ . The three steps carried out by the iterative distributed beamforming algorithm can now be summarized as follows.

- *Step 1. Initialization of the received signal level:* This step pertains to the zeroth frame in the block. Each of the sources  $\mathcal{S}^j$  initializes its beamforming weight according to  $\alpha^j[0] = 1$ , initializes an auxiliary beamforming weight as  $\hat{\alpha}^j[0] = \alpha^j[0]$  and starts transmitting the pilot symbol  $\sqrt{P/N}$ . The corresponding received signal at destination  $\mathcal{D}$  is given by

$$\begin{aligned} y[n] &= \sqrt{\frac{P}{N}} \sum_{j=1}^N h^j \alpha^j[0] + w[n] \\ &= \sqrt{\frac{P}{N}} \sum_{j=1}^N h^j + w[n], \quad n \in \mathcal{F}_0. \end{aligned}$$

The destination  $\mathcal{D}$  then estimates the received signal level by averaging  $y[n]$  over the entire frame, resulting in

$$L^{\text{rx}} = \frac{1}{T_f} \sum_{n \in \mathcal{F}_0} y[n] = \sqrt{\frac{P}{N}} \sum_{j=1}^N h^j.$$

Here, we assumed that the estimate of  $L^{\text{rx}}$  is perfect, which requires that  $T_f$  be sufficiently large. Finally,  $\mathcal{D}$  initializes  $L^{\text{max}} = L^{\text{rx}}$ .

- *Step 2. Iterative distributed beamforming:* This is the iterative step that is performed for each frame, except for the zeroth one, i.e., for  $l = 1, 2, \dots, k_o N - 1$  (recall that the zeroth frame is used to carry out the previous initialization step). The details of this step are as follows. At the beginning of each frame, each of the

sources  $\mathcal{S}^j, j = 1, 2, \dots, N$ , chooses its beamforming weight  $\alpha^j[l]$ , independently across  $j$ , according to:

$$\alpha^j[l] = \begin{cases} \hat{\alpha}^j[l-1], & w.p. \ 1 - \frac{1}{N} \\ -\hat{\alpha}^j[l-1], & w.p. \ \frac{1}{N} \end{cases}. \quad (2)$$

Each of the sources  $\mathcal{S}^j$  then transmits the pilot symbol  $\sqrt{P/N}$  throughout the frame, using the beamforming weight  $\alpha^j[l]$ , i.e.,  $\mathcal{S}^j$  transmits the signal  $\alpha^j[l]\sqrt{P/N}$ . At the end of the frame under consideration,  $\mathcal{D}$  estimates the corresponding received signal level, as in the initialization step, according to<sup>1</sup>

$$L^{\text{rx}} = \frac{1}{T_f} \sum_{n \in \mathcal{F}_i} y[n] = \sqrt{\frac{P}{N}} \sum_{j=1}^N h^j \alpha^j[l] \quad (3)$$

and, through the 1-bit broadcast feedback channel, informs all the sources in  $\mathcal{G}$  whether  $L^{\text{rx}} > L^{\text{max}}$  or not. Based on the received feedback, the sources  $\mathcal{S}^j, j = 1, 2, \dots, N$ , update their auxiliary beamforming weights  $\hat{\alpha}^j[l]$  as follows:

$$\hat{\alpha}^j[l] = \begin{cases} \hat{\alpha}^j[l-1], & \text{if } L^{\text{rx}} \leq L^{\text{max}} \\ \alpha^j[l], & \text{if } L^{\text{rx}} > L^{\text{max}}, \quad j = 1, 2, \dots, N. \end{cases} \quad (4)$$

Finally, if  $L^{\text{rx}} > L^{\text{max}}$ ,  $\mathcal{D}$  updates  $L^{\text{max}} = L^{\text{rx}}$ .

- *Step 3. Silencing:* At the end of the block, the sources in  $\mathcal{G}$  store the current values of their respective auxiliary beamforming weights as  $\tilde{\alpha}^j = \hat{\alpha}^j[k_o N - 1]$  and go silent.

In the proposed protocol, only one group of sources is active during a given block. It is therefore natural to ask whether the groups that have finished their training phase could start their data transmission phase while the remaining groups are “learning” their channels through the three-step procedure above. At first sight, one would be tempted to conclude that such a modified protocol would result in higher spectral efficiency. We note, however, that performing Steps 1 and 2 above in the presence of interference created by the groups already transmitting data would require a longer  $T_f$ , say  $T_f^I$ , in order to achieve the same quality of the received signal level estimate as in the original protocol, i.e., in the absence of interference. The resulting tradeoff can be illustrated roughly by assuming *maximum-likelihood* estimation of the parameter  $L^{\text{rx}}$ . The corresponding variance of the estimation error is given by  $\sigma^2 = \frac{N_o}{T_f}$  in the original protocol and by  $\sigma_M^2 = \frac{(N_o + \sigma_I^2)}{T_f^I}$  in the modified protocol, where, considering the link  $\mathcal{G}_i \rightarrow \mathcal{D}_i$ , we have

$$\sigma_I^2 = \frac{P}{N} \sum_{r=1}^{i-1} \left| \sum_{j=1}^N h_{i,r}^j \tilde{\alpha}_r^j \right|^2. \quad (5)$$

Consequently, if we require the variance of the estimation error in the two protocols to be equal, then

$$\frac{T_f^I}{T_f} = 1 + \frac{\sigma_I^2}{N_o}. \quad (6)$$

<sup>1</sup>Again, we assume the estimate to be perfect.

