Here we present a correction to "Finding NEMO: Near Mutually Orthogonal Sets and Applications to MIMO Broadcast Scheduling" [1]. These corrections have been made formally in [2]. The bounds in Theorem 1 hold if one studies the Euclidean metric. However, in this paper we must consider the chordal metric, and

$$\delta_c(\theta, m) = B(\sin^2\theta; m-1, 1)$$

where

$$B(z, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^z t^{a-1}(1-t)^{b-1} dt$$

Thus, Lemma 1 should be

Lemma 1: Let $\epsilon, \rho^-, \rho^+ \in \mathbb{R}^+$ and $\epsilon \leq \rho^+$. Then,

$$\begin{aligned} p_\perp &\geq \left(1 - (l-1)B(1 - \epsilon^2/\rho^2, m-1, 1)\right)^{l-1} \\ &= \left(1 - (l-1)\left(1 - \frac{\epsilon^2}{\rho^2}\right)^{m-1}\right)^{l-1} \end{aligned} \quad (1)$$

REFERENCES

[1] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Finding NEMO: Near mutually orthogonal sets and applications to MIMO broadcast scheduling," in *Proc. IEEE WIRELESSCOM 2005 : International Conference on Wireless Networks, Communications, and Mobile Computing*, Maui, Hawaii, USA, June 2005.

[2] ——, "Phase transition phenomena in MIMO broadcast scheduler-multiplexer design," *IEEE Transactions on Information Theory*, 2005, submitted (March 2005).

2005 International Conference on Wireless Networks, Communications and Mobile Computing

Finding NEMO: Near Mutually Orthogonal Sets and Applications to MIMO Broadcast Scheduling

Charles Swannack, Elif Uysal-Biyikoglu and Gregory W. Wornell
Signals, Information and Algorithms Laboratory
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge MA 02139

Abstract— **We define a near-orthogonal set of channel vectors as one that meets certain SIR and SNR guarantees. The probability of finding a near-orthogonal set in a pool of n users is characterized. We identify a phase transition phenomenon in channel geometry whereby this probability transitions from 0 to 1 as k, the number of users that have been examined, increases. It is shown that after this transition the probability of failing to find such a set behaves like $\Theta(k^{-m})$. The rate at which SNR and SIR can be scaled while we remain above this threshold is also characterized. The existence results we provide are not specific to the MIMO scheduling problem, but apply to the more general setting of finding a near-orthogonal set in a random collection of isotropic vectors. The proofs make use of new tight bounds we develop to bound the surface content of spherical caps in arbitrary dimensions. Broader implications of these results are discussed. Specifically, in the case of zero-forcing the best sum rate achievable increases at a rate on the order of $\log \log n$.**

I. Introduction

Developing efficient wireless multiuser communication systems is a problem of substantial interest. An example for such a system is the wireless downlink (as depicted in Figure 1) where independent data streams need to be transmitted to users that are geographically distributed. It is well known that using multiple antennas can greatly increase the capacity of the broadcast channel [1]. Multiplexing users (precoding multiple users' data at the same time) can potentially further increase the throughput of the downlink system. Time variation in the channel states of users leads to the question of which users to choose to encode at a given time to satisfy some overall time-averaged performance criterion. Hence the MIMO broadcast channel contains a quite rich joint scheduling and multiplexing problem when the number of users n is larger than the number of antennas m. A large number of mostly heuristic approaches have been proposed [2]–[5] to explore this multiplexing/scheduling problem space.

This paper has been motivated by the seemingly prohibitive complexity of this joint scheduling/multiplexing problem. In a MIMO channel with choice over users, one expects to improve a particular performance criterion as a larger and larger user pool is searched. This could be maximizing total throughput (or sum rate), for example. The complexity of such an optimization is dominated by the underlying search for the best (possibly ordered) user subset to multiplex across the transmitter array, which must be performed each time the system changes state. To reduce this complexity, one may limit the search to a smaller pool of users while ensuring that a channel set that will be found in this restricted pool is close to optimal with high probability.

We shall define near mutually orthogonal sets to be sets with certain SIR and SNR guarantees. The central goal of the paper is to characterize the probability of finding such a set in a pool of n users. We will show that the probability exhibits a quite sharp transition from 0 to 1 with increasing n, which is a consequence of a phase transition phenomenon in channel geometry. We will obtain upper bounds on the rate at which the SNR and SIR guarantees can be increased while maintaing a high probability that a set with those guarantees exists. More specifically, as a function of the number of users, k, that have been examined, the probability of finding a near-orthogonal set passes through a threshold, after which it behaves like $\Theta(k^{-m})$. This behavior is depicted in Figure 2 for different SIR specifications.

