Throughput Maximization with Multiple Codes and Partial Outages

Syed Ali Jafar, Sriram Vishwanath, Andrea Goldsmith
Stanford University, USA

***Abstract*— We provide an information theoretic perspective on the problem of throughput maximization in a block flat fading wireless data system with codeword lengths restricted to be less than the fade block duration. We assume no channel state information at the transmitter (CSIT) and perfect channel state information at the receiver (CSIR). We explore the tradeoffs between using a single codebook vs multiple codebooks (rate-splitting) on Single Input Single Output (SISO) channels, and scalar coding vs vector coding for diagonal Multiple Input Multiple Output (MIMO) channels. For all log-concave scalar channel fade distributions, we show that using multiple codebooks increases the average throughput of the system when the multiple codewords are transmitted simultaneously in time, frequency and space over the same channel. Splitting the channel orthogonally in time, frequency, or among the inputs of a MIMO system and then transmitting different codewords on each orthogonal sub-channel significantly reduces the achievable average throughput.**

I. INTRODUCTION

We consider the problem of average throughput maximization for wireless data communications. For a typical data communication system transmitting packets at rate R with retransmissions and at a packet error rate P_e, the average throughput defined as $\overline{T} = R(1 - P_e)$ achievable while fulfilling a certain delay requirement is a good performance measure. [1] and [2] use this metric to optimize the link adaptation thresholds for EDGE. These works, and the references therein, deal with the throughput maximization problem from a very practical perspective. From a fundamental information theoretic perspective, not much is known about the maximum achievable average throughput with packet errors and retransmissions. This is in part due to the fact that the usual notion of capacity corresponds to a zero error capacity and entails unbounded decoding delay. A more relevant and a relatively new notion to estimate average throughput is outage capacity [3]. Outage probability of a channel is known to be a good approximation to the probability of codeword error for practical coding schemes. In this work we use the notion of outage capacity to provide an information theoretic perspective on the problem of average throughput maximization with block fading.

As explained in greater detail in Section II, in the absence of CSIT, if the transmitter uses a single codebook at a rate R with codeword duration N smaller than the fade block duration, the average throughput can be expressed as $\overline{T} = R$ Prob(channel capacity $> R$). Choosing a higher rate codebook would decrease the probability that the channel block supports that rate. Thus the rate R needs to be optimized to maximize average throughput. However, if the transmitter splits the flat fading channel block orthogonally in time or frequency or in some cases among the inputs of a multiple input channel, then it can use multiple codebooks, possibly at different rates, over these subchannels. multiple codebooks can also be transmitted simultaneously in time, frequency and space so that the codewords interfere with each other, with successive decoding at the receiver.

The use of multiple codebooks leads to the notion of partial outages as the channel realization decides which codewords can be decoded and which cannot. It also leads to several interesting questions. For example, between single and multiple codebooks which one is better to maximize average throughput? Which one allows easier coding and decoding? And what are the other advantages and disadvantages that one scheme has over the other? These are the questions that we answer in this paper. We find that the answers depend on the channel distribution in general. However, by using the notion of log-concavity of distributions we are able to provide results applicable to a broad range of channel fade distributions commonly encountered in wireless communications.

II. SYSTEM MODEL

We assume a block flat-fading channel model with the input output relationship for the k^{th} symbol in the b^{th} fading block given by

$$y_b(k) = H_b x_b(k) + n_b(k), \tag{1}$$

where y, H, x, and n are the output symbol, channel gain, input symbol and zero mean additive white Gaussian noise, respectively. H_b is fixed over a block, and the sequence of channel realizations $H_1, H_2, \cdots$ is i.i.d. The receiver knows the channel H_b perfectly. The transmitter knows only the statistics of the channel and therefore cannot adapt its transmit rate or power to the channel realization from one block to another. For simplicity, we drop the block and symbol indices. We define

$$h \triangleq ||H||^2, \qquad C_h(P) \triangleq \log(1 + hP) \tag{2}$$

and $f_h(x)$ as the distribution of the random variable h. We normalize h and n so that E[h]=1 and E[$|n|^2$]=1. Note that $C_h(P)$ is the capacity of the constant channel with transmit power P.

