DIVERSITY AND DEGREES OF FREEDOM IN WIRELESS COMMUNICATIONS

Mahesh Godavarti and Alfred O. Hero III

godavarti@altrabroadband.com, hero@eecs.umich.edu

ABSTRACT

We introduce rigorous definitions for two quantities of interest, diversity and degrees of freedom, that are used to quantify the advantages of a multiple antenna multiple-input multiple-output (MIMO) system when compared to a single-input single-output (SISO) system. These definitions are in a general setting which will allow the computation of these quantities for systems other than multiple antenna MIMO systems. We verify the effectiveness of the definitions by computing the quantities of interest for various existing examples.

1. INTRODUCTION

The quality of any communications link is indexed over at least 2 dimensions (rate and probability of error). Furthermore, the fundamental limitations of any link can be indexed by the associated channel capacity, relating to max achievable rate, and the minimum achievable probability of coding error, related to the optimal Bayes receiver (possibly constrained). A natural characterization of relative improvement or degradation on link performance due to deployment of different physical channels (e.g. via adding antennas, widening the frequency band, etc) could be the relative improvement of these two quantities. However, it is obvious that it is desirable to parameterize the two performance gains by a single post-processing signal to noise ratio (SNR) gain factor when this is possible. Unfortunately, for all but the most trivial SISO channels the SNR gain alone is not sufficient to specify the capacity and minimum probability of error gains.

The Degrees of freedom (DOF) and Diversity (DIV) represent a simple supplementary index of the relative performance gain which complements the information provided by SNR gain information in these cases.

Diversity [8] and Degrees of Freedom [10] have been widely used as the two measures of performance gains obtained from using multiple antennas at the transmitter and the receiver in the field of space-time coding. In [8] Diversity (DIV) has been defined as the negative exponent of SNR in the probability of error expression for high SNR and in [10] Degrees of Freedom (DOF) has been defined as the co-efficient of $\log \rho$ occurring in the expression for capacity, again for high values of the signal to noise ratio ρ.

Traditionally, DIV has been thought of as the number of independent channels available through which replicas of the same information signal can be transmitted [7, pp. 709-712]. This in essence gives multiple copies of the same signal thereby reducing the probability of decoding error. This can be easily confused with the number of independent channels available for communication that Zheng and Tse [10] define to be the DOF of the system. This can give rise to considerable confusion as DIV and DOF for a MIMO system operating in a Rayleigh/Rician fading environment don't necessarily agree with each other. Even though [8] refers to DIV as redundancy in the system it doesn't necessarily shed light on the nature of the two measures.

We suggest that DIV and DOF should be considered in the context of a particular communication system in the sense that intuitively DIV can be considered as the redundancy or the maximum number of independent copies of the transmitted signal in a particular communication system whereas DOF can be considered as the number of independent channels available for communication again in a particular communication system. A communication system comprises of all the user controllable parameters such as the input signaling scheme, coding, decoding, the particular choice of transmit and receive antennas etc. These intuitive definitions even though well accepted in the literature are not sufficient to understand the distinction between the two measures. Also, the existing definitions [8, 10] are not general enough to be applicable to all situations. In this paper, we address both issues.

In [2] the authors refer to using multiple antennas at the receiver and at the transmitter to increase information capacity as receive diversity and transmit diversity, respectively. This can also lead to confusion as people can associate increase in diversity with an increase in capacity.

2. DIVERSITY AND DEGREES OF FREEDOM

It's obvious that the concepts of diversity and degrees of freedom are basically a measure of the performance of a channel in terms of probability of error and information capacity. Also, it can be seen that it is very difficult to first evaluate DOF and DIV and then translate them into channel capacity and probability of error, respectively. In that respect, it makes sense to define DOF in terms of information capacity and DIV in terms of probability of error directly.

