Reducing Feedback in Broadcast Channels via Thresholding

Matthew Pugh, *Student Member, IEEE*, and Bhaskar D. Rao, *Fellow, IEEE*
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA, 92093-0407, USA
E-mail: {mopugh, brao}@ucsd.edu

Abstract

In a Gaussian multiple-input multiple-output (MIMO) broadcast channel with random transmit beamforming and a greedy scheduling algorithm, the question of which receivers should feed back their channel state information (CSI) is investigated. A thresholding mechanism is applied to the CSI such that receivers observing CSI lower than the threshold do not feed back. Thresholding is of interest because it reduces the unwanted overhead, which is necessary to achieve multiuser diversity, of CSI feedback in broadcast channels. Utilizing the theory of convergence of types, it is shown that if the thresholding function $T(n)$ scales slower than $\log n$, i.e. $T(n) \in o(\log n)$, that asymptotically in the number of users, no performance is lost compared to the system without a threshold. Conversely, using the asymptotic stability of order statistics, it is shown that if the thresholding function $T(n)$ grows faster than $\log n$, i.e. $T(n) \in \omega(\log n)$, eventually no user in the system will feed back their CSI and all multiuser diversity is lost. For a finite number of users, utilizing the closed form SINR distribution expressions, thresholds are designed to meet specific design criterion as a function of the total number of receivers in the system. Three design criterion are considered: limiting the probability that no receivers feed back information for any transmit direction, limiting the average number of receivers feeding back SINR information and limiting the rate loss due to thresholding.

I. Introduction

The multiuser broadcast channel has been the subject of intense study for many years as it has numerous direct applications to modern technologies such as the cell phone downlink channel.

Great strides have been taken in characterizing the theoretical limits of the broadcast channel, especially the vector Gaussian broadcast channel. Specifically there has been an explosion of work regarding the sum-rate capacity of the vector Gaussian broadcast channel starting with [1] and then followed up by [2], [3], and [4]. The capacity region of the Gaussian vector broadcast channel was characterized in [5]. While of great theoretical importance, many of the schemes used to achieve the theoretical optimal capacity are too complex or impractical to implement in real systems. One of the greatest hinderances in achieving the theoretical optimum is having complete CSI for all of the users in the multiuser broadcast channel available at the transmitter. It is well known that without CSI at the transmitter, the system cannot achieve multiuser diversity, which is a significant benefit of multiuser systems. In order to reap the benefits of the more advanced multiuser MIMO architecture, CSI feedback from the receivers to the transmitter is essential.

To address this critical problem, researchers have developed different methods of quantifying important information to feedback and how to best feed back this information. The problem becomes more complicated when one also includes considerations about the reliability of the CSI due to channel estimation errors and channel dynamics. In [6],[7] and [8] the effects of imperfect channel estimation, feedback delay, and finite-rate quantization are considered on bit, symbol, and packet error probabilities in multiple-input single-output (MISO) channels. The effects of imperfect channel estimation and feedback delay on transmit beamforming is considered in [9] and [10].

In this contribution, channel estimation is assumed to be perfect and the feedback channel is delay and error free. Much reseach has been conducted under these assumptions. In [11], the channel vector from the multi-antenna transmitter to the single receive antenna users is quantized using a random vector quantizer and the performance is analyzed while [12] considers codebook design using the Grassmannian manifold. The asymptotic optimality of zero-forcing beamforming is studied in [13]. The question of how much feedback is required so that the system does not reach a sum-rate ceiling is considered in [14].

The first steps towards thresholding feedback are in the works [15], [16], [17] which consider one bit feedback schemes where a binary decision is made as to whether or not the channel is in a good or bad state for transmission. These works were later extended in [18], where the feedback channel is used multiple times per coherence time to allow for refinement of the

scheduling decision. The rate enhancement due to additional feedback bits is analyzed in [19]. Additional work along these lines can be found in [20] as well as an overview of feedback systems in general in [21].

The tranmission scheme analyzed in this contribution is the random beamforming scheme proposed in [22]. The main idea is for the transmitter to transmit on the basis vectors of a random orthonormal basis and each receiver measures the SINR along each of the random directions. The receivers feed back their measured SINR values and the transmitter schedules the users with the largest SINR value for each orthonormal direction. In this transmission scheme, all of the channel information is captured in the SINR values. This scheme is extended in [23] in several ways, including more advanced receive structures, namely LMMSE receivers, when each user is allowed multiple receive antennas. The main result of these works is that random beamforming is simple yet, asymptotically in the number of users, achieves the optimal sum-rate throughput scaling.

In the aforementioned works, all the users in the system feed back some form of CSI. This contribution addresses the question of which user should feed back CSI. To answer this question, the random beamforming scheme is considered. In the random beamforming scheme, the statistics of the SINR values, which is the CSI, are known (e.g. see [22], [23]). With knowledge of the CSI statistics, consider the example where a particular user in the system observers a small SINR value. Should that user feed back this small SINR value? If there are a large number of users in the system, then it is highly probable that someone will have a larger SINR value. In this case it does not affect the system if the small SINR value is not fed back due to the scheduling algorithm always selecting the largest SINR. On the other hand, it is always possible that every other user in the system has a smaller SINR value, in which case not feeding back will prevent the system from achieving the best throughput. The goal of this work is to design a threshold that achieves certain performance metrics as a function of the number of users in the system such that a user feeds back its observed SINR value if the value is above the threshold and does not feed back if the SINR is below the threshold. By allowing users to not feed back, transmit power can be saved as well as relieving congestion on the radio resource.

The contributions of this work can be broken into two parts. First, conditions on the scaling rate of the threshold $T(n)$ are considered as the number of users n in the system asymptotically grows to infinity. It is shown that if the threshold scales as $o(\log n)$, that asymptotically there is no loss

in performances as compared to the case where every user feeds back their SINR information. Conversely, it is shown that if the threshold $T(n)$ scales as $\omega(\log n)$, that eventually no user in the system will feed back their SINR information, i.e. asymptotically no user in the system exceeds the threshold. In this case the transmitter has no CSI information and it is known that multiuser diversity cannot be achieved. The second part of the contribution considers explicit design of the thresholding function as a function of the number of users in the system when the number of users is finite. For a finite number of users, it is always possible that every user will be beneath any selected threshold and thus no one will feed back. Therefore, the design of the threshold will be with respect to achieving certain design criterion. In this work, three metrics are considered: limiting the probability that no user in the system feeds back for any transmit direction, limiting the average number of users feeding back and limiting the rate loss due to thresholding.

The organization of the contribution is as follows: the system model and random beamforming transmit scheme are discussed in Section II. Section III analyzes the asymptotic scaling rate of the thresholding function. Section IV addresses the design of the thresholding function in the case of a finite number of users. Finally, Section V summarizes the results of this contribution.

II. System Model and Random Beamforming

In this section the system model is described. Under the given system model, the random beamforming technique first proposed in [22] is outlined. The key implications of this work will be discussed as well as extensions found in [23].

A. System Model

A block fading channel model is assumed for the Gaussian broadcast channel. The transmitter has M transmit antennas and there are n receivers (users), each with N receive antennas. It is assumed that $n \gg M$ and $M \geq N$, which is typically the case in a cellular system. Let $\mathbf{s}(t) \in \mathbb{C}^{M\times 1}$ be the transmitted vector of symbols at time slot t and $\mathbf{y}_i(t) \in \mathbb{C}^{N\times 1}$ be the received symbols by the ith user at time slot t. The following model is used for the input-output relationship between the transmitter and the ith user:

$$\mathbf{y}_i(t) = \sqrt{\rho_i} H_i \mathbf{s}(t) + \mathbf{w}_i(t), \quad i = 1, \ldots, n. \tag{1}$$

$H_i \in \mathbb{C}^{N\times M}$ is the complex channel matrix which is assumed to be known at the receiver, $\mathbf{w}_i \in \mathbb{C}^{N\times 1}$ is the white additive noise, and the elements of H_i and $\mathbf{w}_i$ are i.i.d. complex Gaussians with zero mean and unit variance. This is the standard Rayleigh fading model, where the channel gain between every transmit-receive antenna pair is a complex circularly symmetric normal random variable. The transmit power is chosen to be M, i.e. $E\{\mathbf{s}^*\mathbf{s}\} = M$, so that the normalized power per antenna is one. The signal-to-noise ratio (SNR) at the receiver is $E\{\rho_i|H_i\mathbf{s}|^2\} = M\rho_i$ and ρ_i is the SNR of the ith user. It is assumed that the network is homogeneous, thus $\rho_i = \rho\ \forall i$.

B. Random Beamforming

The random beamforming transmit scheme proposed in [22] is the scheme analyzed in this contribution. The key elements of this work, as well as extensions found in [23], are now described.