Assuming that in each of the links already in (close-to) beamforming configuration the corresponding group of sources transmits at a rate of  $R$  bits per time slot, the total number of bits transmitted in the modified protocol during the first  $k_o M N T_f^I$  time slots is given by  $R k_o N T_f^I M (M - 1) / 2$  [bit] whereas it is  $R k_o M^2 N (T_f^I - T_f)$  [bit] in the original protocol. Using (6), this implies that the modified protocol would have a higher spectral efficiency if

$$\frac{\sigma_I^2}{N_o} < 1 - \frac{2}{M+1}. \quad (7)$$

Now, since the weights  $\tilde{\alpha}_i^j$  have been optimized to be in beamforming configuration w.r.t.  $h_{i,i}^j$ , they are independent of  $h_{i,r}^j$  for  $r \neq i$ . Furthermore,  $\tilde{\alpha}_i^j = \text{sign}(h_{i,i}^j)$  and  $\tilde{\alpha}_i^j = -\text{sign}(h_{i,i}^j)$  for  $\mathcal{S}_i^j$  in the set of aligned and reverse-aligned sources, respectively. Since  $P(\text{sign}(h_{i,i}^j) = 1) = P(\text{sign}(h_{i,i}^j) = -1) = 1/2$ , it follows that  $P(\tilde{\alpha}_i^j = 1) = P(\tilde{\alpha}_i^j = -1) = 1/2$ , for all  $j$ . Along with the assumption  $h_{i,r}^j \sim \mathcal{N}(0, 1)$ , for all  $i, r, j$ , we therefore get  $h_{i,r}^j \tilde{\alpha}_r^j \sim h_{i,r}^j$ , for all  $j$  and  $r < i$ , which allows us to conclude that  $\left| \sum_{j=1}^N h_{i,r}^j \tilde{\alpha}_r^j \right|^2, r < i$ , scales linearly in  $N$ . Consequently, it follows from (5) that  $\sigma_I^2$  is proportional to  $P$ , which implies that the condition (7) would not be met for any reasonable choice of  $\text{SNR} = P/N_o$ . We can therefore conclude that, provided the variance of the signal level estimation error is the relevant performance measure, the original protocol would in practice always outperform the modified protocol in terms of spectral efficiency.

#### IV. CONVERGENCE OF THE ITERATIVE DISTRIBUTED BEAMFORMING ALGORITHM

We shall next show that the iterative algorithm described in the previous section converges to the beamforming configuration for  $k_o$  large. The proof is rather straightforward and consists of the following sequence of arguments:

- We start by recognizing that through (2) and (4), the coefficient vector  $\hat{\alpha}[l] = [\hat{\alpha}^1[l] \ \hat{\alpha}^2[l] \ \dots \ \hat{\alpha}^N[l]]^T$  depends only on the coefficient vector in the previous frame, i.e., on  $\hat{\alpha}[l-1]$ . The sequence of vectors  $\hat{\alpha}[l]$  therefore follows a Markov chain.
- Each of the elements in the vector  $\hat{\alpha}[l]$  can take on the values 1 or  $-1$ , which implies that the total number of states of the Markov chain is given by  $2^N$ . Let us denote these states by  $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_{2^N}$ . To each of the states, we can associate a value  $\sum_{j=1}^N h^j \hat{\alpha}^j$ .
- The Markov chain has one special state, namely, when  $\hat{\alpha}^j = \text{sign}(h^j)$ , for all  $j$ . W.l.o.g. we let this state be  $\mathcal{Q}_1$  and note that it corresponds to the beamforming configuration, i.e.,

$$\sum_{j=1}^N h^j \hat{\alpha}^j = \sum_{j=1}^N h^j \text{sign}(h^j) = \sum_{j=1}^N |h^j| \quad (8)$$

which implies that once the system is in state  $\mathcal{Q}_1$  the parameter  $L^{\text{max}}$  is maximized over all states  $\mathcal{Q}_i, i = 1, 2, \dots, 2^N$ . From the update rule (4), we can therefore

conclude that once in  $\mathcal{Q}_1$  the Markov chain will remain in  $\mathcal{Q}_1$ , which implies that  $\mathcal{Q}_1$  is an *absorbing state* (see [4, Def. 11.1]).

- Each of the states  $\mathcal{Q}_r, r \neq 1$ , corresponds to a coefficient vector  $\hat{\alpha}$  with at least one reverse-aligned element. Let us denote the number of reverse-aligned elements corresponding to  $\mathcal{Q}_r$  by  $r_o$ , where  $r_o \geq 1$ . Then, the transition probability from  $\mathcal{Q}_r$  to the absorbing state  $\mathcal{Q}_1$  is  $(1/N)^{r_o}(1 - 1/N)^{N-r_o}$ . We can therefore conclude that the absorbing state can be reached from all the states  $\mathcal{Q}_r$  so that the Markov chain is an *absorbing Markov chain* (see [4, Def. 11.1]). It then follows from [4, Th. 11.3] that the system moves to the absorbing state  $\mathcal{Q}_1$  w.p. 1 as the length of the training phase becomes large. Since  $\mathcal{Q}_1$  corresponds to the beamforming configuration, we can conclude that the iterative algorithm eventually converges to the beamforming configuration.