Let $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ be the set of all link parameters that determine the diversity and degrees of freedom in the communication system. For example, P_1 could be M, the number of transmit antennas; P_2 could be N, the number of receive antennas; and P_3 could be t, the number of channel uses. Fixing $\mathcal{P}$ needn't fix the channel used in a communication system. Let $H(P_1, P_2, \ldots, P_n)$ denote one of the choices for the channel when the parameter set is fixed to be $\mathcal{P}$. For example, in the case of M transmit antennas and N receive antennas there are a total of MN choices for $H(1, 1)$. The channel specified by $\mathcal{P}$ in general could be stochastic, for example, H in the Rayleigh fading model. ρ is the measure of reliability in the channel with the channel becoming more reliable as $\rho \to \infty$. For example, in the

Mahesh Godavarti is with Altra Broadband, Inc. and Alfred Hero is with the University of Michigan. This work was partially completed while Mahesh Godavarti was a PhD student under the supervision of Prof. Alfred Hero.

MIMO system the reliability measure ρ is the signal to noise ratio. In the binary symmetric channel with crossover probability p, $p \to 0$ is equivalent to $\rho \to \infty$. Let $P_e(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})$ denote the probability of error in a particular communication system $\mathcal{C}$ when the parameters of interest are $(P_1, P_2, \ldots, P_n)$. $\mathcal{C}$ as stated earlier is specified by all the user controllable parameters like the input probability distribution $p(S)$, the input alphabet $\mathcal{S}$, the choice of the available channel $H(P_1, P_2, \ldots, P_n)$ for a fixed parameter set $(P_1, \ldots, P_n)$, the decoding structure/strategy (coherent vs non-coherent and hard vs soft decoding etc.) and the transmitter structure (coding etc.).

Similarly, let $R(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})$ denote the rate of communication in the communication system $\mathcal{C}$. Then

Definition 1

$$\mathrm{DIV}_{P_1,\ldots,P_n}(\mathcal{C}) = \lim_{\rho\to\infty} \frac{\log P_e(\rho, P_1, \ldots, P_n, \mathcal{C})}{\sup_{\mathcal{C}} \log P_e(\rho, 1, \ldots, 1, \mathcal{C})} \quad (1)$$

and

$$\mathrm{DOF}_{P_1,\ldots,P_n}(\mathcal{C}) = \lim_{\rho\to\infty} \frac{R(\rho, P_1, \ldots, P_n, \mathcal{C})}{\sup_{\mathcal{C}} R(\rho, 1, \ldots, 1)}. \quad (2)$$

We will assume that the above limits exist whenever required. Note that $\sup_{\mathcal{C}} R(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})$ can simply be replaced by $\sup_{H(P_1,P_2,\ldots,P_n)} C_H(\rho, P_1, P_2, \ldots, P_n)$ where $C_H(\rho, P_1, P_2, \ldots, P_n)$ denotes the capacity of the link when the channel is given by $H(P_1, P_2, \ldots, P_n)$.

We will assume that $H(P_1, P_2, \ldots, P_n)$ is unique when P_1, $P_2 \ldots P_n$ take the maximum possible values. Then

Definition 2 DIV *and* DOF *for the channel* $H(P_1, P_2, \ldots, P_n)$, *are defined as*

$$\mathrm{DIV}_{P_1,\ldots,P_n} = \sup_{\mathcal{C}} \mathrm{DIV}_{P_1,\ldots,P_n}(\mathcal{C}) \quad (3)$$

and

$$\mathrm{DOF}_{P_1,\ldots,P_n} = \sup_{\mathcal{C}} \mathrm{DOF}_{P_1,\ldots,P_n}(\mathcal{C}). \quad (4)$$

Note that it can be easily shown that

$$\mathrm{DOF}_{P_1,\ldots,P_n} = \lim_{\rho\to\infty} \frac{C(\rho, P_1, \ldots, P_n)}{\sup_{H(1,\ldots,1)} C_H(\rho, 1, \ldots, 1)}. \quad (5)$$

From now on, we will use (5) as the definition for DOF of a channel.