The random beamforming transmit scheme involves generating an M dimensional random orthonormal basis from an isotropic distribution, with basis vectors $\phi_m \in \mathbb{C}^{M\times 1}$ for $m = 1, \ldots, M$. Let $s_m(t)$ be the mth transmit symbol at time t, then the total transmit signal at time slot t is given by

$$\mathbf{s}(t) = \sum_{m=1}^{M} \phi_m(t) s_m(t). \tag{2}$$

The received signal at the ith user is given by

$$\mathbf{y}_i(t) = \sum_{m=1}^{M} \sqrt{\rho} H_i(t)\phi_m(t)s_m(t) + \mathbf{w}_i(t). \tag{3}$$

In the original formulation in [22], each of the N receive antennas at each user measure the SINR for each of the M transmit directions and the maximum of the observed SINR values for each transmit direction is fed back. Looking at Figure 1, which shows the SINR values for each receive antenna/transmit direction pair, this scheme feeds back the maximum SINR in each column. The transmitter then schedules the users experiencing the largest SINR for each transmit direction, leading to maximal order statistics.

It is assumed that the ith user knows the quantity $H_i(t)\phi_m(t)$ from Equation 3 for each transmit direction. With this knowledge, the SINR observed at the jth receive antenna of the ith user for

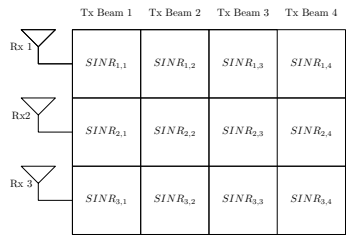

Fig. 1. SINR observations at a single user for $M = 4, N = 3$

the mth transmit direction is given by the following equation:

$$SINR_{i,j,m} = \frac{|H_{i,j}(t)\phi_m(t)|^2}{\frac{2}{\rho} + \sum_{k \neq m} |H_{i,j}(t)\phi_k(t)|^2}. \tag{4}$$

$H_{i,j}$ is the jth row of the ith user's channel matrix. The orthonormal beamforming vectors and i.i.d. zero mean and unit variance complex Gaussian elements of H_i imply the numerator in Equation 4 is distributed as a $\chi^2(2)$ random variable and the denominator is an independent $\chi^2(2(M-1))$ random variable.

When each receiver has a single receive antenna, i.e. $N = 1$, then under the assumed system model, it is shown in [22] that the distribution of the SINRs at the receive antennas is given by

$$f_{SINR}(x) = \frac{e^{-\frac{x}{\rho}}}{(1+x)^M} \left(\frac{1}{\rho}(1+x) + M - 1 \right) u(x), \tag{5}$$

where $u(x)$ is the unit step function. From now on, the $u(x)$ will be dropped from the SINR distribution and SINR density expressions with the understanding that all the SINR random variables of interest are nonnegative. Integrating the density in Equation 5 yields the distribution function for the SINR random variables:

$$F_{SINR}(x) = 1 - \frac{e^{-\frac{x}{\rho}}}{(1+x)^{M-1}}. \tag{6}$$

The main result of [22] is that asymptotically random beamforming has the optimal sum-rate throughput scaling. More specifically, the sum-rate throughput of the random beamforming technique R asymptotically scales at a rate of $M \log\log n$, which is the theoretical optimal scaling rate. This theorem is now stated:

Theorem 1: ([22]) Let M and ρ be fixed and $N = 1$. Then

$$\lim_{n \to \infty} \frac{R}{M \log\log n} = 1 \tag{7}$$

where R is the throughput of the random beamforming technique.

If there are $N \geq 1$ receive antennas at each user, the additional antennas can be utilized in different ways. In [22], each receive antenna at a given user is considered an independent user, in which case all the SINRs observed at each antenna are distributed according to the distribution function in Equation 6. An extension considered in [23] is to feedback only the largest SINR observed across all receive antennas and over all transmit directions, thus limiting the feedback to a single SINR value per user (and the index of the transmit direction that generated the largest SINR value) and reducing the total amount of feedback in the system. Once again referring to Figure 1, this scheme corresponds to feeding back the largest SINR value in the array. It is shown in [23] that the distribution of the SINR value fed back is given by $\left[F_{SINR_{(M)}}(x)\right]^N$ where $F_{SINR_{(M)}}(x)$ is given by

$$F_{SINR_{(M)}}(x) = 1 - \sum_{i=1}^{M} \frac{[d_i(x)]_+^M \exp\left(-\frac{xc}{d_i(x)}\right)}{A_i(x)} \tag{8}$$

and where $d_i(x) = \frac{2[1-x(M-i)]}{M-i+1}$, $A_i(x) = d_i(x)\prod_{j\neq i}^{M}(d_i(x) - d_j(x))$, and $[\cdot]_+$ is the positive part of the argument.

The extra receive antennas at each user can also be utilized by implementing a more complicated receive architecture such as an LMMSE receiver. With this architecture, the receiver will measure the SINR after LMMSE reception along each transmit direction and feed back the values. Under the given system model, it is shown in [23] that the distribution of SINR after LMMSE reception is given by

$$F_{MMSE}(x) = 1 - \frac{\exp\left(-\frac{x}{\rho}\right)}{(1+x)^{M-1}} \sum_{i=1}^{N} \frac{1 + \sum_{j=1}^{N-i} \begin{pmatrix} M-1 \\ j \end{pmatrix} x^j}{(i-1)!} \left(\frac{x}{\rho}\right)^{i-1}. \tag{9}$$

Using these distributions, it is shown in [23] that the results of Theorem 1 hold for these alternative feed back schemes and architectures. With the previous results now reviewed, the properties of the distribution functions can be used to ask questions about designing thresholds to limit which users feed back CSI and the effects these thresholds have on the system performance.

III. Thresholding Asymptotically in the Number of Users

Let $T(n)$ be the function that determines the threshold as a function of the number of users n in the system. For a fixed n, users experiencing a SINR less than $T(n)$ do not report back their

observed SINRs and users experiencing SINR values larger than $T(n)$ feed back their observed values. In this section, the question analyzed is how fast can the thresholding function $T(n)$ grow and still have the thresholded system achieve the same optimal asymptotic sum-rate scaling as the non-thresholded system? The theory of extreme value distributions and the convergence of types are essential to answering this question and will now be briefly reviewed.

A. Extreme Value Distributions

The convergence of types theorem is critical for the analysis of extreme value distributions and is now stated.

Theorem 2: (Convergence of Types)
Suppose $U(x)$ and $V(x)$ are two distribution function neither of which concentrates at a point. Suppose for $n \geq 1, F_n$ is a distribution, $a_n \geq 0, b_n \in \mathbb{R}, \alpha_n \geq 0, \beta_n \in \mathbb{R}$ and

$$F_n\left(a_n x + b_n\right) \to U(x)$$

$$F_n\left(\alpha_n x + \beta_n\right) \to V(x)$$

weakly. If

$$\tfrac{\alpha_n}{a_n} \to A > 0$$

$$\tfrac{\beta_n - b_n}{a_n} \to B \in \mathbb{R}$$

then

$$V(x) = U(Ax + B).$$

Let $X_1, \ldots, X_n$ be i.i.d. random variables with common distribution function $F(x)$ and let $M_n = \bigvee_{i=1}^{n} X_i$. The main theorem concerning extreme value distributions states that if a nondegenerate limiting distribution exists for M_n, then the limiting distribution converges to one of three classes of types:

Theorem 3: (Gnedenko, 1943)
Suppose there exists $a_n > 0, b_n \in \mathbb{R}, n \geq 1$ such that

$$P\left[(M_n - b_n)/a_n \leq x\right] = F^n(a_n x + b_n) \to G(x),$$

weakly as $n \to \infty$, where G is assumed nondegenerate. Then G is of the type of one of the following three classes:

$$\text{Type 1) } \Phi_\alpha(x) = \begin{cases} 0 & x < 0 \\ \exp\{-x^{-\alpha}\} & x \geq 0 \end{cases} \quad \text{for some } \alpha > 0.$$

$$\text{Type 2) } \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{\alpha}\} & x < 0 \\ 1 & x \geq 0 \end{cases} \quad \text{for some } \alpha > 0.$$

$$\text{Type 3) } \Lambda(x) = \exp\{-e^{-x}\} \quad x \in \mathbb{R}$$

The asymptotic scaling of the sum-rate throughput for the random beamforming scheme is established by showing the SINR values scheduled fall in the domain of attraction of the Type 3 extreme value distribution given in Theorem 3. Formally, $F_{SINR} \in D(\Lambda)$, i.e. the distribution function for the SINR, lies in the domain of attraction of Type 3. Once the extreme value distribution is established, the constants a_n and b_n are determined and used to show that the sum-rate throughput of the random beamforming scheme scaled optimally. The selection of the sequence a_n and b_n is not unique and the following lemma shows the relationship between any selection of sequences that satsify $F \in D(\Lambda)$.

