

Capacity-Approaching Rate Function for Layered Multi-Antenna Architectures

Angel Lozano*

January 25, 2012

Abstract

The simultaneous use of multiple transmit and receive antennas can unleash very large capacity increases in rich multipath environments. Although such capacities can be approached by layered multi-antenna architectures with per-antenna rate control, the need for short-term feedback arises as a potential impediment, in particular as the number of antennas—and thus the number of rates to be controlled—increases. What we show, however, is that the need for short-term feedback in fact vanishes as the number of antennas and/or the diversity order increases. Specifically, the rate supported by each transmit antenna becomes deterministic and a sole function of the signal-to-noise, the ratio of transmit and receive antennas, and the decoding order, all of which are either fixed or slowly varying. More generally, we illustrate—through this specific derivation—the relevance of some established random CDMA results to the single-user multi-antenna problem.

Keywords: Channel Capacity, Fading Channels, Multiple-Antenna Information Theory, Multiuser Detection, Antenna Diversity, Adaptive Arrays.

*Angel Lozano is with the Wireless Communication Research Department at Bell Laboratories (Lucent Technologies), Holmdel, NJ07733, USA.

I Introduction

IT has been shown that, with the simultaneous use of multiple transmit and receive antennas, very large capacity increases can be unleashed in rich multipath environments [1]–[4]. As a result, the deployment of arrays at both base stations and terminals appears as an attractive scenario for the evolution of wireless data access.

Analyzing the capacity and performance of such multi-antenna structures typically involves dealing with channel matrices whose entries can be modeled stochastically. Consequently, it is only natural to invoke results from random matrix theory as has been successfully done in related fields of communication theory such as code-division multiple access (CDMA) with random spreading [5]–[8]. This paper shows how, because of the isomorphism of both problems, many of these results apply also to layered multi-antenna architectures.

Focusing on mobile systems, we consider only schemes wherein the transmitter does not have access to the instantaneous state of the channel. Only information that varies slowly with respect to the fading rate is available to the transmitter.

II Layered Multi-Antenna Architectures

a. Multi-Antenna Channel Capacity

With M transmit and N receive antennas, the normalized channel responses from every transmit to every receive antenna can be assembled into an $N \times M$ matrix \mathbf{H} whose unit-variance elements we assume independent and identically distributed [9]–[10]. Furthermore, this channel matrix is presumed to be perfectly known at the receiver [11]–[12].¹ With the channel realization unknown to the transmitter, the mutual information is maximized by transmitting a Gaussian signal with spatial covariance

$$\Phi = \frac{P}{M} \mathbf{I}_M \quad (1)$$

given a total radiated power P [3]. Note that \mathbf{I}_M symbolizes the identity matrix of size M . Assuming flat fading, the corresponding capacity—conditioned on the channel—in the presence of additive white Gaussian noise (AWGN) can be expressed as

$$C(\mathbf{H}, \text{SNR}) = \log_2 \det(\mathbf{I}_N + \frac{\text{SNR}}{M} \mathbf{H} \mathbf{H}^\dagger). \quad (2)$$

¹The capacity penalty associated with channel estimation is small as long as the coherence time of the channel sufficiently large [12].

with SNR the average signal-to-noise ratio per receive antenna [1].

Since the channel is time varying in nature, the capacity fluctuates with it. With a sufficiently long coding horizon, it is possible to code over the short-term fluctuations and approach the ergodic capacity, which is simply the average of (2) over the distribution of \mathbf{H} [3]. It is also possible to code over the short-term randomness in the frequency domain, whereby the ergodic capacity is approached as the signal bandwidth increases [13]. If it is not possible to code over the short-term channel variations, one must resort to the idea of outage capacity, wherein the capacity itself is regarded as a random variable that fluctuates with the channel [14].

Fundamentally, the single-user multi-antenna problem is akin to that of a synchronous CDMA channel with random spreading [8]. The dimensionality of such a channel, represented by the spreading factor, corresponds in our case to the number of receive antennas, while the number of users corresponds to the number of transmit antennas. There are, nonetheless, some differences:

- The spatial signatures for the various transmit antennas are not chosen by the system designer, as in CDMA, but rather imposed by nature. Moreover, their distribution depends on the type of fading being experienced.
- The transmit power constraint, which applies to each individual user in CDMA, becomes a constraint on the sum of powers in the multi-antenna case. This typically leads to a scaled definition of the SNR.
- The transmit antennas are co-located and part of a single transceiver whereas, in CDMA, the multiple users are geographically dispersed. Hence, joint coding of the transmit signals is feasible in our problem, but not in CDMA.