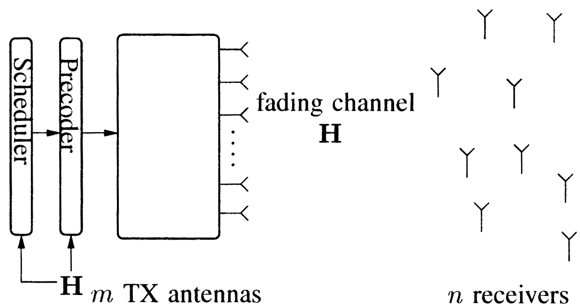

Fig. 1. The MIMO downlink system overview

An outline of the rest of this paper is as follows. We describe our model for the problem in Section II. In section III, near-orthogonal sets are defined, and through examining their geometric properties, the existence probabilities are characterized. To this end, we develop and make use of new tight bounds on the surface content of spherical caps in arbitrary dimensions. Finally, Section IV discusses the broader implications of these results, such as characterizing the throughput of certain low-complexity multiplexing/scheduling techniques such as zero forcing (*i.e.*, interference nulling).

II. System Model

We consider a broadcast channel with an m-antenna transmitter and n uncoordinated receivers each having a single

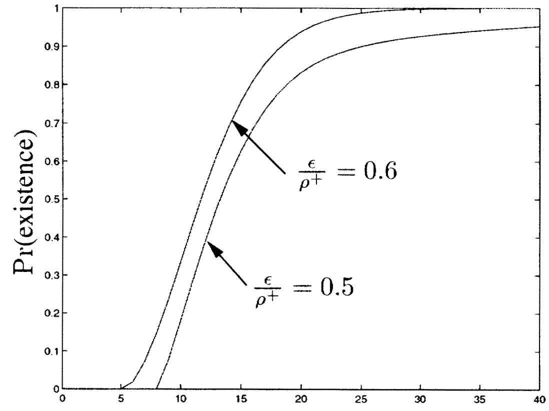

Fig. 2. The probability of finding, among k independent channels, a subset meeting a target SIR requirement. The SIR is lower-bounded by $(1/|\mathcal{A}|)(\epsilon/\rho^-)^{-1}$.

receive antenna. We will denote this set of n users by $\mathcal{U} = \{1, 2, \ldots, n\}$ and let $\mathcal{A} \subset \mathcal{U}$ be an arbitrary subset of users. We will assume the standard input-output model for the channel. Let $\mathbf{x} \in \mathbb{C}^m$ be the transmitted signal vector and $\mathbf{h}_i \in \mathbb{C}^{1\times m}$ be the channel of the ith user. Further, let $\mathbf{H}_\mathcal{A}$ be the channel matrix of the set of users $\mathcal{A}$ and let $\mathbf{x}$ and $\mathbf{y}$ be the input and output respectively. We further assume the channel vectors $\mathbf{h}_i$ are distributed as iid complex Gaussian m-vectors. Under the assumption of complex circularly symmetric Gaussian noise we have,

$$\mathbf{y} = \begin{bmatrix} y_{a_1} \\ \vdots \\ y_{a_{|\mathcal{A}|}} \end{bmatrix} = \mathbf{H}_\mathcal{A}\mathbf{x} + \mathbf{n} \quad \text{where } \mathbf{H}_\mathcal{A} = \begin{bmatrix} \mathbf{h}_{a_1} \\ \vdots \\ \mathbf{h}_{a_{|\mathcal{A}|}} \end{bmatrix}$$

and $n_i \sim \mathcal{N}_\mathbb{C}(0, 1)$. An upper bound on the input covariance is assumed, which corresponds to an total input power constraint of P.

In many applications one is interested in finding a set $\mathcal{A}$ that maximizes some objective function that depends strongly on $\mathbf{H}_\mathcal{A}$. This optimization in general may be very complex, since the user sets do not have an obvious structure. However, if the channel vectors in the set are orthogonal, multiplexing is trivial. More generally, near-orthogonal vectors allow zero-forcing, beamforming or other multiplexers to perform well. In the following section we examine near orthogonal sets as an approximation to any optimal multiplexing set.

III. Near-Orthogonal Sets

In the following, we examine the probability that there is a set $\mathcal{A} \subset \mathcal{U}$ that is near-orthogonal. We will begin by making the definition of an ϵ-orthogonal set, present a geometric interpretation of such a set and discuss some special properties. We will then examine the question of existence of such a set among n independently formed random channels.