Lemma 1: Suppose $F \in D(G)$, where G is one of the limiting distributions from Gnedenko's Theorem, and let $a_n > 0, b_n \in \mathbb{R}, \alpha_n > 0,$ and $\beta_n \in \mathbb{R}$ satisfy $F^n(a_n x + b_n) \to G(x)$ and $F^n(\alpha_n x + \beta_n) \to G(x)$. Then $a_n \sim \alpha_n$ (i.e. $\lim_{n\to\infty} a_n/\alpha_n = 1$) and $(\beta_n - b_n) \in o(a_n)$.
Proof: The existence of a_n, b_n, α_n and β_n is guaranteed by Gnedeno's Theorem since $F \in D(G)$. Also by Gnedenko's Theorem, G is the only possible limiting distribution that F approaches since $F \in D(G)$. Since $F^n(a_n x + b_n) \to G(x)$ and $F^n(\alpha_n x + \beta_n) \to G(x)$, by the Convergence of Types Theorem and letting G play the role of U and V yields the constants $A = 1$ and $B = 0$. Thus, also by the Convergence of Types Theorem, $\frac{\alpha_n}{a_n} \to 1$ and $\frac{\beta_n - b_n}{a_n} \to 0$, which gives the desired result. ■

B. Sufficient Conditions on the Scaling Rate of the Thresholding Function

The effects of thresholding on the distribution function F_{SINR} must be understood to design a threshold that achieves optimal asymptotic scaling of the sum-rate throughput. In [23], it is observed that that the threshold truncates the distribution function of the SINR values fed back

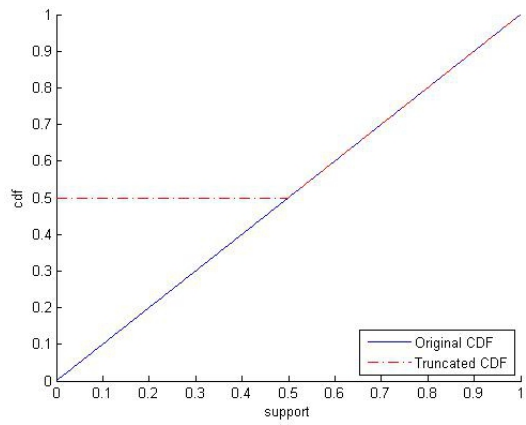

Fig. 2. Uniform Distribution Truncated at $x_{threshold} = 0.5$

to the transmitter in the following manner:

$$F_{SINR}^{threshold}(x) = \begin{cases} F_{SINR}(x) & x \geq x_{threshold} \\ F_{SINR}\left(x_{threshold}\right) & 0 \leq x < x_{threshold} \\ 0 & x < 0 \end{cases} \tag{10}$$

Equation 10 states that the probablity mass less than $x_{threshold}$, the threshold value, concentrates at zero, and this is the probability that the user does not feed back. The mass concentrates at zero because this is the lowest possible value that the nonnegative SINR random variables can attain. With this mapping, the values that do not exceed the threshold do not affect the scheduling decision since they are assumed to be the smallest possible value. To generalize the truncation from nonnegative random variables, let $x_1 = \inf\{x : F(x) > 0\}$ for an arbitrary distribution function F, i.e. the left end point of the support of F. Thresholding an arbitrary distribution function produces the following truncation:

$$F^{threshold}(x) = \begin{cases} F(x) & x \geq x_{threshold} \\ F\left(x_{threshold}\right) & x_1 \leq x < x_{threshold} \\ 0 & x < x_1 \end{cases} \tag{11}$$

Figure 2 shows the truncation of the uniform distribution at $x_{threshold} = 0.5$. The probability mass less than $x_{threshold} = 0.5$ maps to the left end point of the support, which in this case is $x_1 = 0$.

The effect of thresholding by truncation on a distribution function has been mathematically quantified. Time will now be spent to show that this mathematical definition corresponds with the goal of reducing feedback by implementation of a threshold. Let X be a random variable drawn from a distribution function $F \in D(\Lambda)$. Then the distribution function from Equation 11 corresponds to the following random variable Y:

$$Y = \begin{cases} X & X \geq x_{threshold} \\ x_1 & X < x_{threshold} \end{cases} \tag{12}$$

From an engineering perspective, this formulation is motivated by the following scenario. Let there be n users, where user i makes an observation X_i drawn i.i.d. from $F \in D(\Lambda)$. Each user decides to send back their observation X_i if it exceeds some threshold $x_{threshold}$ to a central unit, and the central unit has the task of finding the maximum observation. The users that did not report their observations because they did not exceed the threshold should not effect the decision of finding the maximum. Thus if a user did not report a value, their observation is assumed to be x_1. By taking on the lowest possible value, it cannot affect the task of finding the maximum value among the users that did feed back, and this motivates Equation 11.

Remark: If $x_1 = -\infty$, then $F^{threshold}(x)$ given by Equation 11 is not a proper distribution function because there is non-zero probability mass at $-\infty$. This will not pose a problem in the following analysis. Additionally, any point x_2 chosen such that $x_1 \leq x_2 < x_{threshold}$ could be selected since such a selection would not affect the task of finding the maximum among the fed back values, all of which exceed x_2.

In [23], it is shown that any finite threshold does not affect the asymptotic performance. The goal here is to design a threshold as a function of the number of users, $T(n)$, that grows as the number of users grows, and still has the desired asymptotic scaling. The following theorem is the main result of this section and provides conditions on the scaling rate of the threshold as a function of the number of users that guarantee optimal asymtotic scaling of the sum-rate throughput.

Theorem 4: Consider the system when $N = 1$, i.e. each user has a single receive antenna, under the model and random beamforming scheme described in Section II. If the thresholding function $T(n) \in o(\log n)$, then the sum-rate throughput of the system satisfies $\frac{R}{M \log\log n} \to 1$.

Before presenting the proof of Theorem 4, the following lemma will be required.

Lemma 2: Suppose $F \in D(\Lambda)$ with the sequence of coefficients $a_n > 0$ and $b_n \in \mathbb{R}$ satisfying $F^n(a_n x + b_n) \to G(x)$ and the left end point of the support is x_1. Let X_i be drawn i.i.d. from F and let $M_n = \bigvee_{i=1}^n X_i$. Also let $T(n)$ be the threshold as a function of n. Then $F^{threshold} \in D(\Lambda)$ if and only if $\frac{T(n)-b_n}{a_n} \to -\infty$.

Proof: $\Leftarrow$: Suppose $\frac{T(n)-b_n}{a_n} \to -\infty$. Let Y_i be drawn i.i.d. from $F^{threshold}$ and let $M_n^{th} = \bigvee_{i=1}^n Y_i$. Then

$$\begin{aligned} P\left[\frac{M_n^{th} - b_n}{a_n} \leq x\right] &= \left[F^{threshold}(a_n x + b_n)\right]^n \\ &= \begin{cases} F^n(a_n x + b_n) & a_n x + b_n \geq T(n) \\ F^n(T(n)) & x_1 \leq a_n x + b_n < T(n) \\ 0 & a_n x + b_n < x_1 \end{cases} \end{aligned}$$

As $n \to \infty$, by the hypothesis $\frac{T(n)-b_n}{a_n} \to -\infty$, so $a_n x + b_n \geq T(n)$ for all $x \in \mathbb{R}$, and thus $\left[F^{threhsold}(a_n x + b_n)\right]^n \to F^n(a_n x + b_n) \to G(x)$ as $n \to \infty$ since $F \in D(\Lambda)$. To show that this result holds regardless of which sequence of coefficients were chosen, suppose α_n and β_n are another sequence of coefficients satisfying $F^n(\alpha_n x + b_n) \to G(x)$. By Lemma 1, $\frac{\alpha_n}{a_n} \to 1$ and $\frac{\beta_n - b_n}{a_n} \to 0$. Substituting in the new coefficients yields

$$\begin{aligned} \frac{T(n) - \beta_n}{\alpha_n} &= \left(\frac{1/a_n}{1/a_n}\right) \cdot \left(\frac{T(n) - \beta_n + (b_n - b_n)}{\alpha_n}\right) \\ &= \left[\frac{1}{\frac{\alpha_n}{a_n}}\right] \cdot \left(\frac{T(n) - b_n}{a_n} - \frac{\beta_n - b_n}{a_n}\right) \\ &\to \frac{T(n) - b_n}{a_n} \to -\infty \end{aligned}$$

Thus any coefficients satisfying $F \in D(\Lambda)$ work.