3. EXAMPLES

For the first example in this section, we need the following Gallager error exponent [3]. For a system communicating at a rate R the upper bound on probability of error is given as follows

$$P_e \leq \exp\left(-n \max_{p(S)} \max_{0\leq\gamma\leq 1} \left[E_0(\gamma, p(S)) - \gamma R\right]\right) \quad (6)$$

where n is the length of codebook used and $E_0(\gamma, p(S))$ is as follows

$$E_0(\gamma, p(S)) = -\log \int \left[\int p(S) p(X|S)^{\frac{1}{1+\gamma}} dS\right]^\gamma dX$$

where S is the input to the channel and X is the observed output.

Example 1 *Consider a single-input single-output discrete-time additive white Gaussian noise (AWGN) channel* $x_l = \sqrt{\rho} h s_l + w_l$, $l = 1, \ldots, t$. *h is a deterministic complex number and w_l is a complex circular Gaussian random variable with mean zero and variance 1. We will calculate* DIV *and* DOF *for t-uses of the channel under the average energy constraint* $\sum_{l=1}^{t} E[s_l s_l^*] \leq t$.

First, consider $\mathcal{C}_1$ consisting of a binary input taking values over $\{-1, 1\}$ with equal probability. The code used is a repetition code. That is either $s_l = -1$ for $l = 1, \ldots, t$ or $s_l = 1$ for $l = 1, \ldots, t$. The decoding at the receiver is Maximum A Posteriori (MAP) decoding. Since the probability of error when using the repetition code of length t is given by $Q(t\rho)$ where $Q(x) = \frac{1}{2\pi}\int_x^{-\infty} e^{-y^2/2} dy$ as opposed to $Q(\rho)$ when $t = 1$, we obtain $\mathrm{DIV}_t(\mathcal{C}_1) = t$. *Since $R(\rho, t, \mathcal{C}_1) = 1/t$ irrespective of the value of ρ,* $\mathrm{DOF}_t(\mathcal{C}_1) = 0$.

Next, consider $\mathcal{C}_2$ where s_l is an i.i.d. complex circular Gaussian random variable with mean zero and unit variance for $l = 1, \ldots, t$. This is the capacity achieving signal density. The decoding at the output is MAP decoding. In this case, $\mathrm{DOF}_t(\mathcal{C}_2) = t$ *since $R(\rho, t, \mathcal{C}_2) = t\log(1+\rho)$. For this value of R ($R = C$, the capacity) the value of γ that maximizes Gallager's error exponent is zero and at $\gamma = 0$, $E_0(\gamma, p(S)) = 1$. Therefore,* $\mathrm{DIV}_t(\mathcal{C}_2) = 0$.

Therefore, $\mathrm{DIV}_t = t$ *and* $\mathrm{DOF}_t = t$.

We believe that DIV_t being equal to DOF_t in this example can cause confusion between DIV and DOF.

It is quite intuitive to expect that DIV and DOF depend on each other. From the definitions, it is obvious that they are related to each other parametrically through $\mathcal{C}$. Indeed, in the example given above we see that when a communication system is operating at maximum diversity ($\sup_{\mathcal{C}} \mathrm{DIV}(\mathcal{C})$), the corresponding DOF is zero whereas when the system is operating at maximum degrees of freedom ($\sup_{\mathcal{C}} \mathrm{DOF}(\mathcal{C})$), the corresponding DIV is zero. This point is illustrated further in Example 3.

Example 2 *Now, let's consider the following multiple antenna Rayleigh fading MIMO system*

$$X = \sqrt{\frac{\rho}{M}} SH + W$$

where X, S and H are $T \times N$, $T \times M$ and $M \times N$ sized matrices, respectively. Let H be known at the receiver. In this model T is the time interval over which H remains constant. We will investigate DIV *and* DOF *for $t = T$ channel uses under the average energy constraint* $\mathrm{tr}\{E[SS^\dagger]\} \leq tM$. *We assume that H is known to the receiver. Let's again consider two systems $\mathcal{C}_1$ and $\mathcal{C}_2$.*

In $\mathcal{C}_1$, we use the signaling scheme developed by Tarokh et. al. [8, pp. 747–749]. The decoding at the output is chosen to be MAP decoding. Using our definition for DIV *and Tarokh's development for probability of error [8, (10), p. 749] we conclude that* $\mathrm{DIV}_{M,N,t}(\mathcal{C}_1) = N\min(M, t)$ *which agrees with Tarokh's conclusion about diversity. However, from [8, (18), p. 755] we see that the rate R is bounded above by a constant independent of ρ. Therefore,* $\mathrm{DOF}_{M,N,t}(\mathcal{C}_1) = 0$.

