1

Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels

Lizhong Zheng, *Member, IEEE* David N.C. Tse, *Member, IEEE*

***Abstract*— Multiple antennas can be used for increasing the amount of diversity or the number of degrees of freedom in wireless communication systems. In this paper, we propose the point of view that both types of gains can be simultaneously obtained for a given multiple antenna channel, but there is a fundamental tradeoff between how much of each any coding scheme can get. For the richly scattered Rayleigh fading channel, we give a simple characterization of the optimal tradeoff curve and use it to evaluate the performance of existing multiple antenna schemes.**

***Index Terms*— diversity, multiple antennas, MIMO, spatial multiplexing, space-time codes .**

I. Introduction

Multiple antennas are an important means to improve the performance of wireless systems. It is widely understood that in a system with multiple transmit and receive antennas (MIMO channel), the spectral efficiency is much higher than that of the conventional single antenna channels. Recent research on multiple antenna channels, including the study of channel capacity [1], [2] and the design of communication schemes [3], [4], [5], demonstrates a great improvement of performance.

Traditionally, multiple antennas have been used to increase *diversity* to combat channel fading. Each pair of transmit and receive antennas provides a signal path from the transmitter to the receiver. By sending signals that carry the same information through different paths, multiple independently faded replicas of the data symbol can be obtained at the receiver end; hence more reliable reception is achieved. For example, in a slow Rayleigh fading environment with 1 transmit and n *receive* antennas , the transmitted signal is passed through n different paths. It is well known that if the fading is independent across antenna pairs, a maximal diversity gain (advantage) of n can be achieved: the average error probability can be made to decay like $1/\mathsf{SNR}^n$ at high SNR, in contrast to the SNR^{-1} for the single antenna fading channel. More recent work has concentrated on using multiple *transmit* antennas to get diversity (some examples are trellis-based space-time codes [6], [7] and orthogonal designs [8], [3]). However, the underlying idea is still averaging over multiple path gains (fading coefficients) to increase the reliability. In a system with m transmit and n receive antennas, assuming the path gains between individual antenna pairs are i.i.d. Rayleigh faded, the maximal diversity gain is mn, which is the total number of fading gains that one can average over.

both authors are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720

This research is supported by a National Science Foundation Early Faculty CAREER Award, with matching grants from A.T.&T., Lucent Technologies and Qualcomm Inc., and by the National Science Foundation under grant CCR-01-18784

Transmit or receive diversity is a means to *combat* fading. A different line of thought suggests that in a MIMO channel, fading can in fact be *beneficial* through increasing the *degrees of freedom* available for communication [2], [1]. Essentially, if the path gains between individual transmit-receive antenna pairs fade independently, the channel matrix is well-conditioned with high probability, in which case multiple parallel *spatial channels* are created. By transmitting independent information streams in parallel through the spatial channels, the data rate can be increased. This effect is also called *spatial multiplexing* [5], and is particularly important in the high signal-to-noise ratio (SNR) regime where the system is degree-of-freedom-limited (as opposed to power-limited). Foschini [2] has shown that in the high SNR regime, the capacity of a channel with m transmit, n receive antennas and i.i.d. Rayleigh faded gains between each antenna pair is given by:

$$C(\mathsf{SNR}) = \min\{m,n\} \log \mathsf{SNR} + O(1).$$

The number of degrees of freedom is thus the minimum of m and n. In recent years, several schemes have been proposed to exploit the spatial multiplexing phenomenon(for example BLAST [2]).

In summary, a MIMO system can provide two types of gains: diversity gain and spatial multiplexing gain. Most of current research focuses on designing schemes to extract either maximal diversity gain *or* maximal spatial multiplexing gain. (There are also schemes which switch between the two modes, depending on the instantaneous channel condition [5].) However, maximizing one type of gain may not necessarily maximize the other. For example, it was observed in [9] that the coding structure from the orthogonal designs [3], while achieving the full diversity gain, reduces the achievable spatial multiplexing gain. In fact, each of the two design goals addresses only one aspect of the problem. This makes it difficult to compare the performance between diversity-based and multiplexing-based schemes