Consider the collection of sets $\mathcal{S}_\epsilon$

$$\mathcal{S}_\epsilon = \left\{ \mathcal{A} \,\middle|\, |\mathbf{h}_i\mathbf{h}_j^\dagger| \leq \epsilon \text{ and } \rho^- \leq \|\mathbf{h}_i\|^2 \leq \rho^+ \,, \forall i \neq j \in \mathcal{A} \right\} \tag{1}$$

Note for a given ρ^-, as ϵ decreases toward 0, the channel vectors of any set in $\mathcal{S}_\epsilon$ are increasingly orthogonal. Any set in $\mathcal{S}_\epsilon$ will thus be called an ϵ-orthogonal set. A geometric interpretation is depicted in Figure 3. One can think of any ϵ-orthogonal set as a set of points that lie in the spherical shell between radii ρ^- and ρ^+ such that any two points form an angle no smaller than $\theta_{\epsilon,\rho}$, where

$$\theta_{\epsilon,\rho} \triangleq \cos^{-1}\left(\frac{\epsilon}{\rho^+}\right) \tag{2}$$

From (1), the interference between any two users in a near-orthogonal set is upper-bounded by ϵ, and therefore the SIR for any user is lower-bounded by $\rho^-/(|\mathcal{A}|\epsilon)$. Near-orthogonal sets have been defined in similar ways in [6], [7]. What may seem unusual in the present definition is the upper-bound ρ^+ on channel norms. Indeed, if there was only a single transmit antenna or if the scheduler was constrained to select at most one user at a time (*i.e.* , $|\mathcal{A}| = 1$) such as in [8], limiting oneself to a bounded channel gain would certainly result in a loss.

However, when $|\mathcal{A}| \geq 1$, since a user with a larger channel gain can cause a large interference on other users, the constraint ρ^+ is not by itself a certain restriction. In fact, it is a useful technical constraint in answering the question of main interest: How fast can ρ^- (hence the SNR guarantee) and ρ^+ increase as a function of n for a given ϵ while $\mathcal{S}_\epsilon$ is still non-empty with probability 1 as n tends to infinity? The later part of this section will make this existence question precise.

A. Probability of ϵ-orthogonality and New Bounds on the Content of Spherical Caps

We first derive bounds on the probability that any set is ϵ-orthogonal and use this to further bound the probability of existence of an ϵ-orthogonal set in a pool of n independent channel vectors. Define for a set $\mathcal{A}$ of size l

$$p_\epsilon = \Pr\left(\mathcal{A} \in \mathcal{S}_\epsilon\right) \tag{3}$$

to be the probability that the set is ϵ-orthogonal. Note that we can rewrite p_ϵ as

$$p_\epsilon = p_s^l \, p_\perp \tag{4}$$

where p_s is the probability that a point falls in the spherical shell defined by the parameters ρ^- and ρ^+ and $p_\perp$ is the conditional probability that the set is ϵ-orthogonal given that all the points in the set are inside the spherical shell (Figure 3).

We can lower bound $p_\perp$ by pessimistically assuming all channels to have norm ρ^+, yielding

$$p_\perp \geq \Pr\left(|\tilde{\mathbf{h}}_j\tilde{\mathbf{h}}_i^\dagger| \leq \frac{\epsilon}{\rho^+} \; \forall\, i \neq j \,\middle|\, \{\rho^- \leq \|\mathbf{h}_k\|^2 \leq \rho^+\} \;\forall k \right) \tag{5}$$

where $\tilde{\mathbf{h}}_i = \mathbf{h}_i/\|\mathbf{h}_i\|^2$ and is uniformly distributed on the unit sphere in $\mathbb{C}^m$. Now, $p_\perp$ is simply the probability that each $\mathbf{h}_i$ falls outside of the cone of half angle $\theta_{\epsilon,\rho}$ and apex $\tilde{\mathbf{h}}_i$. Let, $\delta_c(\theta; 2m)$ be the probability of falling in a cone of half angle

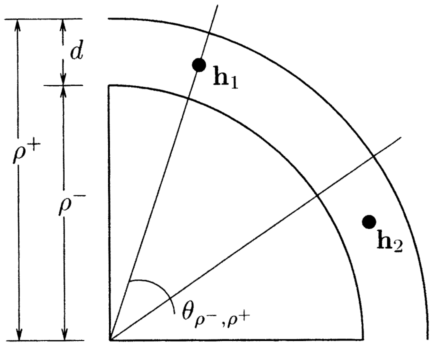

Fig. 3. The geometry of the near-optimal selection procedure

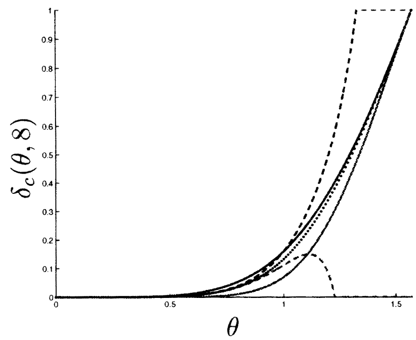

Fig. 4. Bounds on the probability $\delta_c(\theta, 8)$. The new estimates solid; estimates from [9] dashed; exact expression dotted

θ in m complex dimensions as seen in Figure 3. Then, from [9]

$$\delta_c(\theta; 2m) \triangleq 2\frac{\Omega_{2m}(\theta)}{\Omega_{2m}(\pi)} \tag{6}$$

where $\Omega_{2m}(\theta)$ is the area of the spherical cap of half angle θ in $2m$ dimensions.