In $\mathcal{C}_2$, we choose the elements of the matrix S to be i.i.d. complex circular Gaussian random variables with mean zero and variance one. This is the capacity achieving signal. We choosing the decoding strategy to be MAP. Therefore [4],

$$R(\rho, M, N, t, \mathcal{C}_2) = C(\rho, M, N, t) = tE \log\det(I_M + \frac{\rho}{M} HH^\dagger).$$

This shows that $\mathrm{DOF}_{M,N,t}(\mathcal{C}_2) = \mathrm{DOF}_{M,N,t} = t\min(M, N)$.

In Example 2, we see that DIV of a MIMO system is linear in the number of receive antennas. This makes sense intuitively because receive antennas provide natural redundancy in the system. By increasing the number of receive antennas we get many replicas of the transmitted signal and hence greater error protection.

Example 3 *Consider the same system as in Example 2. A lower bound on the error exponent for this system can be calculated as in [9]. By choosing the input distribution $p(S)$, to be i.i.d. complex circular Gaussian $p_G(S)$ (capacity achieving distribution), the error exponent is:*

$$E_0(\gamma, p_G(S), \rho) = -\log E\left[\det\left(I_M + \frac{1}{1+\gamma}\frac{\rho}{M}HH^\dagger\right)^{-\gamma t}\right]$$

where we have chosen to make the dependence of $E_0(\cdot)$ on ρ explicit. Given a rate R, the upper bound on the probability of error is given by

$$P_e \leq \exp\left(-\left[E_0(\gamma, p_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]\right)$$

where γ is chosen so that $R = \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}$.

Fixing γ we see that as $\rho \to \infty$

$$R(\rho, M, N, t, \gamma) = \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}$$

and we obtain a fixed value of DOF*. Similarly, for a fixed value of γ, as $\rho \to \infty$,*

$$P_e(\rho, M, N, t, \gamma) \leq \exp\left(-\left[E_0(\gamma, p_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]\right)$$

and we obtain a fixed value of DIV*. That implies we can treat γ as a particular communication system $\mathcal{C}_\gamma$. We can then plot*

$$\mathrm{DOF}_{M,N,t}(\mathcal{C}_\gamma) = \lim_{\rho\to\infty}\frac{\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}}{\sup_{H(1,1,1)} C_H(\rho, 1, 1, 1, \mathcal{C}_\gamma)}$$

versus

$$\mathrm{DIV}_{M,N,t}(\mathcal{C}_\gamma) = \lim_{\rho\to\infty}\frac{-\left[E_0(\gamma, p_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]}{\sup_{\mathcal{C}_\gamma}\log P_e(\rho, 1, 1, 1, \mathcal{C}_\gamma)}$$

parameterized by γ. One such plot for $t = T = 5$, $N = 1$ and $M = 3$ is shown in Figure 1.

Example 3 shows that if the system is operating at optimum DOF then it can't be operating at optimum DIV and vice-versa.

Example 4 *Consider the following channel*

$$x_1 = \sqrt{\frac{\rho}{2}}s_1 + n_1, \;\; x_2 = \sqrt{\frac{\rho}{2}}hs_2 + n_2$$