Despite these differences, many of the results on random CDMA are very relevant to the multi-antenna problem considered here, where \mathbf{H} is unknown to the multi-antenna transmitter. This is particularly true of asymptotic results such as those presented in [5].² This paper illustrates this relevance, which many in the multi-antenna community appear to be unaware of.

b. Why Layered Architectures ?

While the covariance in (1) implies that the M transmit antennas should radiate uncorrelated signals, these signals can be jointly encoded. This has motivated a blossoming

²The singular values of \mathbf{H} converge (almost surely) to the same asymptotic function regardless of the distribution of its entries.

interest in the design of space-time codes [15]–[?]. However, the complexity of joint detection explodes as the number of transmit antennas grows larger and thus alternative approaches become are of interest.

One such approach is that of layered architectures, wherein each transmit antenna radiates a separately encoded signal. At the receiver, these signals can be successively decoded and their interference cancelled [17]–[19]. The complexity of these architectures increases more gracefully with the number of antennas. Furthermore, they can capitalize on existing scalar coding formats. However, since (i) a successive decoding process favors antennas differently depending on the decoding order and (ii) each transmit antenna encounters a distinct propagation channel, layered architectures can only approach capacity if the radiated signals are cycled around the transmit antennas so that each signal gets exposed, in turn, to every transmit antenna [20]–[21]. This, however, implicitly assumes that the M transmit antennas are constrained to operate at the same data rate. Instead, if the receiver can inform the transmitter of the instantaneous rate that can be supported by each transmit antenna depending on the channel and on the detection order, the capacity in (2) can be approached by a layered architecture without the need to cycle [22]. This possibility, though, is complicated by the need for short-term feedback at the fading rate [23]. However, as we shall show in the remainder, such feedback requirements in fact vanish as the number of antennas and/or diversity order increase.

c. Layered Architectures with per-Antenna Rate Control

It was first shown in [24] (see also [5]) that a minimum mean-square error (MMSE) receiver with successive decoding and interference cancellation can approach the capacity of the CDMA channel. This optimality carries over to layered multi-antenna architectures as long as the rate of each transmit antenna can be appropriately adjusted [22]. Moreover, this is true irrespective of the decoding order. The rate of each individual antenna depends on the decoding order, but the aggregate capacity does not. Hence, we assume—without loss of generality—that the antennas are decoded in accordance with some arbitrary index m .

Let us define \mathbf{h}_m as the m -th column of \mathbf{H} . Let us further define $\mathbf{H}_m = [\mathbf{h}_{m+1} \ \mathbf{h}_{m+2} \ \dots \ \mathbf{h}_M]$. The rate supported by antenna m is given [24, 22] by

$$C_m = \log_2(1 + \mathbf{h}_m^\dagger (\mathbf{H}_m \mathbf{H}_m^\dagger + \frac{M}{\text{SNR}} \mathbf{I}_N)^{-1} \mathbf{h}_m) \quad (3)$$

so that

$$\sum_{m=1}^M C_m = C(\mathbf{H}, \text{SNR}) = \log_2 \det(\mathbf{I}_N + \frac{\text{SNR}}{M} \mathbf{H} \mathbf{H}^\dagger). \quad (4)$$

Eq. (3) can be rewritten as

$$\begin{aligned}
C_m &= \log_2 \det(\mathbf{I}_N + \mathbf{h}_m \mathbf{h}_m^\dagger (\mathbf{H}_m \mathbf{H}_m^\dagger + \frac{M}{\text{SNR}} \mathbf{I}_N)^{-1}) \\
&= \log_2 \det(\frac{M}{\text{SNR}} \mathbf{I}_N + \mathbf{h}_m \mathbf{h}_m^\dagger + \mathbf{H}_m \mathbf{H}_m^\dagger) - \log_2 \det(\frac{M}{\text{SNR}} \mathbf{I}_N + \mathbf{H}_m \mathbf{H}_m^\dagger) \\
&= \log_2 \det(\frac{M}{\text{SNR}} \mathbf{I}_N + [\mathbf{h}_m \ \mathbf{H}_m][\mathbf{h}_m \ \mathbf{H}_m]^\dagger) - \log_2 \det(\frac{M}{\text{SNR}} \mathbf{I}_N + \mathbf{H}_m \mathbf{H}_m^\dagger) \\
&= \log_2 \det(\mathbf{I}_N + \frac{\text{SNR}}{M} \mathbf{H}_{m-1} \mathbf{H}_{m-1}^\dagger) - \log_2 \det(\mathbf{I}_N + \frac{\text{SNR}}{M} \mathbf{H}_m \mathbf{H}_m^\dagger)
\end{aligned} \tag{5}$$