$\Rightarrow$: Suppose $F^{threshold} \in D(\Lambda)$. Then there exists $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\left[F^{threshold}(a_n x + b_n)\right]^n \to G(x) = e^{-e^{-x}}$ for all $x \in \mathbb{R}$. As before

$$\left[F^{threshold}(a_n x + b_n)\right]^n = \begin{cases} F^n(a_n x + b_n) & a_n x + b_n \geq T(n) \\ F^n(T(n)) & x_1 \leq a_n x + b_n < T(n) \\ 0 & a_n x + b_n < x_1 \end{cases}$$

Let $\lim_{n\to\infty} \frac{T(n)-b_n}{a_n} = c \in [-\infty, \infty]$, then in the limit as $n \to \infty$, the above expression becomes

$$\left[F^{threshold}(a_n x + b_n)\right]^n \to \begin{cases} G(x) & x \geq c \\ G(c) & x < c \end{cases}$$

where the left end point of the support need not be of concern because the left end point of the support of $G(x)$ is $-\infty$. The only way the above expression can converge to $G(x)$ for all $x \in \mathbb{R}$ is if $c = -\infty$, and thus $\lim_{n\to\infty} \frac{T(n)-b_n}{a_n} \to -\infty$, the desired result. ■

Lemma 2 shows that if $T(n)$ scales sufficiently slowly, the convergence to the extreme value distribution $\Lambda(x)$ is unaffected. Having established Lemma 2, Theorem 4 can now be proven.

Proof: When $N = 1$, the distribution of the SINR is given by Equation 6 in Section II. It is also shown in [22] that $F_{SINR} \in D(\Lambda)$ and that the so called *auxiliary function* f satifies

$$f(t) = \frac{1 - F_{SINR}(t)}{F'_{SINR}(t)} \to \rho$$

where $F'_{SINR}(t)$ is the density given by Equation 5. Thus $1 - F_{SINR}(x) \sim \rho F'_{SINR}(x)$. Using techniques from [24] (section 1.5), suitable coefficients a_n and b_n can be found so that $F^n_{SINR}(a_n x + b_n) \to G(x)$. The first step is to solve for b_n by solving $\rho F'_{SINR}(b_n) = n^{-1}$. Plugging in the density and taking the $\log$ of both sides produces

$$\begin{aligned} & -\log e^{-b_n/\rho} - M\log(1+b_n) + \log(1+b_n+\rho(M-1)) = -\log n \\ \Rightarrow\ & -\frac{b_n}{\rho} - M\log(1+b_n) + \log(1+b_n+\rho(M-1)) = -\log n \\ \Rightarrow\ & \frac{b_n}{\rho} + M\log(1+b_n) - log(1+b_n+\rho(M-1)) = \log n. \end{aligned}$$

As $n \to \infty, b_n \to \infty$, so $\frac{1}{\rho}b_n$ dominates the above expression and $\frac{1}{\rho}b_n \sim \log n \Rightarrow b_n \sim \rho \log n$. It is known, e.g. [24], that a suitable choice of a_n for this case is $f(b_n)$, where f is the aforementioned auxiliary function, thus a suitable sequence is $a_n = \rho$ for all n. The b_n sequence can be written as $b_n = \rho \log n + r_n$ where r_n is the remainder and $r_n \in o(\log n)$ in this case. Plugging in this formulation of b_n yields

$$\begin{aligned} & \tfrac{\rho\log n + r_n}{\rho} + M\log(1+\rho\log n + r_n) - \log(1+\rho\log n + r_n + \rho(M-1)) = \log n \\ & \Rightarrow r_n + \rho M\log(1+\rho\log n + r_n) - \rho\log(1+\rho\log n + r_n + \rho(M-1)) = 0 \end{aligned}$$

and solving for r_n produces

$$\begin{aligned} r_n &= -\rho M\log\log n - \rho M\log\log\left(\rho + \frac{1+r_n}{\log n}\right) + \\ & \quad \rho\log\log n + \rho\log\log\left(\rho + \frac{1+r_n+\rho(M-1)}{\log n}\right) \\ &= -\rho(M-1)\log\log n + o(1). \end{aligned}$$

Plugging the remainder expression into b_n yields

$$b_n = \rho \log n - \rho(M-1)\log\log n + o(1)$$

which is essentially the same expression found in [22], where it is called u_n. This sequence proved to be the key fact used in determining the sum-rate scaling result in their work.

With a suitable choice of $\{a_n\}$ and $\{b_n\}$ established, Lemma 2 can be utilized to determine the scaling rate of the threshold so that the asymptotic sum-rate throughput scales optimally. Observe that

$$\begin{aligned}
\lim_{n\to\infty} \frac{h_n - b_n}{a_n} &= \lim_{n\to\infty} \frac{T(n) - (\rho\log n - \rho(M-1)\log\log n + o(1))}{\rho} \\
&= \lim_{n\to\infty} \left(\frac{T(n)}{\rho} - \frac{\rho\log n - \rho(M-1)\log\log n + o(1)}{\rho} \right) \\
&= \lim_{n\to\infty} \left(\frac{T(n)}{\rho} - (\log n - (M-1)\log\log n + o(1)) \right) \\
&\to -\infty.
\end{aligned}$$

Therefore $T(n) \in o(\log n)$ allows $\left[F_{SINR}^{threshold}(a_n x + b_n)\right]^n \to G(x)$ and all the conclusions derived in [22] using $F_{SINR} \in D(\Lambda)$ and $\{b_n\}$ hold. That is, selecting $T(n) \in o(\log n)$ still achieves the optimal sum-rate scaling rate. ■

C. Necessary Conditions on the Scaling Rate of the Thresholding Function

In the previous section, it is shown how the threshold must scale in order that F remains in the domain of attraction $D(\Lambda)$. This provides a sufficient condition for the scaling rate of the threshold such that all previously derived results based on the extreme value distribution hold. The goal of this section is finding a necessary condition for the scaling rate of the threshold $T(n)$ such that the asymptotic scaling rate of the sum-rate throughput of the random beamforming scheme is optimal. To state it another way, it may be possible for $T(n)$ to grow faster than the rate required to keep $F \in D(\Lambda)$ and still have the sum-rate throughput of the truncated scheme be optimal.

A different method of analysis will be used to address the necessary conditions. The method is based upon the work in [25]. The following definition is essential for the analysis that follows:

Definition 1: [25] Let a sequence of random variables $\{X_n, n \geq 1\}$ be unbounded above and let $M_n = \bigvee_{i=1}^{n} X_i$. The sequence $\{M_n\}$ is almost surely relatively stable (or stable for short) if there exist constants c_n such that $M_n/c_n \to 1$ almost surely as $n \to \infty$.

In this section it will be shown that if $X_1, \ldots, X_n$ are drawn i.i.d. from F_{SINR}, then the maximum of the SINRs, M_n, is stable, and to find the associated constants c_n. With the known sequence c_n, another sequence d_n can be chosen such that $\lim_{n\to\infty} \frac{c_n}{d_n} \to 0$, or $c_n \in o\left(d_n\right)$. If d_n is substituted for c_n, then $\lim_{n\to\infty} \frac{M_n}{d_n} \to 0$ almost surely. This shows that the maximum of the SINRs is growing slower than the sequence d_n. If the thresholding function $T(n)$ is selected such that $T(n) \sim d_n$, eventually no user in the system will exceed the threshold and thus no users feed back. If no user feeds back SINR information, there is no channel state information at the transmitter, and in this case it is well known that multiuser diversity cannot be achieved. Thus, a threshold $T(n) \sim d_n$ will cause the random beamforming scheduling algorithm to asymptotically fail.

Having discussed how stability plays into the analysis of the random beamforming scheduling algorithm, the following theorem states how fast is too fast for the thresholding function to scale.

Theorem 5: Let the SINR observed at each user be distributed i.i.d. according to F_{SINR}. Then if $T(n) \in \omega(\log n)$, asymptotically no user exceeds the threshold and consequently all multiuser diversity is lost.

Proof: It must now be shown that M_n when $X_i \sim F_{SINR}$ is in fact stable and find the appropriate constants for stability. The following theorem of Resnick's provide necessary and sufficient conditions for stability.

Theorem 6: [25] Let $\{X_n, n \geq 1\}$ be i.i.d. random variables with common distribution function F, where $F(x) < 1$ for all x. There exist suitable normalizing constants $c_n, n \geq 1$ such that $M_n/c_n \to 1$ almost surely as $n \to \infty$ if and only if

$$\int_1^\infty \frac{dF(x)}{1 - F((1-\epsilon)x)} < \infty$$

for all $1 > \epsilon > 0$.