where h, n_1 and n_2 are i.i.d. circular Gaussian random variables with mean zero and unit variance. We calculate the diversity of this channel under the constraint $E[|s_1|^2 + |s_2|^2] \leq 2$. Choose a communication system $\mathcal{C}_\lambda$ such that input the signal is given by $s_1 = \sqrt{\lambda}s$ and $s_2 = \sqrt{1-\lambda}s$ with $E[|s|^2] = 2$. Choose s to be a binary signal taking values over $\{-\sqrt{2}, \sqrt{2}\}$ with equal probability. The probability of decoding $\sqrt{2}$ as $-\sqrt{2}$ is given by $E[Q(\sqrt{\frac{\rho}{2}d^2(-\sqrt{2}, \sqrt{2})})]$ where $d^2(-\sqrt{2}, \sqrt{2}) = 2\lambda + 2(1-\lambda)|h|^2$ and the expectation $E[\cdot]$, is over the density function of $|h|$. Therefore, the probability of error for high ρ behaves like $P_e(\rho, 2, 2, \mathcal{C}_\lambda) \approx \frac{K}{\rho}e^{-\lambda\rho}$ where K is some constant. Therefore

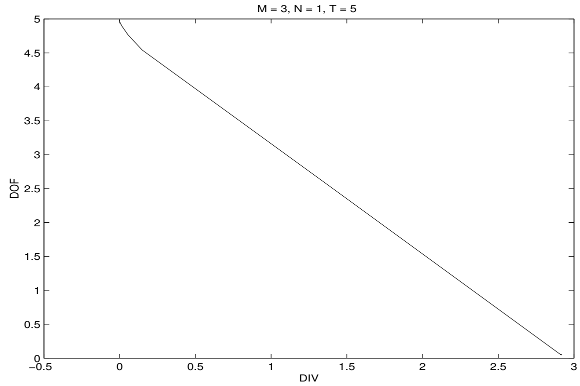

Fig. 1. DOF as a function of DIV

$$\mathrm{DIV}_{\mathcal{C}_\lambda} = \lim_{\rho\to\infty}\frac{\log P_e(\rho, 2, 2, \mathcal{C}_\lambda)}{\sup_{\mathcal{C}}\log P_e(\rho, 1, 1, \mathcal{C})} = \lambda.$$

Maximizing over λ we obtain $\mathrm{DIV} = 1$.

In the example above, we see that eventhough the receiver has two independent copies of the transmitted signal the maximum available DIV is 1 and not 2 as dictated by intuition.

Example 5 *Let's reconsider the case of AWGN channel in Example 1 operating at an optimal rate in a communication system with a binary input ($\{-1, 1\}$) and hard decision decoding at the receiver ($\mathcal{C}_1$). In this case, the channel effectively behaves like a binary symmetric channel with a crossover probability $p = Q(\sqrt{\rho}) \approx \frac{c}{\rho}\exp(-\rho/2)$ where c is some constant. We will calculate* DOF *for t channel uses for $\mathcal{C}_1$. We note that as $\rho \to \infty$, $p \to 0$ and the maximum achievable rate for this system is 1 bit per channel use. That is, $\lim_{\rho\to\infty} R(\rho, t, \mathcal{C}_1) = t$. Therefore, $\mathrm{DOF}_t(\mathcal{C}_1)$ is zero.*

Now consider $\mathcal{C}_2$ which is similar to $\mathcal{C}_1$ except that the channel is no longer operating at the optimal rate and the communication system has repetition coding at the transmitter. We will calculate DIV *for $\mathcal{C}_2$ corresponding to t, t odd, channel uses. Note that the channel is effectively a binary symmetric channel with crossover probability $p \approx \frac{c}{\rho}\exp(-\rho/2)$. The probability of error when using a repetition code of length t, t odd, is*

$$P_e(t, p) = \binom{m}{\frac{t+1}{2}} p^{\frac{t+1}{2}}(1-p)^{\frac{t-1}{2}}$$

Therefore, as $p \to 0$ $P_e(t, p) \approx c'p^{\frac{t+1}{2}}$ where c' is some other constant. Therefore,

$$\mathrm{DIV}_t(\mathcal{C}_2) = \frac{t+1}{2}$$

We see that hard decision decoding at the output reduces DIV *to $(t+1)/2$ as opposed to* DIV *of t in Example 1 that has soft decision decoding at the output.*