The constraints on the decoding delay and complexity at the receiver translate to a constraint on the maximum codeword length N, so that a codeword does not span more than one block. Thus we cannot code across blocks and the ergodic capacity with no CSIT and perfect CSIR is not a relevant measure of the average throughput. Even the delay limited capacity, where one aims to keep the instantaneous mutual information constant at all times, does not apply to this system because of the absence of CSIT. The relevant quantity in our case is the outage probability.

In the absence of CSIT, if the transmitter transmits a code at a rate R, the decoded codeword at the receiver will be in error with probability close to one if the capacity corresponding to the actual channel realization in that block is smaller than R (strong converse to channel coding theorem). We call this an outage. On the other hand, if the channel capacity for the actual channel realization is larger than R the probability of error is negligible. Note that the actual probability of error in this case is bounded away from zero since the code duration is limited to less than a block. However, we make the assumption that the fade block duration is long enough so that the probability of error can be made negligibly small compared to one. This assumption is reasonable because, unlike the time-varying channel where the coding delay is needed to average out over the channel variations which may happen on a much larger time scale, in the time-invariant Gaussian channel, the coding delay is only needed to average out the Gaussian noise to get small error probabilities, and is typically quite short.

Under our assumptions, the maximum achievable average throughput of the system when the transmitter uses one codebook at a rate R can be represented as

$$\overline{T}^{(1)}_{max}(P_t) = \max_{R} \; R \,\text{Prob}(C_h(P_t) > R), \tag{3}$$

where h is the constant channel fade experienced by a codeword, and C_h is the Shannon capacity for the constant channel h with transmit power P_t. Note that since h is drawn randomly according to the distribution $f_h(x)$, C_h is also a random variable, and hence the probability. The superscript (j) in $\overline{T}^{(j)}_{max}$ is used to indicate that j different codebooks are being used. Its significance will be obvious later when we discuss multiple codewords and partial outages. Note that since the receiver knows the channel perfectly, it knows when an outage happens and therefore does not attempt to decode the codeword when the capacity supported by the channel is smaller than the code rate. When this happens, we assume that as in a typical packet transmission system there exists a feedback channel through which the receiver requests the transmitter to retransmit that packet (codeword). Note that even if a fade block is in outage, i.e. $C_h < R$, the receiver can save the received information in that block and possibly use it to aid in the decoding process when the corresponding codewords are retransmitted. However, for simplicity and to model practical systems better, we assume that the receiver simply discards the information received in a fade block during an outage.

Our goal is to achieve the highest possible average throughput over all transmit strategies. Equation (3) represents the average throughput corresponding to just one possible transmit strategy. However it is not the only possible transmit strategy. For instance, what if the transmitter used more than one codeword? The transmitter could perform time-sharing between two codewords, at possibly different rates R_1 and R_2 and different powers P_1 and P_2 while maintaining the same average transmit power. Or it could divide the frequency band into two parts and transmit codewords at different rates and powers on these two sub-channels. For a MIMO system, it could transmit different codewords on different channel inputs. Or it could simply transmit both codewords simultaneously over time, frequency and channel inputs, so that the receiver can recover the data through successive decoding. Thus, the transmitter can transmit multiple codewords at different rates. It is possible then that in a block one of the codewords is in outage while the other is not. We refer to this as a *partial outage*. Thus, while transmitting only one codeword allows the receiver to recover either the full transmit rate or nothing over a block, partial outage allows the receiver to recover a fraction of the total transmit rate.

Our main results contained in Sections IV and V use some notation and results from reliability theory. In particular we use the notions of the reliability function, the failure rate function and log-concavity of the density function $f_h(x)$. The next section presents a summary of these topics.