The condition that $F(x) < 1$ for all x means that the random variable is unbounded from above, which is required in Definition 1. To verify Theorem 6 when $X_i \sim F_{SINR}$, substitute in

the distribution function and density from Equations 5 and 6 into Theorem 6:

$$\int_1^\infty \frac{\frac{e^{-x/\rho}}{(1+x)^M}\left(\frac{1}{\rho}(1+x)+M-1\right)}{1-\left(1-\frac{e^{-(1-\epsilon)x/\rho}}{(1+(1-\epsilon)x)^{M-1}}\right)}dx = \int_1^\infty \frac{\frac{e^{-\frac{x}{\rho}}}{(1+x)^M}\left(\frac{1}{\rho}(1+x)+M-1\right)}{\frac{e^{-(1-\epsilon)x/\rho}}{(1+(1-\epsilon)x)^{M-1}}}dx$$

$$= \int_1^\infty \frac{e^{-\frac{x}{\rho}\epsilon}}{(1+x)^M}(1+(1-\epsilon)x)^{M-1}\left(\frac{1}{\rho}(1+x)+M-1\right)dx$$

$$\leq \int_1^\infty \frac{e^{-\frac{x}{\rho}\epsilon}}{(1+x)^M}(1+x)^{M-1}\left(\frac{1}{\rho}(1+x)+M-1\right)dx \text{ since } 0<\epsilon<1$$

$$= \int_1^\infty \frac{e^{-\frac{x}{\rho}\epsilon}}{1+x}\left(\frac{1}{\rho}(1+x)+M-1\right) < \infty$$

Thus M_n, the maximum of n SINR observations, is stable. In the proof of Theorem 6 in [25], it is shown that the coefficients $c_n \sim \mu_n \equiv F^{-1}(1-n^{-1})$ are suitable for stability to hold. Since distribution functions need only be right continuous, the inverse is defined as $F^{-1}(x) = \inf\{y|F(y) \geq x\}$. In the SINR distribution of interest, the distribution function from Equation 6 is continuous. Similar to example (viii) in [25], the coefficients c_n will be found for this distribution. Define the function $R(x) = -\log(1-F(x))$ and note $\mu_n \equiv R^{-1}(\log n)$. Thus for F_{SINR}, $R(x)$ can be calculated:

$$\begin{aligned}
R(x) &= -\log(1-F_{SINR}(x)) \\
&= -\log\left(\frac{e^{-x/\rho}}{(1+x)^{M-1}}\right) \\
&= -\log(e^{-x/\rho}) + \log(1+x)^{M-1} \\
&= \frac{x}{\rho} + (M-1)\log(1+x) \\
&\sim \frac{x}{\rho} \text{ as } x \to \infty
\end{aligned}$$

and therefore $R^{-1}(x) \sim \rho x$ and a suitable sequence c_n is $R^{-1}(\log n) = \rho \log n$. Thus, if the threshold function $T(n)$ is selected such that $\lim_{n\to\infty} \frac{\rho \log n}{T(n)} \to 0$, or $T(n) \in \omega(\log n)$, then eventually no user in the system will feed back and all multiuser diversity (and thus optimal sum-rate scaling) is lost, proving the theorem. ■

Theorem 5, along with the conclusion from Section III-B provide an interesting result. Section III-B states that if $T(n) \in o(\log n)$, then optimal scaling still occurs, but the previous result states that if $T(n) \in \omega(\log n)$, that optimal scaling fails. Thus, any threshold $T(n) \in O(\log n)$ provides a kind of boundary between thresholds that allow optimal scaling and those that do not. When

$T(n) \in O(\log n)$, stability is achieved, as shown before, and $\lim_{n\to\infty} \frac{M_n}{\log n} \to 1$ almost surely. In this case, it is unknown how the asymptotic scaling is effected.

IV. Thresholds for a Finite Number of Users

The previous sections have established how fast thresholds can scale as a function of the number of users in the system so that asymptotically the sum-rate throughput still maintains the optimal scaling properties. While of theoretical interest, in any real system the number of users is finite. So how should the threshold be chosen in this case? The design of an appropriate threshold depends on the metrics that are to be optimized. This section discusses how to design the threshold for a system with n users to address the following metrics:

1) What should the threshold be so that the probability that no user feeds back information for any transmit beam is less than some design parameter γ_{outage}?
2) What should the threshold be so that on average only k users feed back SINR information for every transmit beam?
3) What should the threshold be so that the rate loss compared to the unthresholded system is less than some design parameter R_{loss}?

The first question is motivated by the fact that for any threshold, as long as one users feeds back SINR information for each transmit beam, then the maximum SINR is conveyed at the transmitter. In the event that the threshold is chosen so that no user feeds back SINR information for a transmit beam, multiuser diversity is lost on that beam, thus the probability that no user feeds back information needs to be controlled. The design parameter γ_{outage} is the "multiuser diversity outage probability", which is the maximum proportion of the time that the system is allowed to have no user feedback CSI information for any transmit direction. The second question will follow as a simple consequence of the analysis of the first question and constrains the average number of users feeding back. The final question is motivated by the fact that feedback is a precious resource, so perhaps one is willing to sacrifice some throughput to reduce the amount of feedback. As a function of the amount of throughput that is sacrificed for the reduction in feedback, an appropriate threshold can be designed. The following subsections provide solutions to the threshold design questions just raised.

A. Designing a Threshold under Outage Constraints

The first design criterion corresponds to each transmit beam having fed back SINR information $1-\gamma_{outage}$ percent of the time. The system is in multiuser diversity outage if at least one transmit beam has no feedback information, and this situation should occur less than γ_{outage} percent of the time. The key to the analysis of designing a threshold to meet this requirement is to recall from Equation 10 the effect thresholding has on the underlying distribution function. Two feedback schemes will be considered. The first feedback scheme has each user feed back their observed SINR value for each transmit direction. The second feedback scheme has each user feed back only the largest observed SINR value over all transmit directions, and the index of the transmit beam that generated the maximum.

1) SINR Information for each Transmit Direction: First, consider the feedback scheme where information is fed back for each transmit beam and $N = 1$, i.e. the number of receive antennas per user is one. Since the SINR is always positive, if $0 \leq SINR < x_{threshold}$, then the user does not feed back the SINR information for that particular transmit beam, and this happens with probability $F_{SINR}(x_{threshold})$. For a system with n users and M transmit beams, let $Y_{i,j}, i \in \{1, \ldots, n\}, j \in \{1, \ldots, M\}$ be a Bernoulli random variable indicating that user i feeds back SINR information for transmit beam j. From the previous statement, $Y_{i,j}$ are i.i.d. with the following distribution:

$$\Pr\left[Y_{i,j} = 1\right] = 1 - \Pr\left[Y_{i,j} = 0\right] = 1 - F_{SINR}\left(x_{threshold}\right). \tag{13}$$

Define the random variable $Z_j = Y_{1,j} + \cdots + Y_{n,j}$, which denotes the total number of users that feedback information for transmit beam j. Since the $Y_{i,j}$ are i.i.d. for all i, Z_j is a binomial random variable, i.e. $Z_j \sim B\left(n, 1 - F_{SINR}\left(x_{threshold}\right)\right)$. Outage occurs for transmit beam j when $Z_j = 0$, and $\Pr\left[Z_j = 0\right] = (1 - (1 - F_{SINR}(x_{threshold})))^n = (F_{SINR}(x_{threshold}))^n$ and the probability that the system is in outage is

$$\begin{aligned}
\Pr[\text{outage}] &= \gamma_{outage} = 1 - \Pr[\text{not in outage}] = 1 - \Pr\left[Z_1 \neq 0, \ldots, Z_N \neq 0\right] \\
&= 1 - \prod_{i=1}^{M} \Pr\left[Z_i \neq 0\right] = 1 - \prod_{i=1}^{M} \left(1 - \Pr\left[Z_i = 0\right]\right) \\
&= 1 - \left(1 - \left(F_{SINR}(x_{threshold})\right)^n\right)^M .
\end{aligned}$$

Thus, for a fixed number of users n, a fixed number of transmit antenna M, and design parameter