In [9], Shannon provides bounds on $\delta_c(\theta; m)$. While the bounds of [9] are very tight at small θ, they diverge for the larger values of θ which are of interest to us here. This is due to the fact that large half angles corresponds to low interference. Therefore, we need new estimates on the surface area of spherical caps and their intersection. Below, we provide new upper and lower bounds. Our new bounds are tight for both $\theta = 0$ and $\theta = \pi/2$ and strictly increasing for $\theta \in (0, \frac{\pi}{2})$.

First define the function $c(\theta; s, \beta)$ to be

$$c(\theta; s, \beta) \triangleq \left(\frac{2}{\pi}\theta\right)^s \sin^\beta \theta \quad \text{and let} \quad \psi_m \triangleq \frac{\sqrt{\pi}}{m}\frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}$$

Then we have the following theorem.

Theorem 1: [10] Let $\delta_c(\theta, m)$ be the density of a spherical cap of half angle θ on the unit m sphere. Then, if $\beta \geq m-1$ and $s = \frac{\pi}{2\psi_{m-2}}$ then

$$\delta_c(\theta, m) \geq c(\theta; s, \beta)$$

Furthermore, if $s \leq \frac{\pi}{2\psi_{m-2}}$ and $0 < \beta \leq \frac{(s-1)s}{\pi}$ then

$$\delta_c(\theta, m) \leq c(\theta; s, \beta)$$

The full proof of this theorem is quite lengthy, but essentially follows the following outline. We guess a function $c(\theta; s, \beta)$, and consider the error term $D_m = \delta_c(\theta, m) - c(\theta; s, \beta)$. Then $c(\theta; s, \beta)$ is an upper bound if the error term D_m is negative over a the compact interval of interest. Further, if D_m is zero on the end-points of the interval then D_m must be decreasing at the left end point and increasing at the right endpoint to be an upper bound. Furthermore, if its derivative has exactly one root in this interval, then the error term can never cross zero and is thus an upper bound over the entire interval. The full derivation and other bounds can be found in [10], [11].

A comparison of these bounds using the optimal exponents can be seen in figure 4. Note that these bounds do not diverge at $\theta = \frac{\pi}{2}$. We can use these bounds and a method similar to the method of exhaustion used in [9] to lower bound $p_\perp$. That is, we can lower bound the probability that a set is ϵ-orthogonal by first placing a single point and deleting all points on the sphere that are within an angle less than $\theta_{\epsilon,\rho}$ of the first. Now, the set is ϵ-orthogonal if every point falls outside the spherical caps about all other points in the set. Using the union bound, we obtain the following lemma:

Lemma 1: Let $\epsilon, \rho^-, \rho^+ \in \mathbb{R}^+$ and $\epsilon \leq \rho^+$. Then,

$$p_\perp \geq (1 - (l-1)\delta_c(\theta_{\epsilon,\rho}, 2m))^{l-1} \tag{7}$$

Certainly, there is some angle θ_0 such that the RHS of equation 7 is 0, such that the bound is not useful for $\theta > \theta_0$. However, it should be clear that for any $0 < \theta_{\epsilon,\rho} < \pi/2$ that $p_\perp > 0$. Obtaining tighter bounds for $p_\perp$ is the subject of ongoing work.

B. Existence Probability of a Near-Orthogonal Set

Using the new bounds on spherical caps we address the probability that $\mathcal{S}_\epsilon$ is non empty. To start, we define the indicator random variable

$$1_\mathcal{A} \triangleq \begin{cases} 1 & \text{if } \mathcal{A} \in \mathcal{S}_\epsilon \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

which is one if $\mathcal{A}$ is ϵ-orthogonal and consider the random variable

$$X_l = \sum_{\mathcal{A}:|\mathcal{A}|=l} 1_\mathcal{A} \tag{9}$$

which counts the number of ϵ-orthogonal sets. Note that the random variables $1_\mathcal{A}$ are not independent. Dependence of $1_\mathcal{A}$ and $1_\mathcal{B}$ occurs when $\mathcal{A} \cap \mathcal{B} \neq \emptyset$. We will employ the concept of a *dependency graph* which has been widely used in the areas of geometric graphs and combinatorial probability [12] for which the condition $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ is a special case. A graph G with vertex set $V = V(G)$ is a *dependency graph* of the family of random variables $\{1_\mathcal{A}\}_{\mathcal{A}\in V}$ if for any two disjoint subsets of V, say $A, B \subset V$, the two sub-families $\{1_\mathcal{A}\}_{\mathcal{A}\in A}$ and $\{1_\mathcal{A}\}_{\mathcal{A}\in B}$ are independent. Clearly our independent Gaussian vectors fit this criterion.