III. Log-concavity and the Failure Rate

One of the main objectives in this paper is to find out if using multiple codebooks increases the achievable average throughput. However, we find that the answer depends on the channel fade distribution. For example, with time sharing between codebooks, we find in Section V-A that for some channel distributions, multiple codebooks strictly decrease the average throughput, while for some other channel distributions multiple codebooks lead to higher average throughput than possible with a single codebook. This makes it impossible to come up with general results that apply to any arbitrary fade distribution. Fortunately, we also find that most commonly encountered distributions lead to the same results. The common property of these distributions that makes this happen turns out to be their log-concavity. This allows us to obtain fairly general results that hold for all log-concave distributions without having to consider each specific distribution separately.

The distribution $f_h(x)$ is said to be log-concave if $\log f_h(x)$ is a concave function of x. The assumption of log-concavity is much used since most commonly encountered distributions are log-concave. For example, the Nakagami-m fading distribution[4], which can model Rayleigh and Ricean distributions as well as more general ones, is log-concave. A list of distributions that always have log-concave density functions is provided in [5].

The mathematical implication of log-concavity that is the key to our results is presented in Theorem 1. To state the theorem we introduce some nomenclature. We define the reliability function $\overline{F}_h(x)$ and the failure rate $r_h(x)$ as

$$\overline{F}_h(x) \triangleq \text{Prob}(h > x) \;\text{ and }\; r_h(x) \triangleq \frac{f_h(x)}{\overline{F}_h(x)}. \tag{4}$$

Note that the reliability function $\overline{F}_h(x)$ is the probability that a random channel realization will support a given rate $R = \log(1+Px)$. The probability that the channel *fails* to support a rate $R + \delta R = \log(1+(x+\delta x)P)$, given that it supports a rate $R = \log(1+Px)$, is given by $r_h(x)\delta x$. Hence the name *failure rate*.

The following theorem relates the log-concavity of $f_h(x)$ to the failure rate.

Theorem 1: If the density function $f_h(x)$ is log-concave on (a, b), then the failure rate is non-decreasing on (a, b).

Proof: This is Corollary 2 to Theorem 2 in [5]. ■

Theorem 1 is central to the proofs of the results in the next two sections.

IV. Characterizing $\overline{T}^{(1)}_{max}(P)$

We begin this section with an alternate representation for $\overline{T}^{(1)}_{max}(P)$ as

$$\overline{T}^{(1)}_{max}(P) = \max_{a_1} \; \log(1+a_1P)\overline{F}_h(a_1) \tag{5}$$

This is easy to see as the rate $R = \log(1+a_1P)$ is achievable whenever $C_h(P) = \log(1+hP) > R$, i.e. $h > a_1$. Further, let us define the function $A(P)$ as the optimal channel threshold that achieves $\overline{T}^{(1)}_{max}(P)$. Mathematically,

$$A(P) \triangleq \arg\max_{a} \log(1+aP)\overline{F}_h(a) \tag{6}$$

Theorem 2 shows that if the channel fade distribution is log-concave, then there is only one unique channel threshold a that maximizes average throughput for a transmit power P. Thus $A(P)$ is unambiguously defined by (6).

To build some intuition, we need to characterize the behavior of the maximum average throughput $\overline{T}^{(1)}_{max}(P)$, the optimal channel threshold $A(P)$ and the optimal code rate $\log(1+PA(P))$ as functions of transmit power P. For example, note that it is not obvious if the optimal rate should increase or decrease with P or if the maximum average throughput is a concave function of P or not. In fact, while Theorem 2 shows that the optimal rate is an increasing function of P and $\overline{T}^{(1)}_{max}(P)$ is an increasing concave function of P for log-concave fade distributions, one can indeed find distributions for which neither of these statements is true.

The following theorem characterizes $\overline{T}^{(1)}_{max}(P)$ for log-concave fade distributions.

Theorem 2: For all log-concave channel fade distributions $f_h(x)$ such that $f_h(x) > 0,\ \forall x \in (0,\infty)$, the following statements are true:

2.1: The optimal channel threshold $A(P)$ is unique. i.e., If $\overline{T}^{(1)}_{max}(P) = \log(1+a_1P)\overline{F}_h(a_1) = \log(1+a_2P)\overline{F}_h(a_2)$, then $a_1 = a_2 = A(P)$.

2.2: The optimal channel threshold $A(P)$ is a strictly decreasing function of P.