which corresponds to

$$C_m = C(\mathbf{H}_{m-1}, \frac{M-m+1}{M} \text{SNR}) - C(\mathbf{H}_m, \frac{M-m}{M} \text{SNR}) \tag{6}$$

indicating that the rate supported by the m -th antenna is given by the difference between the capacity of transmit antennas m to M and the capacity of transmit antennas $m+1$ to M .

III Asymptotic Analysis

As the number of antennas increases, the outage capacity *hardens* and becomes deterministic. This implies that, asymptotically, the outage and ergodic capacities coincide. Mathematically, this hardening process stems from the (almost sure) convergence of the singular values of \mathbf{H} [25]. Let us define the ratio of transmit and receive antennas as

$$\beta \stackrel{\text{def}}{=} \frac{M}{N} \tag{7}$$

and the capacity per receive antenna as

$$\mathcal{C} \stackrel{\text{def}}{=} \frac{C}{N}. \tag{8}$$

As the number of antennas is driven to infinity, \mathcal{C} converges almost surely [5] to

$$\begin{aligned}
\mathcal{C}^\infty(\beta, \text{SNR}) &= \log_2[1 + \text{SNR} - \mathcal{F}(\beta, \frac{\text{SNR}}{\beta})] \\
&+ \beta \log_2[1 + \frac{\text{SNR}}{\beta} - \mathcal{F}(\beta, \frac{\text{SNR}}{\beta})] - \beta \frac{\log_2 e}{\text{SNR}} \mathcal{F}(\beta, \frac{\text{SNR}}{\beta})
\end{aligned} \tag{9}$$

given the auxiliary function

$$\mathcal{F}(x, y) \stackrel{\text{def}}{=} \frac{1}{4} [\sqrt{1 + y(1 + \sqrt{x})^2} - \sqrt{1 + y(1 - \sqrt{x})^2}]^2. \tag{10}$$

Notice that \mathcal{C}^∞ is solely a function of β and SNR. Furthermore, despite being an asymptotic result, it yields an extremely accurate approximation to the ergodic capacity even when the number of antennas is small [26].

Defining a normalized version of the decoding index as

$$b \stackrel{\text{def}}{=} \frac{m}{N} \quad (11)$$

and defining also

$$\Delta \stackrel{\text{def}}{=} \frac{1}{N} \quad (12)$$

the rate supported by the m -th antenna in (6) can be rearranged as

$$\begin{aligned} C_m &= \frac{\mathcal{C}(\mathbf{H}_{m-1}, \frac{M-m+1}{M} \text{SNR}) - \mathcal{C}(\mathbf{H}_m, \frac{M-m}{M} \text{SNR})}{\Delta} \\ &= \frac{\mathcal{C}(\mathbf{H}_{m-1}, \frac{\beta-b+\Delta}{\beta} \text{SNR}) - \mathcal{C}(\mathbf{H}_m, \frac{\beta-b}{\beta} \text{SNR})}{\Delta}. \end{aligned} \quad (13)$$

As the number of antennas grows to infinity, (13) converges almost surely to

$$C_m = \frac{\mathcal{C}^\infty(\beta - b + \Delta, \frac{\beta-b+\Delta}{\beta} \text{SNR}) - \mathcal{C}^\infty(\beta - b, \frac{\beta-b}{\beta} \text{SNR})}{\Delta}. \quad (14)$$

At the same time, the normalized index b becomes continuous within $(0, \beta]$ and thus the rates supported by the transmit antennas become also a continuous function of that index, which we label C_b^∞ . Moreover, since $\Delta \rightarrow 0$, the difference in (14) becomes a derivative and hence C_b^∞ can be asymptotically expressed as

$$C_b^\infty = \frac{d}{dx} \mathcal{C}^\infty(x, x \frac{\text{SNR}}{\beta}) \big|_{x=\beta-b} \quad (15)$$

the converse of which is

$$\int_0^\beta C_b^\infty db = \mathcal{C}^\infty(\beta, \text{SNR}) \quad (16)$$

indicating how the asymptotic capacity per receive antenna can be obtained by integration over the asymptotic rate function C_b^∞ . This is the same result found in [5] for random CDMA.