Theorem 2: [13] Let $\mathcal{P}_l(\mathcal{U})$ be the collection of all unordered sets of size l on n items and let

$$X = \sum_{\mathcal{A}\in\mathcal{P}_l(\mathcal{U})} 1_\mathcal{A} \tag{10}$$

where $\{1_{\mathcal{A}}\}$ is a family of Bernoulli random variables with $\Pr(1_{\mathcal{A}} = 1) = p$, which are independent if $\mathcal{A} \cap \mathcal{B} = \emptyset$. Then,

$$\Pr(X = 0) \le \exp\left(-\max\left\{2p^2\left\lfloor\frac{n}{l}\right\rfloor, \frac{8p}{25}\frac{\binom{n}{l}}{\binom{n}{l-1}}\right\}\right) \quad (11)$$

Theorem 2 can be employed to address the probability of existence of an ϵ-orthogonal set. First, we take a slightly different exponent than the one in Theorem 2 so that the bound is continuous. Let,

$$E(p,l) \triangleq \max\left\{\frac{2p^2}{l}, \frac{8p}{25l}\left(\frac{l-1}{el}\right)^{l-1}\right\} \quad (12)$$

Now, we use Theorem 2 to bound the probability that there is an ϵ-orthogonal set.

Theorem 3: Let $\epsilon, \rho^-, \rho^+ \in \mathbb{R}^+$. Then if $\epsilon \le \rho^+$

$$\Pr(N_\epsilon^{(l)} > 0) \ge \Pr(N_\rho \ge l) - c_1\left(1 + p_s\left(e^{-E(p_\perp, l)} - 1\right)\right)^n \quad (13)$$

where $c_1 = \exp\frac{2p_\perp^2(l-1)}{l}$ if $E(p_\perp, l) = \frac{2p_\perp^2}{l}$ and $c_1 = 1$ otherwise.

Proof: Note by conditioning on the number of users that fall in the spherical shell defined by ρ^- and ρ^+ we have,

$$\begin{aligned}\Pr(|\mathcal{S}_\epsilon| > 0) &= \sum_{j=l}^{n} \Pr(N_\rho = j)\Pr(X_l > 0 | N_\rho = j) \\ &> \sum_{j=l}^{n}\binom{n}{j}p_s^j(1-p_s)^{n-j}\left(1 - c_1 e^{-jE(p_\perp,l)}\right) \\ &= \Pr(N_\rho \ge l) \\ &\quad - c_1\sum_{j=l}^{n}\binom{n}{j}\left(p_s e^{-E(p_\perp,l)}\right)^j(1-p_s)^{n-j} \\ &> \Pr(N_\rho \ge l) - c_1\left(p_s e^{-E(p_\perp,l)} + (1-p_s)\right)^n\end{aligned}$$

where the constant c_1 appears by bounding $\lfloor\frac{n}{l}\rfloor$ by $\frac{n}{l} - \frac{l-1}{l}$. ■

Theorem 3 provides a lower bound on the probability the there is an ϵ-orthogonal set. That is, it provides a lower bound on the probability that we can meet any given SIR and SNR targets. We now use these bounds to examine the rate at which one can grow the SNR and still expect a non-zero probability.

C. Scaling Laws for Near-Orthogonal Selection

Recall that is our ultimate goal to use the bound of Theorem 3 for user selection. In the following section we will show that the constraint on channel norms can grow at a rate at most $\log n$ for $\Pr(|\mathcal{S}_\epsilon| > 0) > 0$. Before we develop that result, let us discuss for a moment its implications: It implies that SNR increases at most on the order of $\log n$, which will correspond to sum-rate increasing on the order of at most $\log\log n$. As such, if we have that $\Pr(|\mathcal{S}_\epsilon| > 0) > 1 - \delta$ for some small δ we will have to examine exponentially more users to realize a relatively small gain in rate. We will call $n_\delta(\epsilon)$ a threshold if for some given $\delta > 0$, we have $\Pr(|\mathcal{S}_\epsilon| > 0) < 1 - \delta$ for $n < n_\delta$ while $\Pr(|\mathcal{S}_\epsilon| > 0) \ge 1 - \delta$ for $n > n_\delta$. From examining Theorem 3, it is easy to see that n_δ is finite for $p_\perp > 0$ and $p_s > 0$ and that the threshold can be sharp.

The sharpness of this threshold has two possible explanations. First, the number of points that fall in a spherical shell follows a binomial distribution which is known to exhibit a thresholding behavior [12]. Secondly, it is reasonable to expect that there will additionally be a rapid emergence of an ϵ-orthogonal set since it can be roughly modeled by the existence of a clique of size m in a binomial random graph [6], [14]. It is difficult to invert the bound of Theorem 3 to fully characterize n_δ. It is useful, however, to consider the level surfaces where the bound of Theorem 3 is equal to δ. This level surface can be seen plotted as a function of ϵ in the case of four transmit antennas in Figure 5.

The existence of such a threshold implies that the search among the set of n users for a set to maximize the objective function can be restricted to n_δ users. This is very significant, of course, in reducing algorithm complexity, and moreover, it does not result in an appreciable loss in throughput provided that ρ^+ and ρ^- can be increased sufficiently fast, which is what we address next.