2.3: The optimal outage probability, 1- $\overline{F}_h(A(P))$ is a strictly decreasing function of P.

2.4: The optimal rate $\log(1+PA(P))$ is a strictly increasing function of P.

2.5: $A(P)r_h(A(P)) \leq 1$ and $r_h(A(P))P\frac{\partial}{\partial P}A(P) + 1 > 1$ for all $P \in (0,\infty)$. This is needed to prove the next statement.

2.6: Most importantly, $\overline{T}^{(1)}_{max}(P)$ is an increasing concave function of P.

Proof: See [6]. ■

V. Average Throughput Maximization

A. Time or Frequency Sharing

Our main result in this section is contained in the following theorem.

Theorem 3: Time or frequency sharing between multiple codebooks does not increase the average throughput if and only if $\overline{T}^{(1)}_{max}(P)$ is a concave function of P for the channel fade distribution.

Proof: See [6]. ■

B. Simultaneous Multiple Codes

Consider a system where the transmitter transmits n codewords simultaneously in time and frequency. Such a system is characterized by the code rates for each codebook, the transmit power in each codeword, the decoding order at the receiver, and channel thresholds corresponding to partial outages. For simplicity we first consider two simultaneous codewords at rates R_1 and R_2 and with powers P_1 and P_2. Also suppose that codeword 1 is decoded before codeword 2, i.e. codeword 2 is treated as noise while codeword 1 is being decoded. The channel thresholds a_1 and a_2 ($a_1 \leq a_2$) are defined such that

$$R_1 = \log\left(1+\frac{P_1a_1}{1+P_2a_1}\right) \quad \text{and} \quad R_2 = \log(1+P_2a_2).$$

Thus when $h \in (0,a_1)$ the receiver cannot decode either codeword, when $h \in [a_1,a_2)$ the receiver can only decode codeword 1 and therefore can only recover a rate R_1 (partial outage), and when $h \in [a_2,\infty)$ the receiver can decode both codewords to recover rate R_1+R_2.

The maximum average throughput with two simultaneous codewords can therefore be expressed as

$$\overline{T}^{(2)}_{max}(P) = \max_{P_1,P_2,a_1,a_2 \,|\, P_1+P_2=P_t} \log(1+\frac{a_1P_1}{1+a_1P_2})\overline{F}_h(a_1) + \log(1+a_2P_2)\overline{F}_h(a_2) \quad (7)$$

Similarly, with n simultaneous codewords, let the rates be given by the vector $\vec{R} = \{R_1,R_2,\cdots,R_n\}$, powers given by the vector $\vec{P} = \{P_1,P_2,\cdots,P_n\}$, and the channel thresholds given by the vector $\vec{a} = \{a_1,a_2,\cdots,a_n\}$ such that $a_1 \leq a_2 \leq \cdots \leq a_n$. So the codewords are decoded at the receiver in ascending order, i.e. codeword 1 is decoded first and codeword n decoded last. Then we have the following relationships:

$$\begin{aligned} R_i &= \log\left(1+\frac{a_iP_i}{1+a_i\sum_{j=i+1}^{n}P_j}\right), \quad i=1,2,\cdots,n-1 \\ R_n &= \log(1+a_nP_n) \end{aligned} \quad (8)$$

The receiver recovers a rate $R_1+R_2+\cdots+R_i$ if the channel $h \in [a_i, a_{i+1})$. Define $a_{n+1} = +\infty$. The maximum average throughput with n simultaneous codewords can therefore be expressed as

$$\overline{T}^{(n)}_{max}(P) = \log(1+a_nP_n)\overline{F}_h(a_n) + \max_{\vec{P},\vec{a} \,|\, \sum_{j=1}^{n}P_j=P_t} \sum_{i=1}^{n-1} \log\left(1+\frac{a_iP_i}{1+a_i\sum_{j=i+1}^{n}P_j}\right)\overline{F}_h(a_i)$$

Next we present the main result of this section.

Theorem 4: The maximum average throughput achievable by transmitting n codes simultaneously, $\overline{T}^{(n)}_{max}(P)$ is a strictly increasing function of the number of codes n for any log-concave channel fade distribution.