If the M transmit antennas are constrained to operate—with no cycling—at the same rate, such rate must be the smallest supported by any of the antennas, that is, the one corresponding to $m=1$ and, in the limit, to $b=0$.³ Consequently, the asymptotic capacity per receive antenna becomes βC_0^∞ and the advantage of enabling different rates at the transmit antennas—or, alternatively, of cycling—can be quantified, in terms of capacity per receive antenna, as the difference between \mathcal{C}^∞ and βC_0^∞ .

³The ordering process advocated in [17] becomes asymptotically irrelevant.

IV Results

The derivative in (15) can be calculated from (9) to yield

$$C_b^\infty = \log_2[1 + \frac{\text{SNR}}{\beta} - \mathcal{F}(\beta - b, \frac{\text{SNR}}{\beta})] \quad (17)$$

which coincides with the asymptotic MMSE capacity derived in [5] for CDMA.⁴

The asymptotic convergence of the transmit rates to the asymptotic expression in (17) is verified in Fig. 1. A sequence of three random realizations of the rates supported by 16, 32 and 64 antennas respectively are plotted over the asymptotic function for $\beta=1$ and SNR=10 dB. The progressive convergence is evident as the number of antennas increases.

The ergodic convergence, on the other hand, is validated in Fig. 2. The average rates in a Rayleigh fading channel, with only 4 transmit and 4 receive antennas, are shown to match exactly those predicted by the asymptotic analysis, again at SNR=10 dB.

Finally, using (9) and (17), the advantage of per-antenna rate control in the presence of AWGN is found to be

$$\mathcal{C}^\infty - \beta C_0^\infty = \log_2[1 + \text{SNR} - \mathcal{F}(\beta, \frac{\text{SNR}}{\beta})] - \beta \frac{\log_2 e}{\text{SNR}} \mathcal{F}(\beta, \frac{\text{SNR}}{\beta}). \quad (18)$$

This advantage vanishes for diminishing β while, for growing β , it becomes

$$\lim_{\beta \rightarrow \infty} [\mathcal{C}^\infty - \beta C_0^\infty] = \log_2(1 + \text{SNR}) - \frac{\text{SNR}}{1 + \text{SNR}} \log_2 e \quad (19)$$

per receive antenna. For $\beta=1$, on the other hand, it approaches

$$\mathcal{C}^\infty - C_0^\infty \approx \log_2(\frac{\sqrt{\text{SNR}}}{e}) \quad (20)$$

as the SNR increases.

V Summary

While it is known that layered multi-antenna architectures (MMSE with successive detection and interference cancellation) can approach capacity through per-antenna rate control, the need for short-term feedback arises as a serious impediment, in particular

⁴While the capacity varies if the noise is Gaussian but spatially non-white (see [?] for the random CDMA asymptotic capacity in the presence of colored noise and [26] for its multi-antenna counterpart), the main result in (15) holds and thus the rate function can still be derived from \mathcal{C}^∞ .

as the number of antennas (and thus the number of distinct rates) increases. However, the need for short-term feedback in fact vanishes as the number of antennas and/or the diversity order increases. More specifically, the rates supported by the transmit antennas converge to a deterministic function that depends solely on the SNR, the antenna ratio and the decoding order, all of which are either fixed or slowly varying.