In this paper, we put forth a different viewpoint: given a MIMO channel, both gains can in fact be *simultaneously* obtained, but there is a *fundamental tradeoff* between how much of each type of gain any coding scheme can extract: higher spatial multiplexing gain comes at the price of sacrificing diversity. Our main result is a simple characterization of the optimal tradeoff curve achievable by *any* scheme. To be more specific, we focus on the high SNR regime, and think of a *scheme* as a family of codes, one for each SNR level. A scheme is said to have a spatial multiplexing gain r and a diversity advantage d if the rate of the scheme scales like $r \log \mathsf{SNR}$ and the average error probability decays like $1/\mathsf{SNR}^d$. The optimal tradeoff curve yields for each multiplexing gain r the optimal diversity advantage $d^*(r)$ achievable by *any* scheme. Clearly, r cannot exceed the total number of degrees of freedom $\min\{m, n\}$ provided by the channel; and $d^*(r)$ cannot exceed the maximal diversity gain mn of the channel. The tradeoff curve bridges between these two extremes. By studying the optimal tradeoff, we reveal the relation between the two types of gains, and obtain insights to understand the overall resources provided by multiple antenna channels.

For the i.i.d. Rayleigh flat fading channel, the optimal tradeoff turns out to be very simple for most system parameters of interest. Consider a slow fading environment in which the channel gain is random but remains constant for a duration of l symbols. We show that as long as the block length $l \geq m+n-1$, the optimal diversity gain $d^*(r)$ achievable by any coding scheme of block length l and multiplexing gain r (r integer) is precisely $(m-r)(n-r)$. This suggests an appealing interpretation: out of the total resource of m transmit and n receive antennas, it is *as though* r transmit and r receive antennas were used for multiplexing and the remaining $m-r$ transmit and $n-r$ receive antennas provided the diversity. Thus, by adding one transmit and one receive antenna, the spatial multiplexing gain can be increased by one while maintaining the *same* diversity level. It should also be observed that this optimal tradeoff does not depend on l as long as $l \geq m+n-1$; hence, no more diversity gain can be extracted by coding over block lengths greater than $m+n-1$ than using a block length equal to $m+n-1$.

The tradeoff curve can be used as a unified framework to compare the performance of many existing diversity-based and multiplexing-based schemes. For several well-known schemes, we compute the achieved tradeoff curves $d(r)$ and compare it to the optimal tradeoff curve. That is, the performance of a scheme is evaluated by the tradeoff it achieves. By doing this, we take into consideration not only the capability of the scheme to combat against fading, but also its ability to accommodate higher data rate as SNR increases, and therefore provide a more complete view.

The diversity-multiplexing tradeoff is essentially the tradeoff between the error probability and the data rate of a system. A common way to study this tradeoff is to compute the *reliability function* from the theory of *error exponents* [10]. However, there is a basic difference between the two formulations: while the traditional reliability function approach focuses on the asymptotics of *large block lengths*, our formulation is based on the asymptotics of *high SNR* (but fixed block length). Thus, instead of using the machinery of the error exponent theory, we exploit the special properties of fading channels and develop a simple approach, based on the outage capacity formulation [11], to analyze the diversity-multiplexing tradeoff in the high SNR regime. On the other hand, even though the asymptotic regime is different, we do conjecture an intimate connection between our results and the theory of error exponents.

The rest of the paper is outlined as follows. Section II presents the system model and the precise problem formulation. The main result on the optimal diversity-multiplexing tradeoff curve is given in Section III, for block length $l \geq m+n-1$. In Section IV, we derive bounds on the tradeoff curve when the block length is less than $m+n-1$. While the analysis in this section is more technical in nature, it provides more insights to the problem. Section V studies the case when spatial diversity is combined with other forms of diversity. Section VI discusses the connection between our results and the theory of error exponents. We compare the performance of several schemes with the optimal tradeoff curve in Section VII. Section VIII contains the conclusions.

II. System Model and Problem Formulation

A. Channel Model

We consider the wireless link with m transmit and n receive antennas. The fading coefficient $\mathbf{h}_{ij}$ is the complex path gain from transmit antenna j to receive antenna i. We assume that the coefficients are independently Rayleigh distributed with unit variance, and write $\mathbf{H} = [\mathbf{h}_{ij}] \in \mathcal{C}^{n \times m}$. $\mathbf{H}$ is assumed to be known to the receiver, but not at the transmitter. We also assume that the channel matrix $\mathbf{H}$ remains constant within a block of l symbols, i.e. the block length is much small than the channel coherence time. Under these assumptions, the channel, within one block, can be written as:

$$\mathbf{Y} = \sqrt{\frac{\mathsf{SNR}}{m}}\mathbf{H}\mathbf{X} + \mathbf{W} \tag{1}$$