Proof: See [6]. ■

In summary, the results of this section and the previous section imply that for log-concave densities, using different rate codebooks on different orthogonal (non-interfering) channels obtained by splitting the flat block fading channel into non-overlapping multiple channels in time or frequency does not improve the average throughput. However, using multiple codebooks simultaneously over the same time and frequency yields maximum achievable average throughputs that are strictly increasing in the number of codebooks.

We considered only SISO channels so far. The next two sections show how the results extend to Single Input Multiple Output (SIMO), Multiple Input Single Output (MISO), and diagonal MIMO systems.

C. Extension to SIMO and MISO systems

First, consider a SIMO system. So the channel H, the output y and the noise n in equation (1) are vectors. The input x is still a scalar. However note that nothing else changes. The definitions of h, $C_h(P)$ and $\overline{T}^{(n)}_{max}(P)$ are still the same. Thus the results of the previous sections hold for SIMO systems with no change.

Next, we consider MISO systems. So the channel H and the input x are vectors while the output y and the noise n are still scalars. We assume that each component of H is i.i.d. Now, since there are multiple transmit antennas available to the transmitter, it can map multiple codebooks to multiple antennas in different ways. The maximum achievable average throughput with multiple codebooks and optimal mapping of codebooks to antennas are interesting questions that remain unsolved. However, if we consider simultaneous transmission of codebooks over space, time and frequency, i.e. if the input covariance matrix of each codebook is a multiple of an $n_T \times n_T$ identity matrix - a reasonable assumption in the absence of CSIT - then once again the derivation of previous sections holds with a minor change in the definition of $C_h(P)$ to $C_h^{MISO} \triangleq \log\left(1 + h\frac{P}{n_T}\right)$, where n_T is the number of inputs in the MISO system. Thus the results of the previous sections apply to MISO systems where all codebooks are used simultaneously over all inputs if we divide all codebook powers by n_T.

D. Diagonal MIMO systems

For a general $M \times M$ MIMO channel H with i.i.d. elements, and no CSIT, it is reasonable to assume that the input covariance matrix is a multiple of the $M \times M$ identity matrix. With such an input covariance matrix, the maximum rate supported by the channel over a fade block is

$$C_H(P) = \log\left|I + \frac{P}{M}HH^\dagger\right| \tag{9}$$

Note that the diagonal covariance matrix alone does not uniquely determine the coding strategy. The same capacity can be achieved by scalar coding or vector coding. With scalar coding M scalar codes are transmitted simultaneously on the M inputs. By scalar code we mean that the code symbol at each instant is a scalar. Vector coding on the other hand uses just one codebook and the code symbol at each instant is an M dimensional vector. The M components of this symbol are transmitted on different inputs. While the Shannon capacity with either of these schemes and an input covariance matrix that is a multiple of the $M \times M$ identity matrix is the same, the schemes are not equivalent in terms of maximum achievable average throughput. As before, while transmitting multiple scalar codes allows partial outages, transmitting one vector code only allows the receiver to either recover the full transmitted rate or nothing. A comparison of vector coding vs scalar coding for maximizing average throughput with no CSIT is therefore an interesting problem, but hard to solve in the general case.

We compare scalar coding vs vector coding for the simpler case of diagonal MIMO channels. So H is a diagonal matrix over each fading block with i.i.d. diagonal elements drawn according to the distribution $f_h(x)$. Notationally H=diag$\{H_{11}, H_{22}, \cdots, H_{MM}\}$ and $h_i \triangleq |H_{ii}|^2$. Note that the channels do not interfere with each other. This model corresponds to the situation when the transmitter uses multiple flat block fading frequency slots that are separated far enough so that the fade level in each is independent of the others.