Example 6 (Unknown fading channel, [10]) *Let's consider the channel in Example 2. The parameters in this channel are t the number of channel uses, M the number of transmit antennas and N the number of receive antennas. We assume that H is unknown to both the transmitter and the receiver and $T \geq 2M$ so that we can apply the results of [10]. In this example, $t = T$ where T is the fading block length. From [10], $C(\rho, M, N, T)$ behaves as $M(T - M)\log\rho$ as $\rho \to \infty$. Also $C_H(\rho, 1, 1, T)$ behaves as $(T-1)\log\rho$ irrespective of the value of $H(1,1)$. Therefore, $\sup_{H(1,1)} C(\rho, 1, 1, 1)$ behaves as $\frac{T-1}{T}\log\rho$ as $\rho \to \infty$. This implies that DOF for this channel with M transmit antennas, N receive antennas and T channel uses is $M(T-M)\frac{T}{T-1}$. Note that this is different from $\mathrm{DOF} = M(T-M)$ that was obtained in [10].*

Intuitively, we would expect the DIV $\geq$ DOF for a channel as DOF in a channel can be used to transmit redundant information (repetition coding) thus adding to "natural" redundancy (multiple receive antennas) in the channel. This intuition however, breaks down for some examples as with the case of multiple antenna channels operating in a coherent Rayleigh fading environment where we have seen the diversity is $\min(M, t)N$ [8] whereas the degrees of freedom is $\min(M, N)t$ [4, 10].

4. DISCUSSION

The concept of degrees of freedom has been well studied in the optical community [1, 5, 6]. The author in [6] calls DOF the number of independent parameters needed to represent the output of the channel. It has also been referred to as the number of independent parameters that can be extracted from the transmitted signal after it has passed through the channel. It is possible by borrowing from the optical community to give a definition that is consistent with the implicit definition prevalent in the multiple-antenna literature.

5. CONCLUSIONS AND FUTURE WORK

We have introduced rigorous definitions for diversity and degrees of freedom in a more general setting than MIMO communication system operating in a fading environment. We have illustrated how these definitions can be used to evaluate DIV and DOF through different examples.

Intuitively diversity has been considered as the redundancy in the system and degrees of freedom as the number of independent channels available for communication. Any rigorized definitions that agree with the intuition may not be easy to translate to actual figures of merit, information capacity and probability of error. In that respect the definitions in Section 2 are the first step in tackling this problem.

For future work, we address the following issues:

- A rigorous definition for DIV that agrees with intuition can be formulated along the lines of (3). If possible the definition can be formulated in such a way so as to be able to derive the minimum probability of error directly from the definition.
- A rigorous definition for DOF that agrees with intuition can be formulated along the lines of (4) and a relationship between this definition and information capacity can be derived.
- Is it possible that the definitions based on intuition are such that we obtain DIV to be always greater than DOF? We expect the answer to be "yes" because by transmitting redundant information over the independent channels of communication we can obtain as many independent copies of the transmitted signal as DOF of the channel.

6. REFERENCES

[1] I. J. Cox and C. J. R. Sheppard, "Information capacity and resolution in an optical system," *Journal of the Optical Society of America A*, vol. 3, no. 8, pp. 1152–1158, Aug. 1986.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.

[3] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[4] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading channel," *IEEE Trans. on Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan 1999.

[5] D. A. B. Miller, "Communicating with waves between volumes: Evaluating orthogonal spatial channels and limits on coupling strengths," *Applied Optics*, vol. 39, no. 11, pp. 1681–1699, Apr 2000.

[6] R. Piestun and D. A. B. Miller, "Electromagnetic degrees of freedom of an optical system," *Journal of the Optical Society of America A*, vol. 17, no. 5, pp. 892–902, May 2000.

[7] J. G. Proakis and M. Salehi, *Communication Systems Engineering*, Prentice-Hall and Englewood Cliffs and NJ, 1994.

[8] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Communications*, vol. 44, no. 2, pp. 744–765, March 1998.

[9] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, Nov/Dec 1999.

[10] L. Zheng and D. N. C. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the non-coherent multi-antenna channel," *To appear in IEEE Trans. on Information Theory*.