References

- [1] G. J. Foschini and M. J. Gans, “On the limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, pp. 315–335, 1998.
- [2] G. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communications,” *IEEE Trans. Commun.*, vol. 46, no. 3, 1998.
- [3] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Eur. Trans. Telecom*, vol. 10, pp. 585–595, Nov. 1999.
- [4] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela, “Lifting the limits in high-speed wireless data access using antenna arrays,” *IEEE Commun. Magazine*, vol. 39, no. 9, Sept. 2001.
- [5] S. Verdú and S. Shamai, “Spectral efficiency of CDMA with random spreading,” *IEEE Trans. on Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
- [6] D. N. C. Tse and S. Hanly, “Linear multiuser receivers: Effective interference, effective bandwidth and user capacity,” *IEEE Trans. on Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.
- [7] P. Rapajic and D. Popescu, “Information capacity of a random signature multiple-input multiple-output channel,” *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1245–1248, Aug. 2000.
- [8] S. Shamai and Sergio Verdú, “The effect of frequency-flat fading on the spectral efficiency of CDMA,” *IEEE Trans. on Inform. Theory*, vol. 47, no. 5, May 2001.
- [9] D. Chizhik, F. R. Farrokhi, J. Ling, and A. Lozano, “Effect of antenna separation on the capacity of BLAST in correlated channels,” *IEEE Commun. Letters*, vol. 4, no. 11, Nov. 2000.
- [10] C. C. Martin, J. H. Winters, and N. R. Sollenberger, “Multiple-input multiple-output (MIMO) radio channel measurements,” *Proc. of IEEE Vehic. Techn. Conf. (VTC’2000)*, Boston, MA, USA, Sept. 2000.

- [11] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high capacity space-time wireless," *Proc. of 37th Annual Allerton Conf. on Communication, Control, Computing, Monticello, IL*, Sept. 1999.
- [12] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *Submitted to IEEE Trans. on Inform. Theory*, 2000.
- [13] H. Bolcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–2234, Feb. 2002.
- [14] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communication aspects," *IEEE Trans. on Informat. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [15] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [16] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
- [17] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolnianski, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.
- [18] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communications architecture," *Electronic Letters*, Nov. 1998.
- [19] A. Lozano and C. B. Papadias, "Space-time receiver for wideband BLAST in rich-scattering wireless channels," *Proc. of IEEE Vehic. Technol. Conf. (VTC'2000)*, May 2000.
- [20] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.
- [21] S. L. Ariyavisitakul, "Turbo space-time processing to improve wireless channel capacity," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1347–1358, Aug. 2000.

- [22] S. T. Chung, A. Lozano, and H. C. Huang, "Approaching eigenmode BLAST channel capacity using V-BLAST with rate and power feedback," *IEEE Vehic. Techn. Conf. (VTC'2001 Fall)*, Atlantic City, NJ, Oct. 2001.
- [23] P. Sebastian, H. Sampath, and A. Paulraj, "Adaptive modulation for multiple antenna systems," *Proc. of the Internat. Conference in Commun. (ICC'01)*, Helsinki, Finland, June 2001.
- [24] M. K. Varanasi and T. Guess, "Optimum decision-feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," *Asilomar Conference*, pp. 1405–1409, 1998.
- [25] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *Journal of Multivariate Analysis*, vol. 12, pp. 1–38, 1982.
- [26] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. on Inform. Theory*, vol. 48, Dec. 2002.