We emphasize that due to our simplifying assumptions, compared to [1], of the fading coefficients and the signals being real-valued, we were able to prove convergence of the iterative algorithm to the beamforming configuration in a simple fashion using basic results from Markov chain theory. The proof above, however, does not reveal anything about the rate of convergence of the algorithm. It seems difficult to obtain results on the actual convergence rate because the state-transition probabilities in the Markov chain, implicitly, depend on the actual realizations of the fading coefficients  $h^j$ . However, interesting insights into the behavior of the convergence rate, as a function of  $N$ , can be obtained by considering convergence in expectation according to

$$\mathbb{E}\left[\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t]\right] = \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t]\right]\right] \xrightarrow{t \rightarrow k_o N} \mathbb{E}_{h^j}\left[\sum_{S^j \in \mathcal{G}} |h^j|\right]. \quad (9)$$

In particular, we shall show that convergence in the sense of (9) can be obtained if the length of the training phase  $T_{tr}$  scales linearly in  $N$ , i.e.,  $k_o$  is independent of  $N$ . To be precise, we shall establish convergence to the beamforming configuration up to a certain level in the sense that we will allow for a (small) fraction of the sources to be reverse-aligned. The corresponding concept and the associated convergence proof are provided next.

#### A. Convergence to $\epsilon$ -level and convergence proof

In the following, we shall be interested in convergence of the iterative algorithm in the sense of (9) up to a certain level. Concretely, we shall allow that, on average, a (small) fraction  $\epsilon$  of the sources in a given group is reverse-aligned, which trivially implies that  $(1 - \epsilon)$  is the fraction of sources that are aligned, on average. Throughout the paper, we shall assume that  $\epsilon$  is independent of  $N$  and  $2\epsilon N \in \mathbb{Z}^+$ . Denoting the sets

<sup>2</sup>Strictly speaking, for a given  $N$ , this requires that  $\epsilon$  be an integer multiple of  $1/N$ .

of aligned and reverse-aligned sources after the update (4) at the end of the  $t_{th}$  frame as  $\mathcal{A}[t]$  and  $\bar{\mathcal{A}}[t]$ , respectively, we have

$$\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t] = \sum_{S^j \in \mathcal{A}[t]} |h^j| - \sum_{S^j \in \bar{\mathcal{A}}[t]} |h^j|. \quad (10)$$

We say that convergence in the sense of (9) up to an  $\epsilon$ -level has been achieved if

$$\begin{aligned} & \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t]\right]\right] \\ &= \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[\sum_{S^j \in \mathcal{A}[t]} |h^j| - \sum_{S^j \in \bar{\mathcal{A}}[t]} |h^j|\right]\right] \end{aligned} \quad (11)$$

$$\geq N(1 - 2\epsilon) \mathbb{E}[|h^j|]. \quad (12)$$

We can now formalize our convergence result as follows.

*Theorem 1:* For any  $\epsilon_o > 0$  and large  $N$ , the number of iterations required in the second step of the distributed beamforming algorithm in Section III to achieve *convergence to  $\epsilon_o$ -level* is at most  $k_o N$ , where  $k_o$  is a constant independent of  $N$ .

*Proof:* We start by defining

$$S^{\text{new}}[t] = \sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t] - \sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t-1] \quad (13)$$

so that  $S^{\text{new}}[t] \geq 0$ , for all  $t$ . With this definition, we have

$$\begin{aligned} & \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^{k_o N-1}|h^j}\left[\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j\right]\right] \\ &= \sum_{t=1}^{k_o N-1} \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t]\right]\right] + \mathbb{E}_{h^j}\left[\sum_{S^j \in \mathcal{G}} h^j\right] \end{aligned} \quad (14)$$

$$= \sum_{t=1}^{k_o N-1} \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t]\right]\right]. \quad (15)$$

Let there be  $(1 - \epsilon)N$  and  $\epsilon N$  sources in  $\mathcal{A}[t-1]$  and  $\bar{\mathcal{A}}[t-1]$ , respectively, with  $0 < \epsilon_o < \epsilon < 1$ . Moreover, let  $q_a[t]$  and  $\bar{q}_a[t]$  denote the number of sources in  $\mathcal{A}[t-1]$  and  $\bar{\mathcal{A}}[t-1]$ , respectively, that alter their beamforming weights from 1 to  $-1$  or vice-versa according to (2) at the beginning of the  $t_{th}$  frame. Then, we get

$$\begin{aligned} & \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t]\right]\right] \\ &= P(q_a[t] = 0) \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t] | q_a[t] = 0\right]\right] \\ & \quad + P(q_a[t] > 0) \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t] | q_a[t] > 0\right]\right] \\ &\geq P(q_a[t] = 0) \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t] | q_a[t] = 0\right]\right] \\ &= P(q_a[t] = 0) \sum_{s=0}^{\epsilon N} P(\bar{q}_a[t] = s) \\ & \quad \cdot \mathbb{E}_{h^j}\left[\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t|h^j}\left[S^{\text{new}}[t] | q_a[t] = 0, \bar{q}_a[t] = s\right]\right]. \end{aligned} \quad (16)$$