Lemma 2: Let $m\rho^+(n) = c\log(n)$ and $m\rho^-(n) = \log(n) - d$ where $c \ge 1$ and $d > 0$. Then,

$$2m\left(e^d - 1\right) \le \lim_{n\to\infty} np_s \le 2me^d$$

Further, if $\log n = o(\rho^-(n))$ then

$$\lim_{n\to\infty} np_s = 0$$

Proof: See appendix I. ■

We can now state our main result.

Theorem 4: Let $\theta_{\epsilon,\rho}$ be fixed. Then if, $p_\perp > 0$ then we can achieve a probability of failure $\delta(n) = \Theta(n^{-m})$ with $m\rho^-(n) = \log(n) - \log(\log(n) + 1)$

Proof: First note that for a Binomial random variable, say N, we have the Chernoff bound [12]

$$\Pr(N \ge l) \ge 1 - \exp\left(-\frac{(np - l)^2}{2np}\right)$$

if $l \le np = \mathbb{E}N$. Thus,

$$\log\left(\frac{1}{\Pr(N_\rho = 0)}\right) \ge \frac{(np_s)^2}{2np_s} \quad (14)$$

Now applying lemma 2, for sufficiently large n

$$\log\left(\frac{1}{\Pr(N_\rho = 0)}\right) \ge m\left(e^d - 1\right)$$

Substituting $d = \log(\log(n) + 1)$ proves that with this choice of ρ^- we have $\Pr(N_\rho = 0) = O(n^{-m})$. We additionally have

$$\Pr(N_\rho = 0) = (1 - p_s)^n \ge \exp\left(-\frac{np_s}{1 - p_s}\right)$$

using the inequality $1 - x \ge \exp\frac{-x}{1-x}$ [12]. Repeating the argument above, $\Pr(N_\rho = 0)$ is easily shown to be $\Omega(n^{-m})$. ■

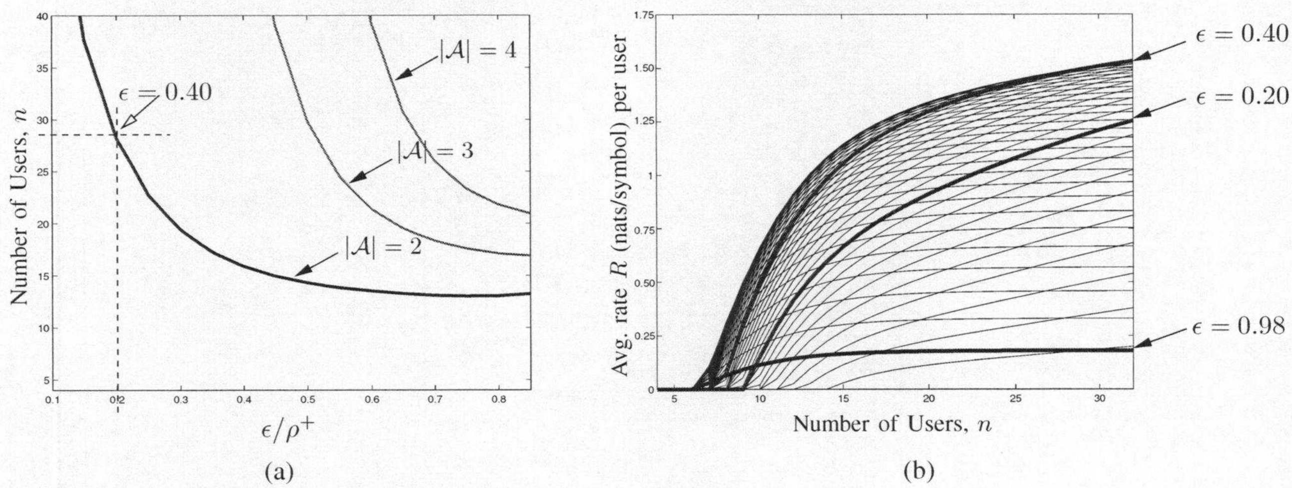

Fig. 5. An example of using ϵ-orthogonal selection for a 4 transmit antenna system with $\rho^- = 1$ and $\rho^+ = 2$ and zero forcing precoding. (a) The level curves for the lower bound on the probability of existence for $\delta = 0.1$. (b) The lower bound on the expected throughput of selecting 2 users for various ϵ. The simulated throughput using optimal selection can be seen above the convex hull of these bounds

Corollary 1: The sequence $m\rho^+(n) = \log n$ and $m\rho^-(n) = \log n - \log(\log(n) + 1)$ achieve a threshold n_δ. Further, for any $0 \leq \theta_{\epsilon,\tau} < \frac{\pi}{2}$ and probability of failure $0 < \delta < 1$ we can achieve as $n \to \infty$

$$\frac{\rho^-(n)}{\rho^+(n)} \to 1$$

We note that the discussion, theorem and corollary say something fundamental about multiple antenna channels in which we can construct our channel matrix. That is in a *channel with choice* we can, for sufficiently large n, consider only channel vectors with norms of the order $\log(n)$ and for any target interference we can achieve a rapidly decaying probability of failure.