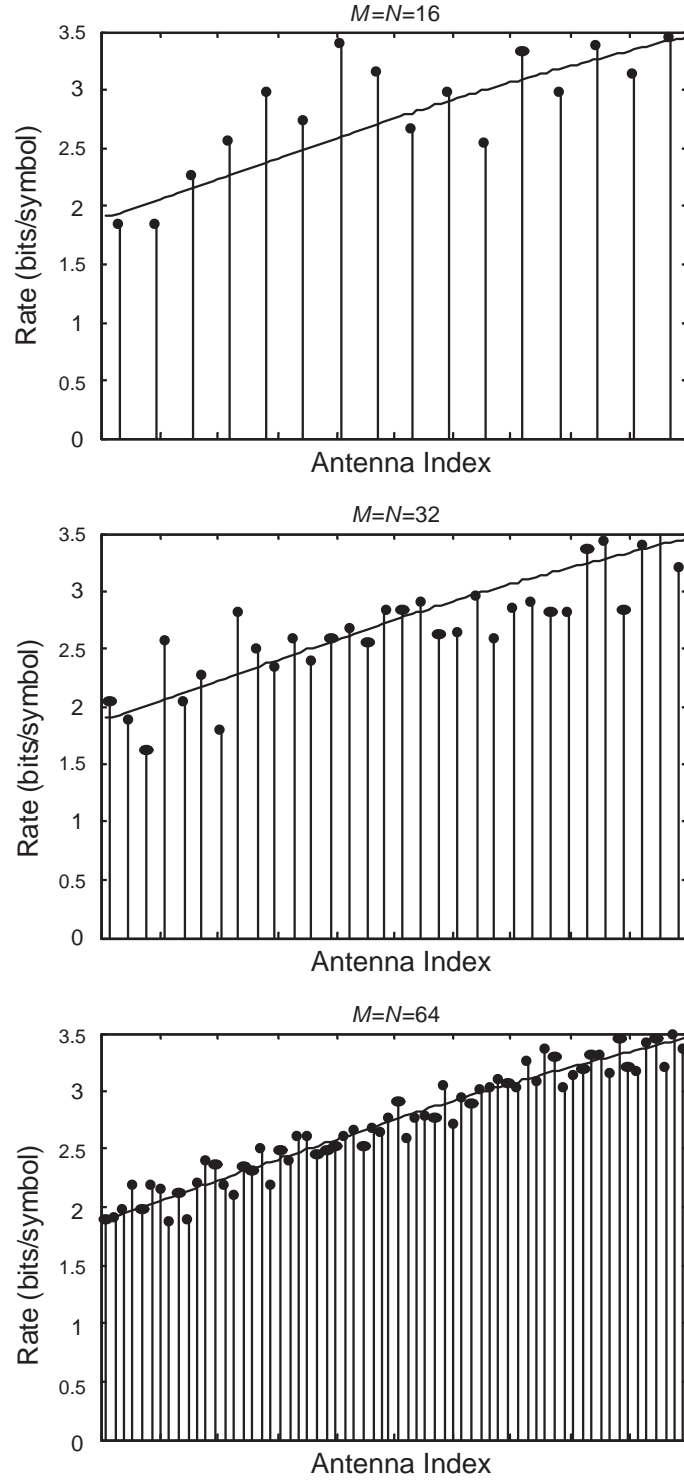


Figure 1: Asymptotic convergence: rate vs. transmit antenna index for $\beta=1$ and SNR=10 dB. Solid line corresponds to asymptotic function. Circles indicate the supported rates (numerically computed) for individual random channel realizations with 16, 32 and 64 antennas.

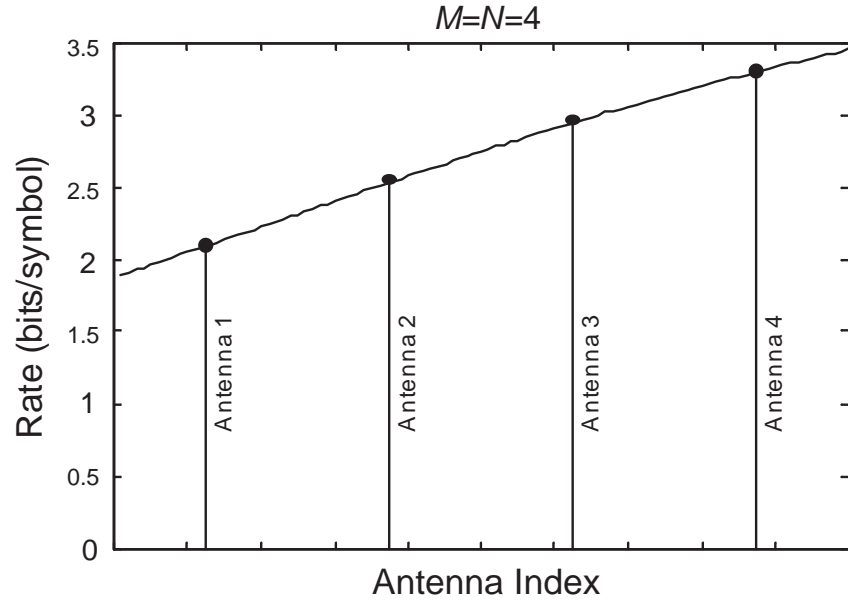


Figure 2: Ergodic convergence: rate vs. transmit antenna index with $\beta=1$ and SNR=10 dB. Solid line corresponds to asymptotic function. Circles indicate average rates (numerically computed) with 4 antennas and Rayleigh fading.