We next show that the expected value inside the summation in (16) can be lower-bounded by  $c_o s$ , where  $c_o$  is a constant independent of  $N$ . To this end, we start by noting that  $q_a[t] = 0, \bar{q}_a[t] = s$  corresponds to the case where precisely  $s$  sources move from  $\bar{\mathcal{A}}[t-1]$  to  $\mathcal{A}[t]$  and none of the sources moves

from  $\mathcal{A}[t-1]$  to  $\bar{\mathcal{A}}[t]$ . We denote the sets of these  $s$  sources as  $\mathcal{R}$ , and note that since the  $s$  sources in  $\mathcal{R}$  are chosen from the  $\epsilon N$  sources in  $\bar{\mathcal{A}}[t-1]$ , there are precisely  $\binom{\epsilon N}{s}$  possible choices for  $\mathcal{R}$ , with the corresponding sets denoted as  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\binom{\epsilon N}{s}}$ . Furthermore, each source in  $\bar{\mathcal{A}}[t-1]$  alters its beamforming weight independently and with equal probability, and hence, each of the  $\binom{\epsilon N}{s}$  choices is equally likely. From (13) it therefore follows that for a given set  $\mathcal{R}_i$ , we have

$$\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t | h^j} [S^{\text{new}}[t] | \mathcal{R}_i] = 2 \sum_{S^j \in \mathcal{R}_i} |h^j| \quad (17)$$

which implies

$$\begin{aligned} & \mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t | h^j} [S^{\text{new}}[t] | q_a[t] = 0, \bar{q}_a[t] = s] \\ &= \frac{1}{\binom{\epsilon N}{s}} \sum_{\mathcal{R}_i \in \{\mathcal{R}_1, \dots, \mathcal{R}_{\binom{\epsilon N}{s}}\}} \mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t | h^j} [S^{\text{new}}[t] | \mathcal{R}_i] \\ &= \frac{2}{\binom{\epsilon N}{s}} \sum_{\mathcal{R}_i \in \{\mathcal{R}_1, \dots, \mathcal{R}_{\binom{\epsilon N}{s}}\}} \sum_{S^j \in \mathcal{R}_i} |h^j| \\ &\stackrel{(a)}{=} \frac{2}{\binom{\epsilon N}{s}} \sum_{S^j \in \bar{\mathcal{A}}[t-1]} \binom{\epsilon N - 1}{s - 1} |h^j| \\ &= \frac{2s}{\epsilon N} \sum_{S^j \in \bar{\mathcal{A}}[t-1]} |h^j| \end{aligned}$$

where Step (a) is a result of the fact that each source in  $\bar{\mathcal{A}}[t-1]$  is present in precisely  $\binom{\epsilon N - 1}{s - 1}$  of the sets  $\mathcal{R}_i$ . We therefore get

$$\begin{aligned} & \mathbb{E}_{h^j} [\mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t | h^j} [S^{\text{new}}[t] | q_a[t] = 0, \bar{q}_a[t] = s]] \\ &= \frac{2s}{\epsilon N} \mathbb{E}_{h^j} \left[ \sum_{S^j \in \bar{\mathcal{A}}[t-1]} |h^j| \right] \\ &\stackrel{(a)}{=} 2s \mathbb{E}_{h^j} [|h^j| | S^j \in \bar{\mathcal{A}}[t-1]] \\ &= 2s \int_x x f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) dx \quad (18) \end{aligned}$$

where Step (a) follows from  $|\bar{\mathcal{A}}[t-1]| = \epsilon N$  and the fact that the  $h^j$  are identically distributed. Next, we need to show that the integral on the right hand side (RHS) of (18) can be lower-bounded by a constant independent of  $N$ . To this end, we start by noting that

$$\begin{aligned} f_{|h^j|}(x) &= P(\bar{\mathcal{A}}[t-1]) f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) \\ &\quad + P(\mathcal{A}[t-1]) f_{|h^j| | \mathcal{A}[t-1]}(x) \\ &\geq P(\bar{\mathcal{A}}[t-1]) f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) \\ &= \epsilon f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) \end{aligned}$$

which, using  $\epsilon > \epsilon_o$ , implies

$$f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) \leq \frac{f_{|h^j|}(x)}{\epsilon_o}. \quad (19)$$

We can then argue that finding the minimum of  $\int_x x f(x) dx$  over the class of functions  $f(x)$  that satisfy

$$\begin{aligned} f(x) &= 0, & x < 0, \\ f(x) &\geq 0, & x \geq 0, \\ f(x) &\leq \frac{f_{|h^j|}(x)}{\epsilon_o} & \text{and} \\ \int_x f(x) dx &= 1 \end{aligned} \quad (20)$$

guarantees that this minimum is a lower bound on  $\int_x x f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x) dx$  as  $f_{|h^j| | \bar{\mathcal{A}}[t-1]}(x)$  is a member of this class of functions. Concretely, we want to determine

$$\min_{f(\cdot)} \int_x x f(x) dx \quad (21)$$

where the minimization is under the constraints (20). This minimization problem can be solved as follows. We start by setting  $x_o = \mathcal{Q}^{-1}(\frac{1-\epsilon_o}{2})$ , where  $\mathcal{Q}(v) = \int_v^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$ , so that using  $f_{|h^j|}(x) = \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}}$ ,  $x \geq 0$ , and  $f_{|h^j|}(x) = 0$ ,  $x < 0$ , we have

$$\int_0^{x_o} \frac{f_{|h^j|}(x)}{\epsilon_o} dx = 1.$$

Next, we note that

$$\begin{aligned} \int_0^\infty x f(x) dx &= \int_0^{x_o} x f(x) dx + \int_{x_o}^\infty x f(x) dx \\ &\geq \int_0^{x_o} x f(x) dx + x_o \int_{x_o}^\infty f(x) dx \\ &= \int_0^{x_o} x f(x) dx + x_o \left[ 1 - \int_0^{x_o} f(x) dx \right] \\ &= \int_0^{x_o} x f(x) dx + x_o \left[ \int_0^{x_o} \frac{f_{|h^j|}(x)}{\epsilon_o} dx - \int_0^{x_o} f(x) dx \right] \\ &= x_o \int_0^{x_o} \frac{f_{|h^j|}(x)}{\epsilon_o} dx - \int_0^{x_o} (x_o - x) f(x) dx \\ &= \int_0^{x_o} x \frac{f_{|h^j|}(x)}{\epsilon_o} dx + \int_0^{x_o} (x_o - x) \frac{f_{|h^j|}(x)}{\epsilon_o} dx \\ &\quad - \int_0^{x_o} (x_o - x) f(x) dx \\ &= \int_0^{x_o} x \frac{f_{|h^j|}(x)}{\epsilon_o} dx + \int_0^{x_o} (x_o - x) \left( \frac{f_{|h^j|}(x)}{\epsilon_o} - f(x) \right) dx. \end{aligned} \quad (22)$$