IV. Discussion and Conclusions

The probability of existence of a near-orthogonal subset among n independently formed m-dimensional complex circularly symmetric channel vectors have been bounded in the previous section. It is evident from the bounds and the empirical results exhibited in Figure 2 that for any given $\theta_{\epsilon,\rho}$, the probability of finding an ϵ-orthogonal set of channels quickly jumps to 1. This phase transition behavior has direct application to algorithm design for scheduling in the MIMO broadcast channel. Indeed, we have shown that any algorithm that tries to approximate the optimal set using say k users and a criterion based inner products and norms of the channel vectors will have a small probability of success if $k \ll n_\delta$. However, if the same algorithm is employed in the scenario where $k \gg n_\delta$, then we can expect a high probability of success.

The threshold n_δ also has a valuable application in providing lower bounds for the expected rate of any multiplexer in a channel with choice. If we have that $n > n_\delta$, then we can randomly choose any set of users $\mathcal{A} \in \mathcal{S}_\epsilon$ and expect good performance. To be more precise, let f_{rate} be the sum rate expression for a given multiplexer. Then, we can use our SIR and SNR guarantees to provide simple bounds for $f_{\text{rate}}(\mathcal{A})$, say $f_{\text{bnd}}(\text{SNR}, \text{SIR})$. Then,

$$\begin{aligned} \mathbb{E} \max_{\mathcal{A}} f_{\text{rate}}(\mathcal{A}) &\geq \mathbb{E} \max_{\mathcal{A} \in \mathcal{S}_\epsilon} f_{\text{rate}}(\mathcal{A}) && (15) \\ &\geq \Pr(|\mathcal{S}_\epsilon| > 0) f_{\text{bnd}}(\text{SNR}, \text{SIR}) && (16) \\ &\geq (1 - \delta(n)) f_{\text{bnd}}(\text{SNR}, \text{SIR}) && (17) \end{aligned}$$

We can further optimize the above bound over the parameters ϵ, ρ^- and ρ^+. In the case of zero forcing multiplexing, with power constraint P, using the results from [6], it is easy to show that for large n [10]

$$f_{\text{bnd}} = m \log\left(1 + \frac{\text{SNR}}{m}\left(1 - \frac{1}{\text{SIR}}\right)^m\right) \quad (18)$$

Note that there is a tradeoff between the probability of existence and the SIR constraint in (17). That is, the probability of existence is decreasing in SIR while the rate function of (18) is increasing in SIR. An example of how one may achieve a good point on this trade-off (in the case of zero-forcing precoding and four transmit antennas) can be seen in Figure 5. Part (a) of the figure provides contours of equal existence probability 0.9. The curves show that to find an ϵ-orthogonal set with this probability we must either select only pairs of users ($|\mathcal{A}| = 2$) or have $\epsilon/\rho^+ > 0.8$. Clearly, then, a high rate (*i.e.* small ϵ) requires taking $|\mathcal{A}| = 2$ in this example.

One would like to take ϵ at the knee of the level contour. Such an ϵ not only guarantees high rate at a low number of users (which lowers complexity, and, when the number of users can vary, provides robust performance.) The effect of choosing a large ϵ is exhibited on part (b) of the figure where for $\epsilon = 0.98$, the probability quickly jumps to one, but the rate remains a small constant for practically all n thereafter. Alternatively, if ϵ is taken to be too small then the existence probability will be too low, as seen in the $\epsilon = 0.2$ curve: the

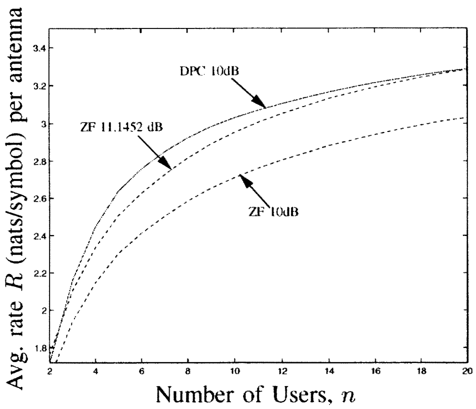

Fig. 6. The empirical average rate of optimal zero forcing multiplexing and dirty paper coding with nominal SNR $= P$ of 10dB

climb in rate indicates that the probability of existence is still increasing.

Now, note that in the right hand side of (18) the SNR is multiplied by a constant that reflects the effects of our choice of target SIR on the expected rate under zero forcing multiplexing. We expect this term to vary depending on how we choose to multiplex our signal. Thus, if we can not grow the SIR to infinity faster than $\log n$ we expect there to be an asymptotic SNR gap in the expected rate under different multiplexers. Figure 6 bears on this question. We continue to explore this using the improved bounds presented in Theorem 1.