where $\mathbf{X} \in \mathcal{C}^{m \times l}$ has entries $\mathbf{x}_{ij}, i = 1, \ldots, m, j = 1, \ldots, l$ being the signals transmitted from antenna i at time j; $\mathbf{Y} \in \mathcal{C}^{n \times l}$ has entries $\mathbf{y}_{ij}, i = 1, \ldots, n, j = 1, \ldots, l$ being the signals received from antenna i at time j; the additive noise $\mathbf{W}$ has i.i.d. entries $\mathbf{w}_{ij} \sim \mathcal{CN}(0, 1)$; SNR is the average signal to noise ratio at each receive antenna.

We will first focus on studying the channel within this single block of l symbol times. In section V, our results are generalized to the case when there is a multiple of such blocks, each of which experiences independent fading.

A rate R bps/Hz codebook $\mathcal{C}$ has $|\mathcal{C}| = \lfloor 2^{Rl} \rfloor$ codewords $\{X(1), \ldots, X(|\mathcal{C}|)\}$, each of which is an $m \times l$ matrix. The transmitted signal $\mathbf{X}$ is normalized to have the average transmit power at each antenna in each symbol period to be 1. We interpret this as an overall power constraint on the codebook $\mathcal{C}$:

$$\frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} \|X(i)\|_F^2 \leq ml. \tag{2}$$

where $\|.\|_F$ is the Frobenius norm of a matrix: $\|R\|_F^2 \triangleq \sum_{i,j} \|R_{ij}\|^2 = \text{trace}(RR^\dagger)$.

B. Diversity and Multiplexing

Multiple antenna channels provide *spatial diversity*, which can be used to improve the reliability of the link. The basic idea is to supply to the receiver multiple independently faded replicas of the same information symbol, so that the probability that all the signal components fade simultaneously is reduced.

As an example, consider uncoded binary PSK signals over a single antenna fading channel ($m = n = l = 1$ in the above model). It is well known [12] that the probability of error at high SNR (averaged over the fading gain $\mathbf{H}$ as well as the additive noise) is

$$P_e(\mathsf{SNR}) \approx \frac{1}{4}\mathsf{SNR}^{-1}.$$

In contrast, transmitting the same signal to a receiver equipped with 2 antennas, the error probability is

$$P_e(\mathsf{SNR}) \approx \frac{3}{16}\mathsf{SNR}^{-2}.$$

Here we observe that by having the extra receive antenna, the error probability decreases with SNR at a faster speed of SNR^{-2}. This phenomenon implies that at high SNR, the error probability is much smaller. Similar results can be obtained if we change the binary PSK signals to other constellations. Since the performance gain at high SNR is dictated by the SNR exponent of the error probability, this exponent is called the *diversity gain*. Intuitively, it corresponds to the number of independently faded paths that a symbol passes through; in other words, the number of independent fading coefficients that can be averaged over to detect the symbol. In a general system with m transmit and n receive antennas, there are in total $m \times n$ random fading coefficients to be averaged over; hence the *maximal (full) diversity gain* provided by the channel is mn.

Besides providing diversity to improve reliability, multiple antenna channels can also support a higher data rate than single antenna channels. As an evidence of this, consider an ergodic block fading channel in which each block is as in (1) and the channel matrix is independent and identically distributed across blocks. The ergodic capacity (bps/Hz) of this channel is well-known [1], [2]:

$$C(\mathsf{SNR}) = \mathcal{E}\left[\log\det\left(I + \frac{\mathsf{SNR}}{m}\mathbf{H}\mathbf{H}^\dagger\right)\right]$$

At high SNR

$$C(\mathsf{SNR}) = \min\{m,n\}\log\frac{\mathsf{SNR}}{m} + \sum_{i=|m-n|+1}^{\max\{m,n\}} \mathcal{E}[\log\chi^2_{2i}] + o(1),$$

where χ^2_{2i} is Chi-square distributed with $2i$ degrees of freedom. We observe that at high SNR, the channel capacity increases with SNR as $\min\{m,n\}\log\mathsf{SNR}$ (bps/Hz), in contrast to $\log\mathsf{SNR}$ for single antenna channels. This result suggests that the multiple antenna channel can be viewed as $\min\{m,n\}$ parallel *spatial channels*; hence the number $\min\{m,n\}$ is the total *number of degrees of freedom* to communicate. Now one can transmit independent information symbols in parallel through the spatial channels. This idea is also called *spatial multiplexing*.