Since the second term in the last line of (22) is positive due to the constraint  $f(x) \leq (1/\epsilon_o) f_{|h^j|}(x)$ , it follows that setting

$$f(x) = \begin{cases} \frac{f_{|h^j|}(x)}{\epsilon_o}, & 0 \leq x \leq x_o \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

yields the desired lower bound. Further, corresponding to  $f(x)$  in (23), we get

$$\int_x x f(x) dx = \left( \frac{1 - e^{-\frac{(\mathcal{Q}^{-1}(\frac{1-\epsilon_o}{2}))^2}{2}}}{\epsilon_o} \right) \sqrt{\frac{2}{\pi}} \triangleq \frac{c_o}{2}. \quad (24)$$

Substituting (24) into (18) and the result thereof into (16), we obtain

$$\begin{aligned} & \mathbb{E}_{h^j} \left[ \mathbb{E}_{\{\alpha^j[l]\}_{l=0}^t} h^j [S^{\text{new}}[t]] \right] \\ & \geq c_o P(q_a[t] = 0) \sum_{s=0}^{\epsilon N} s P(\bar{q}_a[t] = s) \end{aligned} \quad (25)$$

$$= c_o \left(1 - \frac{1}{N}\right)^{(1-\epsilon)N} \quad (26)$$

$$\cdot \sum_{s=1}^{\epsilon N} s \binom{\epsilon N}{s} \left(\frac{1}{N}\right)^s \left(1 - \frac{1}{N}\right)^{\epsilon N - s} \quad (27)$$

$$= c_o \left(1 - \frac{1}{N}\right)^N \sum_{s=1}^{\epsilon N} s \binom{\epsilon N}{s} \left(\frac{1}{1 - \frac{1}{N}}\right)^s \quad (28)$$

$$= c_o \left(1 - \frac{1}{N}\right)^N \sum_{s=1}^{\epsilon N} s \binom{\epsilon N}{s} \left(\frac{1}{N-1}\right)^s \quad (29)$$

$$= c_o \left(1 - \frac{1}{N}\right)^N \frac{1}{N-1} \epsilon N \left(1 + \frac{1}{N-1}\right)^{\epsilon N - 1} \quad (30)$$

$$\left(\text{since } \sum_{k=1}^K \binom{K}{k} k x^{k-1} = K(1+x)^{K-1}\right)$$

$$= c_o \left(1 - \frac{1}{N}\right)^N \frac{\epsilon N}{N-1} \left(\frac{N}{N-1}\right)^{\epsilon N - 1} \quad (31)$$

$$= c_o \epsilon \left(1 - \frac{1}{N}\right)^{(1-\epsilon)N} \quad (32)$$

$$\approx c_o \epsilon e^{-(1-\epsilon)} \quad (\text{for large } N) \quad (33)$$

$$> c_o \epsilon_o e^{-(1-\epsilon_o)} \quad (\text{since } \epsilon > \epsilon_o). \quad (34)$$

Substituting (34) into (15) finally yields

$$\begin{aligned} \mathbb{E}_{h^j} \left[ \mathbb{E}_{\{\alpha^j[l]\}_{l=0}^{k_o N - 1}} h^j \left[ \sum_{S^j \in \mathcal{G}} h^j \tilde{\alpha}^j \right] \right] & > \sum_{t=1}^{k_o N - 1} \frac{c_o \epsilon_o}{e^{(1-\epsilon_o)}} \\ & = (k_o N - 1) \frac{c_o \epsilon_o}{e^{(1-\epsilon_o)}} \\ & \approx k_o N \frac{c_o \epsilon_o}{e^{(1-\epsilon_o)}} \end{aligned} \quad (35)$$

which upon setting

$$k_o = \frac{(1 - 2\epsilon_o)e^{1-\epsilon_o}}{c_o \epsilon_o} \mathbb{E}[|h^j|]$$

and noting that  $k_o$  does not depend on  $N$  establishes the desired result.  $\square$

### B. Numerical Results

Next, we present simulation results related to the convergence behavior of the iterative distributed beamforming algorithm described in Section III. In particular, Fig.1 shows how  $\mathbb{E}[\sum_{S^j \in \mathcal{G}} h^j \hat{\alpha}^j[t]]$  (averaged over 50 realizations of the fading coefficients  $h^j$ ) evolves as a function of  $t$  for  $N = 100, 500, 1000$  and  $5000$ , respectively. In all four cases, convergence occurs within approximately  $10N$  iterations, thereby corroborating the fact that the convergence time is linear in  $N$ .