To transmit at a total rate R, with scalar coding, and for code duration N, the transmitter has M codebooks, each with $2^{\frac{NR}{M}}$ codewords with power $\frac{P}{M}$. M codewords are transmitted over every N symbol durations, such that each codeword sees one fixed scalar channel. The receiver decodes each of these separately as they are sent on non-interfering channels. The fade value on each channel decides whether the corresponding codeword can be decoded. Note that if one of the channels is bad the corresponding codeword cannot be decoded even if all the other channels are strong. However if a channel is good the corresponding codeword can be decoded even if all the other channels are bad.

With vector coding at a rate R and for code duration N, the transmitter has one codebook with 2^{NR} codewords with power P. The codewords are $M \times N$ matrices and over N symbol durations one of these matrices is transmitted, at the rate of 1 column per symbol period. The elements in the columns are mapped to different inputs of the MIMO diagonal channel. To contrast this case with the scalar coding case, note that with vector coding if one of the channels is bad, it is still possible to decode the full rate R if the other channels are good enough. On the other hand if one channel is good, the codeword cannot be decoded if the other channels are bad.

To summarize the difference, and to strike an analogy with gambling, vector coding bets the entire rate on the 'average' (in some sense) channel and either gets it all or nothing. Scalar coding, on the other hand, bets equally on all the components of the channel and gets a rate proportional to the number of 'good' channels. While one would expect a smaller variance in the rate over each block with scalar coding, it is not obvious which scheme would do better in terms of average throughput.

With our assumptions, the maximum average throughput with scalar codes is given by

$$\begin{aligned} T^s_{max} &= \max_R \sum_{i=1}^{M} \frac{R}{M} \text{Prob}\left(\log\left[1 + \frac{P}{M}h_i\right] > \frac{R}{M}\right) \\ &= \max_R MR \text{ Prob}\left(\log\left[1 + \frac{P}{M}h_1\right] > R\right), \end{aligned}$$

and the maximum average throughput with vector coding is given by

$$\begin{aligned} T^v_{max} &= \max_R R \text{ Prob}\left(\log\left|I + \frac{P}{M}HH^\dagger\right| > R\right) \\ &= \max_R MR \text{ Prob}\left(\frac{\sum_{i=1}^{M}\log\left[1 + \frac{P}{M}h_i\right]}{M} > R\right). \end{aligned}$$

Using Markov's inequality we find that both T^s_{max} and T^v_{max} can be upper-bounded by

$$T_{max} \le M \text{ E}\left[\log\left(1 + \frac{P}{M}h_1\right)\right]. \tag{10}$$

Note that Markov's inequality is a very loose bound for most common distributions. However, at least for large M, applying the law of large numbers we find that T^v_{max} comes close to the bound in (10). Thus we conclude that, asymptotically for large M, vector coding (single codebook) yields higher throughputs than scalar coding (multiple codebooks). Numerical results included in the next section indicate that for Rayleigh fading, vector coding always outperforms scalar coding.

In the next section we present a bigger picture of the tradeoffs between single and multiple codebooks.

VI. Numerical Results and Discussion

First we address the question of single codebook versus multiple codebooks. Note that vector and scalar coding as described in the previous section correspond to using single and multiple codebooks, respectively. Thus the following discussion applies to scalar coding versus vector coding as well.

There are several advantages to using multiple codebooks. Firstly for the same average throughput, partial outages make sure that the variance of the rates achieved over different blocks is smaller with multiple codebooks. Further, splitting the same total rate R over multiple codebooks and using successive decoding at the receiver reduces the decoding complexity at the receiver for the same codeword lengths N. For example, with a single codebook, the receiver needs to compare the received sequence y to 2^{NR} codewords in the codebook to find the closest. However, if we split the rate equally among M codewords for example, then the receiver needs to compare the received sequence y to $2^{\frac{NR}{M}}$ codewords to decode each codeword. In all M such comparisons need to be made. The complexity is still exponentially smaller than that for a single codebook, as $M2^{\frac{NR}{M}} < 2^{NR}$.

There is also a disadvantage associated with multiple codebooks. The number of codebooks used by the transmitter directly increases the feedback bandwidth required. This is because for each transmitted codeword the receiver needs to request a retransmission or to acknowledge successful decoding to the transmitter.