Reliable communication at rates arbitrarily close to the ergodic capacity requires averaging across many independent realizations of the channel gains over time. Since we are considering coding over only a single block, we must lower the data rate and step back from the ergodic capacity to cater for the randomness of the channel $\mathbf{H}$. Since the channel capacity increases linearly with $\log\mathsf{SNR}$, in order to achieve a certain fraction of the capacity at high SNR, we should consider schemes that support a data rate which also increases with SNR. Here, we think of a *scheme* as a family of codes $\{\mathcal{C}(\mathsf{SNR})\}$ of block length l, one at each SNR level. Let $R(\mathsf{SNR})$ (bits/symbol) be the rate of the code $\mathcal{C}(\mathsf{SNR})$. We say that a scheme achieves a *spatial multiplexing gain* of r if the supported data rate

$$R(\mathsf{SNR}) \approx r\log\mathsf{SNR} \text{ (bps/Hz)}$$

One can think of spatial multiplexing as achieving a *non-vanishing* fraction of the degrees of freedom in the channel. According to this definition, any fixed-rate scheme has a zero multiplexing gain, since eventually at high SNR, any fixed data rate is only a vanishing fraction of the capacity.

Now to formalize, we have the following definition.

Definition 1: A scheme $\{\mathcal{C}(\mathsf{SNR})\}$ is said to achieve *spatial multiplexing gain* r and *diversity gain* d if the data rate

$$\lim_{\mathsf{SNR}\to\infty}\frac{R(\mathsf{SNR})}{\log\mathsf{SNR}} = r$$

and the average error probability

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log P_e(\mathsf{SNR})}{\log\mathsf{SNR}} = -d \qquad (3)$$

For each r, define $d^*(r)$ to be the supremum of the diversity advantage achieved over all schemes. We also define

$$d^*_{max} \triangleq d^*(0)$$
$$r^*_{max} \triangleq \sup\{r : d^*(r) > 0\}$$

which are respectively the maximal diversity gain and the maximal spatial multiplexing gain in the channel.

Throughout the rest of the paper, we will use the special symbol $\doteq$ to denote *exponential equality*, i.e., we write $f(\mathsf{SNR}) \doteq \mathsf{SNR}^b$ to denote

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log f(\mathsf{SNR})}{\log\mathsf{SNR}} = b$$

and $\dot{\geq}, \dot{\leq}$ are similarly defined. (3) can thus be written as

$$P_e(\mathsf{SNR}) \doteq \mathsf{SNR}^{-d}.$$

The error probability $P_e(\mathsf{SNR})$ is averaged over the additive noise $\mathbf{W}$, the channel matrix $\mathbf{H}$ and the transmitted codewords (assumed equally likely). The definition of diversity gain here differs from the standard definition in the space-time coding literature (see for example [7]) in two important ways:

- This is the *actual* error probability of a code, and not the *pairwise* error probability between two codewords as is commonly used as a diversity criterion in space-time code design.

- In the standard formulation, diversity gain is an asymptotic performance metric of one *fixed* code. To be specific, the input of the fading channel is fixed to be a particular code, while SNR increases. The speed that the error probability (of a ML detector) decays as SNR increases is called the diversity gain. In our formulation, we notice that the channel capacity increases linearly with $\log \mathsf{SNR}$. Hence in order to achieve a non-trivial fraction of the capacity at high SNR, the input data rate must also *increase* with SNR, which requires a sequence of codebooks with increasing size. The diversity gain here is use as a performance metric of such a sequence of codes, which is formulated as a "scheme". Under this formulation, any fixed code has 0 spatial multiplexing gain. *Allowing both the data rate and the error probability scale with the* SNR *is the crucial element of our formulation and, as we will see, allows us to to talk about their tradeoff in a meaningful way.*

The spatial multiplexing gain can also be thought as the data rate normalized with respect to the SNR level. A common way to characterize the performance of a communication scheme is to compute the error probability as a function of SNR for a fixed data rate. However, different designs may support different data rate. In order to compare these schemes fairly, Forney [13] proposed to plot the error probability against the *normalized* SNR:

$$\mathsf{SNR}_{norm} \triangleq \frac{\mathsf{SNR}}{C^{-1}(R)}.$$

where $C(\mathsf{SNR})$ is the capacity of the channel as a function of SNR. That is, SNR_{norm} measures how far the SNR is above the minimal required to support the target data rate.