## V. ACHIEVABILITY OF MULTIPLEXING GAIN AND CRYSTALLIZATION

The aim of this section is to show that performing data transmission, i.e., spatial multiplexing, based on the beamforming weights  $\tilde{\alpha}_i^j$  obtained from the iterative distributed beamforming algorithm described in Section III results in a spatial multiplexing gain of  $M$  and a per-stream (distributed) array gain proportional to  $N$ . Moreover, using the framework introduced in [3], we prove that the network ‘‘crystallizes’’, i.e., the individual fading links  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  converge to non-fading links as  $N \rightarrow \infty$ . Finally, we quantify the impact of  $M$  and  $\epsilon_o$  on the crystallization rate, i.e., the rate, as a function of  $N$ , at which the fading links  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  converge to non-fading links.

Consider the group  $\mathcal{G}_i$  and let  $\tilde{\mathcal{A}}_i$  and  $\bar{\mathcal{A}}_i$  denote the set of aligned and reverse-aligned sources, respectively, corresponding to the beamforming weights  $\tilde{\alpha}_i^j$ , so that  $|\tilde{\mathcal{A}}_i| = (1 - \epsilon_o)N$  and  $|\bar{\mathcal{A}}_i| = \epsilon_o N$ . Then, the corresponding input-output relations for the individual links  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  are

$$\begin{aligned} y_i[n] & = x_i[n] \left( \sum_{\tilde{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| \right) \\ & + \underbrace{\sum_{\substack{r=1 \\ r \neq i}}^M \left( \sum_{j=1}^N h_{i,r}^j \tilde{\alpha}_r^j \right)}_{\text{interference}} x_r[n] + w_i[n], \quad i = 1, 2, \dots, M. \end{aligned}$$

Note that here we assume that for any given realization of the fading coefficients  $h_{i,i}^j$ , precisely  $(1 - \epsilon_o)N$  of the sources are aligned. Convergence of the iterative distributed beamforming algorithm in expectation according to (12), however, guarantees only that this is the case on average. We furthermore assume Gaussian codebooks in what follows. Under the assumption of the destination  $\mathcal{D}_i$  having perfect knowledge of its effective channel coefficient  $(\sum_{\tilde{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|)$  and of the coefficients  $\sum_{j=1}^N h_{i,r}^j \tilde{\alpha}_r^j$ , for all  $r \neq i$ , corresponding to the effective interference channels, the outage probability for the link  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  is given by

$$P_i^{\text{out}}(R_i) = P\left(\frac{1}{2} \log(1 + \text{SINR}_i) < R_i\right) \quad (36)$$

$$= P(\text{SINR}_i < 2^{2R_i} - 1) \quad (37)$$

$$= P\left(\frac{\frac{P}{N} (\sum_{\tilde{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|)^2}{\sum_{r \neq i} \frac{P}{N} |\sum_j h_{i,r}^j \tilde{\alpha}_r^j|^2 + N_o} < 2^{2R_i} - 1\right) \quad (38)$$

$$= P\left(\frac{(\sum_{\tilde{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|)^2}{\sum_{r \neq i} |\sum_j h_{i,r}^j \tilde{\alpha}_r^j|^2 + \frac{NN_o}{P}} < 2^{2R_i} - 1\right). \quad (39)$$

To upper-bound the outage probability, we use the (union) bounding techniques introduced in [3]. Skipping the details, we note that for any positive  $k_1, k_2$  and  $k_3$  such that  $k_1 > k_2$ ,