Appendix I
Achievable Rates for Channel Scaling

We now prove the rate at which one can hope to scale channel norms and asymptotically have a non-zero probability. In this direction note that from Alzer's bound [15] we have for $m > 1$

$$\left(1 - e^{-s_l x}\right)^m \leq \tilde{\gamma}_{\text{sf}}(m, x) \leq \left(1 - e^{-x}\right)^m$$

where $s_l \triangleq \Gamma(1+m)^{-1/m}$ and

$$\tilde{\gamma}_{\text{sf}}(m, x) = \frac{1}{\Gamma(1+m)} \int_0^x t^{m-1} e^{-t} dt$$

So,

$$\begin{aligned} p_s &\geq \left(1 - e^{-s_l \rho^+}\right)^{2m} - \left(1 - e^{-\rho^-}\right)^{2m} \\ &= \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} \left(e^{j\rho^-} - e^{j s_l \rho^+}\right) \end{aligned}$$

Now we note that in order for the bound to be non-zero we must have $\rho^- < s_l \rho^+$ so that the probability is non-zero. However, implicit in the proof of the bound given in [15] if we replace the constant s_l in the lower bound by any number $s \in (s_l, 1)$ then there exists a x^* such that

$$\left(1 - e^{-sx}\right)^m \leq \tilde{\gamma}_{\text{sf}}(m, x)$$

for all $x \in [x^*, \infty)$. So, asymptotically we can replace the constant s_l by $1 - \epsilon$ for any ϵ such that $1 > \epsilon > 0$. Now, taking $s < 1$ and $\rho^+(n) = c \log n$ and $\rho^-(n) = \log n - d$ yields

$$\begin{aligned} p_s &\geq \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} \left(e^{-j \log n + jd} - e^{-jcs \log n}\right) \\ &\geq \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} e^{-j \log n} \left(e^{jd} - e^{-j(cs-1) \log n}\right) \\ &= \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} n^{-j} \left(e^{jd} - n^{-j(cs-1)}\right) \end{aligned} \tag{19}$$

Thus for $cs < 1$ as $n \to \infty$ then $np_s \to -\infty$ Further, for $cs \geq 1$ as $n \to \infty$ then $2m(e^d - 1) \leq np_s \leq 2me^d$ where the lower bound corresponds to $cs = 1$ and the upper bound corresponds to $cs = \infty$. From the above derivation (interchanging the role of s in the upper and lower bound) it should be clear that if $\log(n) = o(\rho^-(n))$, then

$$\lim_{n \to \infty} np_s \to 0$$

References

[1] I. E. Teletar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, November 1999.

[2] D. Aktas and H. El Gamal, "Multiuser scheduling for multiple antenna systems," vol. 3, pp. 1743–1747, October 2003.

[3] P. Fernandes, L. T. Berger, J. Mrtires, and P. Kyritsi, "Effects of multi user MIMO scheduling freedom on cellular downlink system throughput," in *Proc. IEEE 60th Vehicular Technology Conference*, Los Angeles, USA, September 2004.

[4] S. Serbetli and A. Yener, "Time-slotted multiuser MIMO systems: Beamforming and scheduling strategies," *EURASIP Journal on Wireless Communications and Networking*, no. 2, pp. 286–296, 2004.

[5] M. Airy, S. Shakkottai, and R. W. H. Jr, "Spatially greedy scheduling in multi-user MIMO wireless systems," in *Proc. of IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, November 2003.

[6] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Low complexity multiuser scheduling for maximizing throughput in the MIMO broadcast channel," in *Proc. 42nd Annual Allerton Conf. on Communication, Control, and Computing*, Monticello, Illinois, 2004.

[7] M. A. Maddah-Ali, M. Ansari, and A. K. Khandani, "An efficient signaling method over MIMO broadcast channels," in *Proc. 42nd Annual Allerton Conf. on Communication, Control, and Computing*, Monticello, Illinois, 2004.

[8] B. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Transactions on Information Theory*, 2004, to appear.

[9] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, no. 3, May 1959.

[10] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Phase transition phenomena in MIMO broadcast scheduler-multiplexer design," *IEEE Transactions on Information Theory*, 2005, submitted.

[11] C. Swannack, "Throughput and complexity tradeoffs in the multi-antenna downlink," Master's thesis, MIT, Cambridge, Massachusetts, 2005.

[12] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*. New York: John Wiley, 2000.

[13] S. Janson, "Large deviations for sums of partly dependent random variables," *Random Structures Algorithms*, vol. 24, no. 3, pp. 234–248, 2004.

[14] M. Penrose, *Random Geometric Graphs*, ser. Oxford Studies in Probability. Oxford: Oxford University Press, 2003, no. 5.

[15] H. Alzer, "On some inequalities for the incomplete Gamma function," *Mathematics of Computation*, vol. 66, no. 218, pp. 771–778, April 1997.