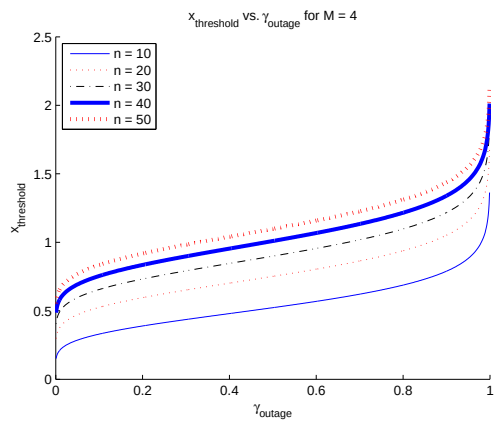

Fig. 3. Threshold as a function of γ_{outage} for $M = 4$ and $\rho = 1$

$\gamma_{outage} \in [0, 1]$, the threshold can be determined:

$$\gamma_{outage} = 1 - \left(1 - \left(F_{SINR}(x_{threshold})\right)^n\right)^M$$
$$\Rightarrow \quad x_{threshold} = F_{SINR}^{-1}\left(\left(1 - (1 - \gamma_{outage})^{\frac{1}{M}}\right)^{\frac{1}{n}}\right)$$

where F_{SINR}^{-1} is the inverse distribution function. The solution to the inverse distribution function can be expressed in terms of Lambert functions. An alternative solution to evaluating the inverse is to use an iterative algorithm. Because the distribution function is a function of a scalar and monotonic, iterative algorithms are well suited for the computation of the inverse. Figure 3 shows the value of $x_{threshold}$ as a function of γ_{outage} for varying numbers of users in the system. As expected, for γ_{outage} approaching zero, the threshold also has to go to zero, and for γ_{outage} approaching one, the threshold begins to blow up. For larger numbers of users in the system, the threshold can be set higher and achieve the same outage probability due to the increased maximum order statistics.

If the receivers are allowed to have $N \geq 1$ number of receive antenna, as mentioned in Section II-B, there are different ways to handle feedback. Originally in [22], one can consider each receive antenna as an independent user. In this case, the previous analysis follows verbatim except substituting n with nN since the system effectively has nN users under this methodology. Another approach considered in [23] is to have each user perform LMMSE reception, and then feedback the post-processed SINR for each transmit direction. In this case, the previous analysis

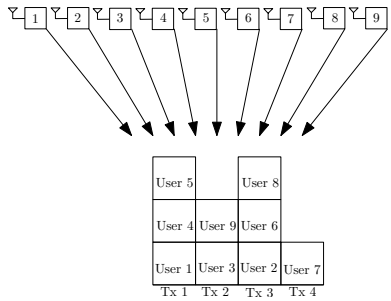

Fig. 4. Example of Maximization at the Transmitter under Alternate Feedback Scheme

is identical except substituting F_{SINR} with F_{MMSE} from Equation 9. In this case, the closed form solution of the inverse distribution function does not exist, and an iterative approach is required.

2) SINR Information for only the Best Transmit Direction: An alternative feedback scheme considered in [23] has each user only feed back the maximum SINR across all transmit beams, as well as the index for the transmit beam that generated the maximum. Recall from Figure 1, this corresponds to feeding back the largest value in the array. For the time being, let $N = 1$. The number of SINR values fed back for a particular transmit beam is a random variable under this scheme. To help understand this randomness, consider Figure 4 which shows an example for a system with nine users and four transmit beams. Each user feeds back the largest SINR observed across all transit beams, so in this example, users 1,4 and 5 report back their SINR values and the fact that transmit beam 1 generated the maximum. The number of users that feed back SINR values corresponding to transmit beam 1 is not known a priori and is thus random. Since each user is restricted to feed back only one SINR value, and due to the symmetry of the system model, the number of SINR values for each transmit beam received at the transmitter is multinomially distributed. Using the traditional balls and bins analogy for the multinomial, the n users are the n balls, the M transmit beams are the M bins, yielding a probability vector $p = \left[\frac{1}{M}, \ldots, \frac{1}{M}\right] \in \mathbb{R}^M$. The effect of thresholding on this problem is that rather than each user's feedback being in a bin associated with a transmit beam, that value may not actually be there if it did not exceed the threshold. Leveraging the balls and bins analogy again, each user can be thought of as putting a ball into a bin, but then there is a chance that the ball may be invisible and not counted due to the SINR not exceeding the threshold.

Recall the probability mass function of the multinomial of interest is

$$\Pr\left[Y_1 = y_1, \ldots, Y_M = y_m\right] = \begin{cases} \frac{n!}{y_1! \cdots y_M!} \left(\frac{1}{M}\right)^{y_1} \cdots \left(\frac{1}{M}\right)^{y_M} & \text{when } \sum_i^M Y_i = n \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where Y_i are the number of users who sent back information for the ith transmit beam. For any given realization of the multinomial, the probability of an outage event occuring is of interest. Let X_i, drawn from $F_{SINR_{(M)}}$ (as given by Equation 8), be the value that is to be fed back by the ith user. In this problem, if $X_i < x_{threshold}$, then the ball is considered invisible. An outage event occurs when all of the balls in at least one bin are invisible. Suppose, for concreteness, that four users are to feed back information regarding transmit beam one. Then there are four balls in bin one, and $Y_1 = 4$. The case where all of the balls are invisible corresponds to the error event, and this event occurs with probability $1 - \left(F_{SINR_{(M)}}(x_{threshold})\right)^4$, where the exponent of four came from the fact that $Y_1 = 4$. This argument can be extended to all the other bins. Thus, conditioned on knowing the number of balls in each bin yields the following probability:

$$\Pr\left[\text{not being in outage} | Y_1 = y_1, \ldots, Y_M = y_m\right] = \prod_{i=1}^{M} \left(1 - (F_{SINR_{(M)}}(x_{threshold}))^{y_i}\right). \tag{15}$$

Removing the conditioning produces the total probability of not being in outage

$$\begin{aligned} &\Pr \quad [\text{not being in outage}] \\ &= \sum \Pr\left[\text{not being in outage} | Y_1 = y_1, \ldots, Y_M = y_m\right] \cdot \Pr\left[Y_1 = y_1, \ldots, Y_M = y_m\right] \\ &= \sum \Pr\left[Y_1 = y_1, \ldots, Y_M = y_m\right] \cdot \prod_{i=1}^{M} \left(1 - (F_{SINR_{(M)}}(x_{threshold}))^{y_i}\right) \end{aligned}$$

where the summation is taken over all possible configurations of the multinomial. With the probability of not being in outage calculated, the threshold can be solved for via the following equation:

$$1 - \gamma_{outage} = \sum \Pr\left[Y_1 = y_1, \ldots, Y_M = y_m\right] \cdot \prod_{i=1}^{M} \left(1 - (F_{SINR_{(M)}}(x_{threshold}))^{y_i}\right). \tag{16}$$

The closed form inverse does not exist, so once again iteratve approaches must be taken. Equation 16 thus provides a way of finding the threshold under a multiuser diversity outage constraint when using a feedback algorithm that only feeds back local maxima. For the case where $N > 1$, the extension of the previous feedback scheme considered in [23] takes the maximum over transmit beams and antennas to reduce the feedback per use to a single SINR value (and the

beam index). In this case, the prior analysis follows substituting in $\left(F_{SINR_{(M)}}\right)^N$ for $F_{SINR_{(M)}}$, which arises due to the additional maximization over receive antennas.

B. Designing a Threshold Constraining the Average Number of Users Feeding Back

Because the SINR information is distributed across geographically separated receivers, it is impossible to design a threshold that can guarantee that exactly $k < n$ users feed back information for each transmit beam without further sharing of information. The best one can hope to do is use the statistical information about the channel metric, the observed SINR, to create a threshold that guarantees on average only k users feed back. First, consider the feedback schemes from Section IV-A.1 where each user feeds back SINR information for each transmit direction. Consider the case where $N = 1$. From Section IV-A.1 it is known that the probability of feeding back is a Bernoulli random variable, and the number of users feeding back for a particular transmit beam is a binomial random variable with parameters n and $1 - F_{SINR}\left(x_{threshold}\right)$. With these observations, the design of the threshold under the average number of users feeding back per transmit beam metric is straightforward. Once again, let $Z_j \sim B\left(n, 1 - F_{SINR}\left(x_{threshold}\right)\right)$ be the number of users feeding back for transmit beam j. Then

$$\begin{aligned} k &= E\left[\text{\# of users feeding back for transmit beam} j\right] \\ &= E[Z_j] = n\left(1 - F_{SINR}\left(x_{threshold}\right)\right) \\ &\Rightarrow x_{threshold} = F_{SINR}^{-1}\left(\frac{n-k}{n}\right). \end{aligned} \tag{17}$$

Figure 5 shows the threshold as a function of the constraint on the average number of users feeding back for varying number of users in the system. As the average number of users allowed is increased, the threshold decreases as expected to accomodate more users exceeding the threshold. When the average number of users equals the number of users in the system, the threshold is equal to zero.