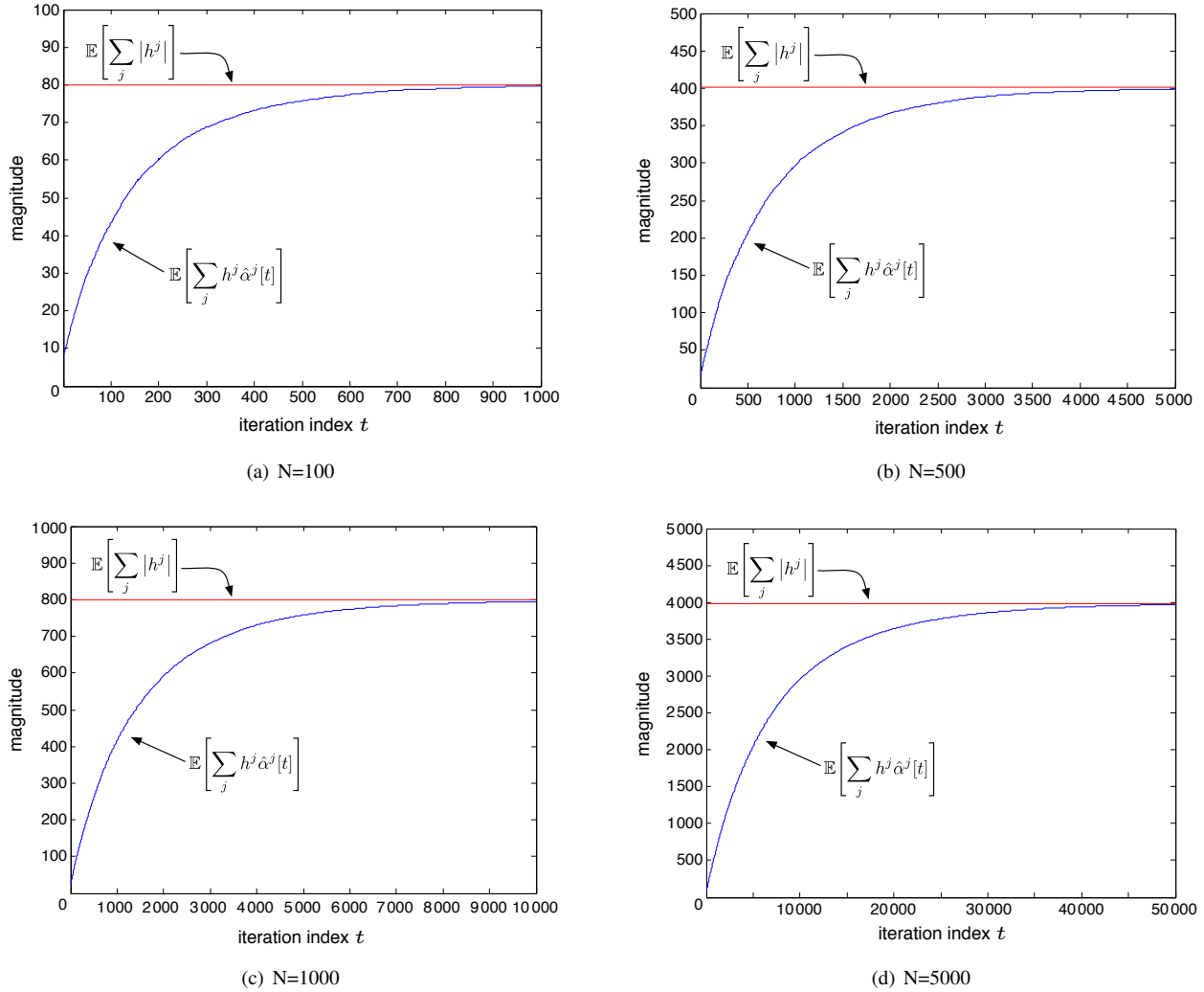


Fig. 1. Convergence of the iterative distributed beamforming algorithm.

the following upper bound holds:

$$\begin{aligned}
& P\left(\frac{\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|\right)^2}{\sum_{r \neq i} \left|\sum_j h_{i,r}^j \tilde{\alpha}_r^j\right|^2 + \frac{NN_o}{P}} < \frac{(k_1 - k_2)^2}{k_3}\right) \\
& \leq P\left(\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|\right)^2 < (k_1 - k_2)^2\right) \\
& \quad + P\left(\sum_{r \neq i} \left|\sum_j h_{i,r}^j \tilde{\alpha}_r^j\right|^2 + \frac{NN_o}{P} > k_3\right) \quad (40) \\
& \leq P\left(\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|\right)^2 < (k_1 - k_2)^2\right) \\
& \quad + \sum_{r \neq i} P\left(\left|\sum_j h_{i,r}^j \tilde{\alpha}_r^j\right|^2 > \frac{k_3 - \frac{NN_o}{P}}{M-1}\right) \quad (41)
\end{aligned}$$

$$\begin{aligned}
& \leq P\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| < k_1\right) + P\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| > k_2\right) \\
& \quad + \sum_{r \neq i} P\left(\left|\sum_j h_{i,r}^j \tilde{\alpha}_r^j\right| > \sqrt{\frac{k_3 - \frac{NN_o}{P}}{M-1}}\right) \quad (42)
\end{aligned}$$

where we implicitly assumed that  $M > 1$  and  $k_3 \geq (NN_o)/P$ . Next, we use the fact, noted previously at the end of Section III, that  $h_{i,r}^j \tilde{\alpha}_r^j \sim h_{i,r}^j$ , for  $r \neq i$ . Consequently, we have

$$\begin{aligned}
& P\left(\frac{\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| - \sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j|\right)^2}{\sum_{r \neq i} \left|\sum_j h_{i,r}^j \tilde{\alpha}_r^j\right|^2 + \frac{NN_o}{P}} < \frac{(k_1 - k_2)^2}{k_3}\right) \\
& \leq P\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| < k_1\right) + P\left(\sum_{\bar{\mathcal{A}}_i} |h_{i,i}^j| > k_2\right) \\
& \quad + (M-1)P\left(\left|\sum_j h_{i,r}^j\right| > \sqrt{\frac{k_3 - \frac{NN_o}{P}}{M-1}}\right). \quad (43)
\end{aligned}$$

Each of the three terms in (43) can be upper-bounded using large deviations bounds [5]. In particular, employing

$$\begin{aligned} P\left(\sum_{i=1}^S |h_i| < k\right) &\leq \left(\frac{ek}{S} \sqrt{\frac{2}{\pi}}\right)^S \\ P\left(\sum_{i=1}^S |h_i| > k\right) &\leq \left(2e^{-\frac{k^2}{2S^2}}\right)^S \\ P\left(\sum_{i=1}^S h_i \geq k\right) &\approx e^{-\frac{k^2}{2S}} \end{aligned}$$

in (43), we get

$$\begin{aligned} P_i^{\text{out}}(R_i) &\leq \left(\frac{ek_1}{(1-\epsilon_o)N} \sqrt{\frac{2}{\pi}}\right)^{(1-\epsilon_o)N} + \left(2e^{-\frac{k_2^2}{2(\epsilon_o N)^2}}\right)^{\epsilon_o N} \\ &\quad + 2(M-1)e^{-\frac{k_3 - \frac{NN_o}{P}}{2N(M-1)}}. \end{aligned} \quad (44)$$