A numerical comparison between achievable average throughputs with single and multiple codebooks on a scalar Rayleigh fading channel is presented in Figure 1. We notice that at low SNRs there is practically no gain from using multiple codebooks, and even at reasonably high SNRs the gains are modest. Also noticeable are the diminishing gains as we increase the number of codewords from 1 to 2 and then to 3.

Figure 1

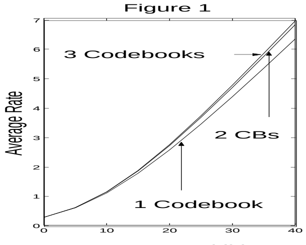
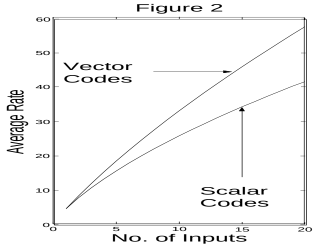


Figure 2



A numerical comparison between scalar and vector coding for Rayleigh fading diagonal MIMO systems is drawn in Figure 2. We notice that the gains with vector coding are significant. The average throughput for 20 inputs at an average transmit SNR of 30 db is about 40% higher for vector coding.

Finally we conclude with a summary of our results in the next section.

VII. Conclusions

We use an information theoretic approach to the problem of average throughput maximization in a block flat fading channel with codeword lengths restricted to less than the fade block duration. We begin by characterizing the maximum achievable average throughput with just one codebook. For log-concave fade distributions, we show that the maximum achievable average throughput with just one codebook is a concave increasing function of the transmit power. As the power increases the average throughput increases both because the optimal code rate increases and because the optimal outage probability decreases. Then we show that using multiple codebooks over orthogonal channels created by splitting the block fading channel in time or frequency does not increase the average throughput. This is essentially a consequence of the concavity of the average throughput as a function of transmit power. However we find that simultaneous transmission of multiple codebooks necessarily increases the maximum achievable average throughput as the number of codebooks is increased. This increase comes at the cost of an increased feedback rate from the receiver as for each codewords the receiver needs to send either a retransmit request or a successful decoding acknowledgment. Numerically, at least for Rayleigh fading, we find that although the average throughput increases, the gains are modest. At low SNRs using multiple codebooks does not increase the average throughput. These results also hold for SIMO systems and can be extended to MISO systems under the assumption that all codewords are transmitted simultaneously over time, frequency, and over all inputs.

For MIMO systems, the optimal mapping of multiple codewords to multiple inputs is hard to find.So we restrict our discussion to some simple cases. We compare the throughputs of scalar coding to vector coding with equal power allocation to all inputs for a diagonal MIMO channel. We find analytically that, asymptotically as the number of inputs increases, vector coding does better. Our numerical results show that for Rayleigh fading, vector coding always does better than scalar coding and the gains are significant.

Thus we conclude that multiple codebooks help to increase the achievable average throughput when they are simultaneously transmitted over the same channel. However, splitting the channel orthogonally in time, frequency, or among inputs of a MIMO channel, and then transmitting different codebooks on each orthogonal sub-channel significantly reduces the achievable average throughput.

References

[1] P. Schramm, H. Andreasson, C. Edholm, N. Edvarcissona, M. Hook, S. Javerbring, F. Muller, J. Skold, "Radio interface performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems", IEEE VTC'98, pp. 1064-1068, 1998.

[2] A. Furuskar, M. Frodigh, H. Olofsson, J. Skold, "System performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems", IEEE VTC'98, pp. 1284-1289.

[3] M. Effros, A. Goldsmith, "Capacity definitions and coding strategies for general channels with receiver side information", in Proc. ISIT 1998.

[4] M.-S, Alouini, A. G. Goldsmith, "Adaptive modulation over Nakagami fading channels", Kluwer WPC, May 2000.

[5] M. Bagnoli, T. Bergstrom, "Log-Concave probability and its applications",ewp-mic/9410002, unpublished.

[6] S. A. Jafar, S. Vishwanath, A. G. Goldsmith, "Throughput Maximization with Multiple Codes and Partial Outages", www.stanford.edu/ syed/globecom2001fullpaper.ps, unpublished.