If one is interested in constraining the average total amount of feedback to k (rather than on average k values per transmit beam), by the symmetry of the problem, substitute k with $k' = \frac{k}{M}$ in Equation 17. This is saying that each transmit direction gets on average $\frac{k}{M}$ values, so for all M transmit beams, the total on average is k. These thresholds extend to the case when $N > 1$ in two ways. If one considers feeding back the maximum observed SINR over the receive antennas for each transmit direction, then the previous analysis holds substituting

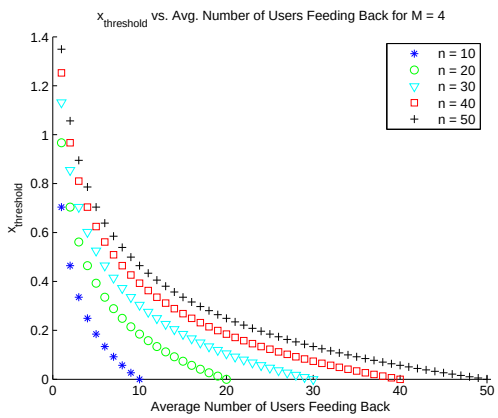

Fig. 5. Threshold as a function of the Average Number of Users Feeding Back for $M = 4$ and $\rho = 1$

$F_{SINR}(x)$ with $(F_{SINR}(x))^N$. If one allows for LMMSE reception, the previous results hold by using the distribution function given in Equation 9.

As in Section IV-A.2, consider the feedback scheme where only the local maximum SINR over the transmit beams is fed back. Recall that $F_{SINR_{(M)}}$ denotes the distribution function of the SINR values fed back under this scheme. The analysis in this case is similar to the previous case, but now one must consider that the transmit beam whose SINR value is being fed back is a random variable that is uniform over the transmit beams, i.e. every transmit direction is equally likely to produce the local maximum SINR. From Figure 1, this corresponds to the fact that when finding the maximum value in the array, it is equally likely to occur in any given column. Let Z'_j be the number of values fed back for transmit beam j. Then $Z'_j \sim B\left(n, \frac{1-F_{SINR}(x_{threshold})}{M}\right)$, where the factor of $\frac{1}{M}$ comes from the previously mentioned fact that associated transmit beam is a uniform random variable over the M transmit beams. The threshold design is given by

$$\begin{aligned} k &= E\left[\text{\# of users feeding back for transmit beam}\right] \\ &= E[Z'_j] = n\left(\frac{1 - F_{SINR_{(M)}}\left(x_{threshold}\right)}{M}\right) \\ \Rightarrow\ & x_{threshold} = F^{-1}_{SINR_{(M)}}\left(1 - \frac{kM}{n}\right). \end{aligned} \tag{18}$$

Notice that in this case it is required that $k \leq \frac{n}{M}$. This constraint makes sense under this feedback scheme since without thresholding, on average there will be $\frac{n}{M}$ values fed back per

transmit beam. These results can be extended to the case where $N > 1$ and the maximum SINR is taken over transmit directions and receive antennas in a straight forward manner. Substituting $\left(F_{SINR_{(M)}}\right)^N$ for $F_{SINR_{(M)}}$ produces $x_{threshold} = F^{-1}_{SINR_{(M)}}\left(1 - \frac{kM}{n}\right)^{\frac{1}{N}}$.

C. Designing a Threshold with Rate Loss Constraints

A third design constraint of interest is to select a threshold that bounds the rate lost in the system as compared to the system without a threshold. Let R be the throughput of the unthresholded system and consider the feedback scheme where $N = 1$ and the SINR is fed back for each transmit direction as in Section IV-A.1. From [23], it is known for this scheme that R can be expressed as

$$R \approx M \int_0^\infty \frac{1}{1+x}\left(1 - \left(F_{SINR}(x)\right)^n\right) dx, \tag{19}$$

where the approximation symbol becomes very accurate for a moderate number of users in the system. Equation 19 is equivalent to writing out the integral equation to evaluate $ME\left[\log\left(1 + \max\left\{X_1, \ldots, X_n\right\}\right)\right]$, where the X_i are drawn i.i.d. from F_{SINR}. The exponent of n on the distribution function is the effect of the maximum order statistics. The sum-rate throughput of the thresholded system is more complicated. The thresholded distribution given in Equation 10 cannot be directly plugged into the above Equation 19. This is because the thresholded distribution defined maps the rate to $x_l = 0$ when the SINR does not exceed the threshold. This mapping works when analyzing the question of optimal sum-rate scaling because the primary concern is whether multiuser diversity is achieved or not. With a finite number of users, even if no user exceeds the threshold and thus there is no feedback, this does not mean the system achieves zero throughput. If no user feeds back for a particular direction, the scheduler can randomly select a user and transmit to it. This random selection does not benefit from multiuser diversity since there is no CSI, but it can achieve a non-zero throughput.

The thresholded sum-rate throughput consists of two parts: the rate when a transmit direction receives CSI and the rate when no user feeds back information for a transmit direction and thus a random user is scheduled on that direction. Let $F_{threshold}$ be the distribution of the SINR for a transmit direction under the scheme where it schedules the maximum user if at least one user feeds back SINR information, and schedules a random user if no CSI is fed back. Thus, as with Equation 10, when at least one user feeds back CSI information for a transmit direction, the

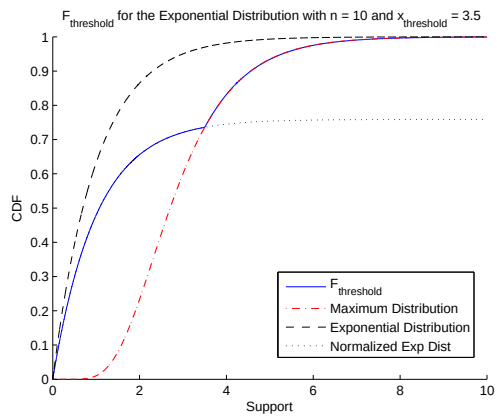

Fig. 6. Thresholded Distribution Function for the Exponential Distribution with $n = 10$ and $x_{threshold} = 3.5$

true maximum is observed and the thresholded distribution is given by

$$F_{threshold}(x) = F_{SINR}^{n} \text{ for } x \in [x_{threshold}, \infty). \tag{20}$$

Now consider the situation when no CSI is fed back. Conditioned on the fact that no SINR value exceeds the threshold,when the scheduling algorithm selects a user at random to transmit to, from [26] the conditional distribution of the SINR of the randomly selected user is given by

$$F_{random}(x) = \frac{F_{SINR}(x)}{F_{SINR}(x_{threshold})} \text{ for } x \in [0, x_{threshold}). \tag{21}$$

The distribution is defined on $[0, x_{threshold})$ since this distribution occurs when no users exceeds the threshold. The denominator in Equation 21 renormalizes the distribution F_{SINR} so that F_{random} is a true distribution function. The probability that no user exceeds the threshold is given once again by $F_{SINR}(x_{threshold})^n$. Putting together these facts with Equation 20 produces the total distribution function for this thresholded scheme:

$$F_{threshold} = \begin{cases} F_{SINR}(x)^n & \text{for } x \in [x_{threshold}, \infty) \\ F_{SINR}(x_{threshold})^n \cdot F_{random}(x) & \text{for } x \in [0, x_{threshold}) \end{cases} \tag{22}$$

Figure 6 verifies this equation with the exponential distribution for $n = 10$ and $x_{threshold}$ = 3.5. Substituting the distribution in Equation 22 for $F_{SINR}(x)^n$ in Equation 19 provides the sum-rate throughput of this scheme $R_{threshold}$:

$$R_{threshold} \approx M \int_0^{\infty} \frac{1}{1+x} \left(1 - F_{threshold}(x)\right) dx, \tag{23}$$

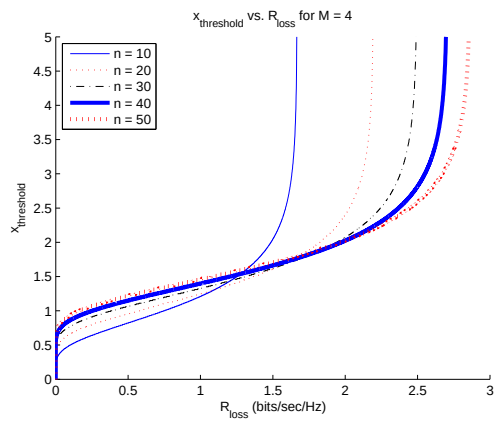

Fig. 7. Threshold as a function of R_{loss} for $M = 4$ and $\rho = 1$

where the integral can be multiplied by M since all the transmit beams statistics are independent and identically distributed.