The key to obtaining meaningful upper bounds from (44) lies in a judicious choice of the constants  $k_1, k_2$  and  $k_3$  ensuring that, for  $R_i = (1/2) \log(1 + (k_1 - k_2)^2/k_3)$ ,  $P_i^{\text{out}}(R_i) \rightarrow 0$  as  $N \rightarrow \infty$  while satisfying the conditions  $k_1 > k_2$  and  $k_3 \geq (NN_o)/P$ . Since  $\mathbb{E}\left[\sum_{\bar{A}_i} |h_{i,i}^j|\right] \propto N$  and  $\text{VAR}\left[\sum_{\bar{A}_i} |h_{i,i}^j|\right] \propto N$ , motivated by the above large deviations bounds, it is sensible to set

$$k_1 = \frac{(1-\epsilon_o)(N-\sqrt{N})}{e} \sqrt{\frac{\pi}{2}}. \quad (45)$$

Similarly, since  $\sum_{\bar{A}_i} |h_{i,i}^j|$  deviates around a mean value proportional to  $N$  with a variance proportional to  $N$  and  $|\sum_j h_{i,r}^j|$  deviates around a mean value proportional to  $\sqrt{N}$  with a variance proportional to  $N$ , again motivated by the above large deviations bounds, we set

$$\begin{aligned} k_2 &= \sqrt{2(1+\ln 2)}(\epsilon_o N + \sqrt{N}) \\ \text{and } k_3 &= (M-1)N^{1+\delta} + \frac{NN_o}{P} \end{aligned} \quad (46)$$

with the constant  $\delta > 0$ . Note that the condition  $k_1 > k_2$  implies that  $\epsilon_o < \frac{1}{1+e\sqrt{\frac{4}{\pi}(1+\ln 2)}} = 0.2419$ . With the above choices for the parameters  $k_1, k_2$  and  $k_3$ , in the limit  $N \rightarrow \infty$ , we get

$$\begin{aligned} 2^{2R_i-1} &= \frac{(k_1 - k_2)^2}{k_3} \\ &= \frac{N^{1-\delta}}{(M-1)} \left(\frac{1-\epsilon_o}{e} \sqrt{\frac{\pi}{2}} - \sqrt{2(1+\ln 2)}\epsilon_o\right)^2 \\ &\triangleq c_1 N^{1-\delta} \end{aligned} \quad (47)$$

so that  $R_i = (1/2) \log(1 + c_1 N^{1-\delta})$ . Substituting (45) and (46) into (44), in the large- $N$  limit, we finally obtain

$$\begin{aligned} P_i^{\text{out}} &\left(\frac{1}{2} \log(1 + c_1 N^{1-\delta})\right) \\ &\leq \left(1 - \frac{1}{\sqrt{N}}\right)^N + e^{-\epsilon_o N - 2\sqrt{N}} + 2(M-1)e^{-\frac{N^\delta}{2}} \\ &= e^{-\sqrt{N}} + e^{-\epsilon_o N - 2\sqrt{N}} + 2(M-1)e^{-\frac{N^\delta}{2}}. \end{aligned} \quad (48)$$

We can therefore conclude that  $P_i^{\text{out}}(R_i) \rightarrow 0$  as  $N \rightarrow \infty$  for any rate  $R_i \leq (1/2) \log(1 + c_1 N)$  (recall that  $\delta$  can be arbitrarily small). Since this holds true for all groups  $\mathcal{G}_i$ , we can choose  $R_i = (1/2) \log(1 + c_i N)$  with  $c_i < c_1$ , for all  $i$ , and get  $P_i^{\text{out}}(R_i) \rightarrow 0$ , for all  $i$ , as  $N \rightarrow \infty$ , which implies full spatial multiplexing gain of  $M$ , a per-stream array gain proportional to  $N$ , and convergence of each of the links  $\mathcal{G}_i \rightarrow \mathcal{D}_i$  to a non-fading link. In summary, in the language of [3], we can conclude that the network ‘‘crystallizes’’ as  $N \rightarrow \infty$ . The third term on the RHS of (48) nicely reflects the impact of interference on the crystallization rate. Specifically, for  $\delta < 1/2$ , this term dominates the decay rate as a function of  $N$ . A smaller  $\delta$  corresponds, through  $R_i = (1/2) \log(1 + c_1 N^{1-\delta})$ , to higher data rates, but results in a reduced crystallization rate. In the single-user case, i.e., for  $M = 1$ , the third term equals zero reflecting the absence of interference. We can therefore conclude that the crystallization rate in the presence of interference is significantly smaller than in the single-user case  $M = 1$ . Finally, regarding the proportionality constant  $c_1$ , it can be observed that the smaller  $\epsilon_o$  (i.e., the smaller the fraction of reverse-aligned sources) the larger  $c_1$ , and hence the larger the individual rates  $R_i = (1/2) \log(1 + c_i N)$  still guaranteeing crystallization. On the other hand, for smaller  $\epsilon_o$  the second term on the RHS of (48) becomes larger, again reflecting that a higher data rate comes at the cost of increased outage probability.

## REFERENCES

- [1] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, ‘‘Distributed transmit beamforming using feedback control,’’ *IEEE Trans. Inf. Th.*, 2006, submitted.
- [2] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, ‘‘Capacity scaling laws in MIMO relay networks,’’ *IEEE Trans. Wireless Comm.*, vol. 5, no. 6, pp. 1433–1444, June 2006.
- [3] V. I. Morgenshtern and H. Bölcskei, ‘‘Crystallization in large wireless networks,’’ *IEEE Trans. Inf. Th.*, vol. 53, no. 10, Oct. 2007, to appear.
- [4] C. M. Grinstead, *Introduction to Probability*, 2nd ed. American Mathematical Society, 1997.
- [5] A. Shwartz and A. Weiss, *Large Deviations For Performance Analysis*. Chapman and Hall, London, UK, 1995.