A dual way to characterize the performance is to plot the error probability as a function of the data rate, for a fixed SNR level. Analogous to Forney's formulation, to take into consideration the effect of the SNR, one should use the *normalized data rate* R_{norm} instead of R:

$$R_{norm} \triangleq \frac{R}{C(\mathsf{SNR})}$$

which indicates how far a system is operating from the Shannon limit. Notice that at high SNR, the capacity of the multiple antenna channel is $C(\mathsf{SNR}) \approx \min\{m,n\} \log \mathsf{SNR}$; hence the spatial multiplexing gain

$$r = \frac{R}{\log \mathsf{SNR}} \approx \min\{m,n\} R_{norm}$$

is just a constant multiple of R_{norm}.

III. Optimal Tradeoff: $l \geq m+n-1$ case

In this section, we will derive the optimal tradeoff between the diversity gain and the spatial multiplexing gain that any scheme can achieve in the Rayleigh fading multiple antenna channel. We will first focus on the case that the block length $l \geq m+n-1$, and discuss the other cases in section IV.

A. Optimal Tradeoff Curve

The main result is given in the following theorem.

Theorem 2: Assume $l \geq m+n-1$. The optimal tradeoff curve $d^*(r)$ is given by the piecewise linear function connecting the points $(k, d^*(k)), k = 0, 1, \ldots, \min\{m,n\}$, where

$$d^*(k) = (m-k)(n-k) \quad (4)$$

In particular, $d^*_{max} = mn$, and $r^*_{max} = \min\{m,n\}$.

The function $d^*(r)$ is plotted in Figure 1.

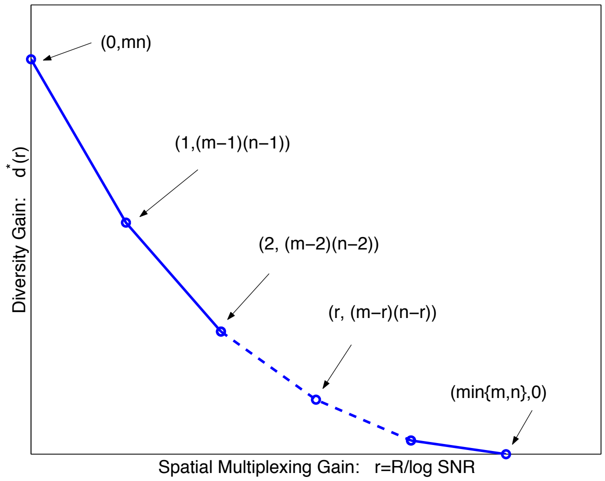

Fig. 1. Diversity-multiplexing tradeoff, $d^*(r)$ for general m, n and $l \geq m+n-1$.

The optimal tradeoff curve intersects the r axis at $\min\{m,n\}$. This means that the maximum achievable spatial multiplexing gain r^*_{max} is the total number of degrees of freedom provided by the channel as suggested by the ergodic capacity result in (3). Theorem 2 says that at this point, however, no positive diversity gain can be achieved. Intuitively, as $r \to r^*_{max}$, the data rate approaches the ergodic capacity and there is no protection against the randomness in the fading channel.

On the other hand, the curve intersects the d axis at the maximal diversity gain $d^*_{max} = mn$, corresponding to the total number of random fading coefficients that a scheme can average over. There are known designs that achieve the maximal diversity gain at a fixed data rate [8]. Theorem 2 says that in order to achieve the maximal diversity gain, no positive spatial multiplexing gain can be obtained at the same time.

The optimal tradeoff curve $d^*(r)$ bridges the gap between the above two design criteria, by connecting the two extreme points: $(0, d^*_{max})$ and $(r^*_{max}, 0)$. This result says that positive diversity gain and spatial multiplexing gain can be achieved simultaneously. However, increasing the diversity advantage comes at a price of decreasing the spatial multiplexing gain, and vice versa. The tradeoff curve provides a more complete picture of the achievable performance over multiple antenna channels than the two extreme points corresponding to the maximum diversity gain and multiplexing gain. For example, the ergodic capacity result suggests that by increasing the minimum of the number of transmit and receive antennas, $\min\{m,n\}$, by one, the channel gains one more degree of

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase® Smarter legal research.