Define the throughput of the system when random users are selected on each transmit beam as R_{random}. Then from Equation 19, R_{random} is given by

$$R_{random} \approx M \int_0^\infty \frac{1}{1+x} \left(1 - F_{SINR}(x)\right) dx. \tag{24}$$

Let $R_{loss} = R - R_{threshold} \in [0, R_{random}]$ be the design parameter that quantifies the amount of rate that can be sacrificed in order to reduce feedback. The upper bound on R_{loss} is R_{random} since in the worse case scenario this throughput can be achieved. Since F_{SINR} for this system is continuous and monotonically increasing, R_{loss} is monotonic increasing as a function of $x_{threshold}$ and there exists one solution for $x_{threshold}$ that solves the above equation for R_{loss}. R_{loss} is a complicated function of $x_{threshold}$, but due to its monotonicity, an iterative algorithm can be used to find $x_{threshold}$ to sufficient accuracy. Figure 7 shows the threshold as a function of the design parameter R_{loss}. For small values of R_{loss}, the larger the number of users, the higher the threshold can be set and still achieve the same rate loss, which is expected since for a larger number of users, the probability of exceeding the threshold increases. All the thresholds eventually approach infinity as R_{loss} approaches the unthresholded sum-rate throughput given by Equation 19.

Adapting the preceeding arguments for the feedback schemes considered in Section IV-A.1

when $N > 1$ is straightforward. For the case where each antenna is considered an individual user, change the exponent of F_{SINR} in Equation 22 from n to nN. When LMMSE receivers are implemented, substitute F_{MMSE} in Equations 21 and 22. The more difficult scheme to analyze is when only the local maximum SINR values and the associated beam index are fed back as in Section IV-A.2. Recall the multinomial distribution given by Equation 14. Define Y_i' to be the number of users that feed back CSI for the ith transmit beam and exceed the threshold. Thus, Y_i' is equal to Y_i minus the the number of values that would have been fed back but did not exceed the threshold. The difficulty in analyzing the rate of this scheme from a theoretical perspective is two-fold. First, $\sum Y_i' \leq n$ since some values may not exceed the threshold, and hence the Y_i' do not form a multinomial. This is a problem since the exponent of the distribution is no longer known. For example, consider a system with $n = 20$ users, and suppose $\sum Y_i' = 15$ and for concreteness $Y_1 = 5$. When theoretically analyzing the rate, the exponent for the first transmit beam could have been the maximum of five values that were fed back, or it could have been a maximum of up to 10 values ($Y_1 + n - \sum Y_i'$), since there is no information available on five of the users. The second problem is that if no information is available for a specific transmit direction, then a random user is scheduled. Suppose that the user randomly selected fed back CSI information, but for a different transmit direction. For example, suppose $N = 1$ and the SINR values at the randomly selected user are 1 and 2 for transmit directions 1 and 2 respectively. If the threshold was $x_{threshold} = 0.5$, both exceed the threshold, but since only the local maximum is fed back, the value of 2 on the first transmit direction is never observed. Thus, when this user is randomly selected for the first transmit direction, the SINR value is greater than the threshold, but is conditionally known to be less than 3 since the that value is known at the scheduler. Therefore the distribution of the randomly selected users' SINRs is quite complicated. For these reasons, the analysis will not be performed on this scheme. Simulations can be used to design an appropriate threshold.

V. Conclusion

This paper addresses the question of which user should feed back the observed SINR given statistical knowledge of the channel metric. When every receiver feeds back their SINR information, it is shown in [22] that random beamforming achieves the optimal sum-rate throughput scaling. In [23], it is shown that any finite threshold such that users experiencing SINRs below

this threshold do not feed back does not affect this optimal scaling rate. The questions raised in this contribution discuss how the thresholds can be designed as a function of the number of users in the system to achieve certain performance metrics.

The first performance metric considered is how fast can the threshold grow yet still exhibit the optimal scaling. It is shown that any thresholding function $T(n)$ as a function of the number of users n that grows slower than $\log n$, i.e. $T(n) \in o(\log n)$, achieves the optimal scaling rate. Conversely, any thresholding function $T(n) \in \omega(\log n)$ causes the system to lose all multiuser diversity.

The second part of the contribution considers how to optimally design thresholds for systems with a finite number of users. Using the statistics of the channel metric, thresholds for an n user system under three different design criterion are found. Under the first criterion, a threshold is designed to limit the probability that no user in the system feeds back CSI information for any transmit beam to be less than the design parameter γ_{outage}. The second criterion considers choosing a threshold such that on average only k out of n users in the system feed back CSI information for each transmit beam. Finally, a threshold is designed to limit the rate lost compared to a random scheduling algorithm to be below a design parameter R_{loss}.

REFERENCES

[1] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *Information Theory, IEEE Transactions on*, vol. 49, no. 7, pp. 1691 – 1706, 2003.

[2] P. Viswanath and D.N.C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *Information Theory, IEEE Transactions on*, vol. 49, no. 8, pp. 1912 – 1921, 2003.

[3] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2658 – 2668, 2003.

[4] Wei Yu and J.M. Cioffi, "Sum capacity of gaussian vector broadcast channels," *Information Theory, IEEE Transactions on*, vol. 50, no. 9, pp. 1875 – 1892, 2004.

[5] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 3936 –3964, 2006.

[6] Y. Isukapalli and B.D. Rao, "Finite rate feedback for spatially and temporally correlated miso channels in the presence of estimation errors and feedback delay," in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, 2007, pp. 2791 –2795.

[7] Y. Isukapalli, R. Annavajjala, and B. Rao, "Performance analysis of transmit beamforming for miso systems with imperfect feedback," *Communications, IEEE Transactions on*, vol. 57, no. 1, pp. 222 –231, 2009.

[8] Y. Isukapalli and B.D. Rao, "Packet error probability of a transmit beamforming system with imperfect feedback," *Signal Processing, IEEE Transactions on*, vol. 58, no. 4, pp. 2298 –2314, 2010.

[9] Yao Ma and Dongbo Zhang, "Error rate of transmit beamforming with delayed and limited feedback," in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, 2007, pp. 4071 –4075.

[10] Yao Ma, A. Leith, and R. Schober, "Predictive feedback for transmit beamforming with delayed feedback and channel estimation errors," in *Communications, 2008. ICC '08. IEEE International Conference on*, May 2008, pp. 4678 –4682.

[11] N. Jindal, "Mimo broadcast channels with finite-rate feedback," *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5045 –5060, nov. 2006.

[12] D.J. Love, Jr. Heath, R.W., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2735 – 2747, 2003.

[13] Taesang Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 3, pp. 528 – 541, 2006.

[14] A. Bayesteh and A.K. Khandani, "How much feedback is required in mimo broadcast channels?," in *Information Theory, 2006 IEEE International Symposium on*, 2006, pp. 1310 –1314.

[15] O. Edfors F. Floren and B. Molin, "The effect of feedback quantization on the throughput of a multiuser diversity scheme," in *Proc. IEEE Glob. Telecom. Conf.*, Dec. 2003, vol. 1, pp. 497–501.

[16] D. Gesbert and M. Alouini, "Selective multi-user diversity," in *Proc. IEEE Int. Symp. Signal Proc. Info. Theory*, Dec. 2003, pp. 162–165.

[17] D. Gesbert and M. Alouini, "How much feedback is multi-user diversity really worth?," in *Proc. IEEE Int. Conf. on Commun.*, June 2004, vol. 1, pp. 234–238.

[18] M. Alouini V. Hassel, D. Gesbert and G.E. Oien, "A threshold-based channel state feedback algorithm for modern cellular systems," *IEEE Trans. Wireless Comm.*, vol. 6, no. 7, pp. 2422–2426, July 2007.

[19] S. Sanayei and A. Nosratinia, "Opportunistic downlink transmission with limited feedback," *IEEE Trans. Info. Th.*, vol. 53, no. 11, pp. 4363–4372, November 2007.

[20] A. Tewfik Y. Al-Harthi and M. Alouini, "Multiuser diversity for wireless communications," in *Proc. IEEE Glob. Telecom. Conf.*

[21] V.K.N. Lau D. Gesbert B.D. Rao D.J. Love, R.W. Heath Jr. and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Jour. on Sel. Areas in Comm.*, vol. 26, no. 8, Oct. 2008.

[22] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[23] M. Pugh and B. D. Rao, "Reduced feedback schemes using random beamforming in MIMO broadcast channels," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1821–1832, Mar. 2010.

[24] Sidney I. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Springer-Verlag, 1987.

[25] Sidney I. Resnick and R. J. Tomkins, "Almost sure stability of maxima," *Journal of Applied Probability*, vol. 10, no. 2, pp. 387–401, 1973.

[26] H. A. David, *Order Statistics*, New York: Wiley, 1970.