On the Capacity of a Binary MIMO Channel with Random Interference

Răzvan Cristescu
Center for Mathematics of Information
California Institute of Technology
Pasadena, CA 91125
razvanc@caltech.edu

Michelle Effros
Data Compression Laboratory
California Institute of Technology
Pasadena, CA 91125
effros@caltech.edu

Muriel Médard
Dept. of Elec. Eng. and Comp. Science
Massachusetts Institute of Technology
Cambridge, MA 02139
medard@mit.edu

Abstract

We study the capacity of a binary multiple input multiple output (MIMO) channel with interference. Interference is modelled as a non-linear relation between the input and the output words of the system. Namely, we assume that the fading coefficients in the channel matrix are binary and drawn with i.i.d. Bernoulli probabilities, and any element of the output word is given by the minimum value (the AND function) of the input elements connected to that output. We find the expected number of output codewords for a given dimension of the MIMO system and interference level. Using this result and Jensen's inequality, we derive an upper bound on the capacity of this system. We show by numerical simulations that the capacity upper bound is tight for most of the interference levels.

1 Introduction

1.1 Motivation

Multiple input multiple output systems (MIMO) have been shown to increase the capacity of various channels [2, 5, 7]. The computation of the capacity as a function of the time-varying channel parameters is in general a hard problem [4], even in the case of a single-user MIMO channel. Most research in the field has concentrated on the case where the channel coefficients are modelled as complex jointly Gaussian random variables. We study a simple MIMO system where the channel is modeled by a non-linear relation between the input and the output terminals.

Consider a MIMO system where the input and output words are binary, namely the input word $\mathbf{x} \in \mathcal{B}^M$ and the output word $\mathbf{y} \in \mathcal{B}^N$, where $\mathcal{B} = \{0, 1\}$. The relationship between the channel inputs and outputs is determined by a randomly chosen interference

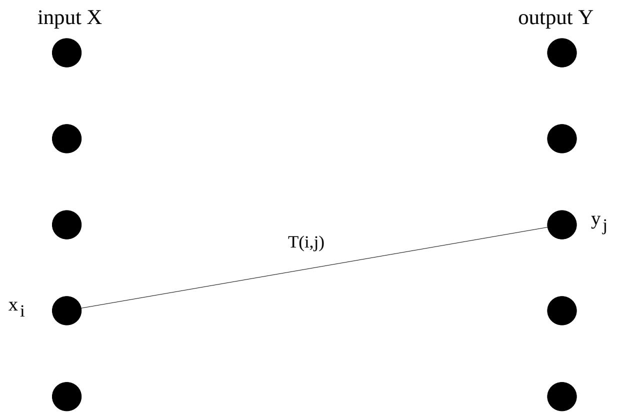

Figure 1: A MIMO system. The fading coefficient between input x_i and output y_j is denoted by $T(i, j) \in \{0, 1\}$.

channel. More precisely, the channel is modelled as a channel matrix $T \in \mathcal{B}[M \times N]$, in which the elements of the matrix $T(i, j)$ represent the fading coefficients between the input and output elements. Each matrix coefficient indicates whether a certain input terminal is connected (matrix coefficient is `1`) or not (matrix coefficient is `0`) to a certain output terminal. We call this system a *binary* MIMO system. Given a fixed channel matrix T, the relationship between the channel input and the output is deterministic. Each channel output is the minimum over all connected inputs.

A recurrent problem in the study of MIMO systems is interference. For most channel models, little is known about the effect of various levels of interference on the performance of the channel. We consider a system with an interference model that spans all the intensity levels. Namely, in our model the interference level is modelled by the random number of active links that are present between the input and the output.

1.2 Related Work

The extensive study of MIMO systems has been motivated by the pioneering work of [2, 5, 7]. An outline of the state of the art in the field is provided in [4]. Gaussian random matrices have been extensively studied (e.g. [6]) as a model for the channel matrix.

The work in this paper is most related to the study of the rank of random matrices over finite fields [1, 3]. However, due to the non-linearity of our problem, we will see that our setting is not solved by this theory.

1.3 Main Contributions and Outline of the Paper

We define a MIMO system with binary input and output words for which the interference level is modelled by the number of connections that are present between the input and the output terminals (the number of `1`-s in the random channel matrix). We describe the capacity of this system, and use Jensen's inequality to upper bound the capacity by the logarithm of the expected number of output words, for a given dimension of the system and interference level.

Section 2 introduces the binary MIMO system we study in this paper and defines the capacity of this system. In Section 3 we find the expected number of output words, from which we derive an upper bound on the capacity. Section 4 illustrates with numerical

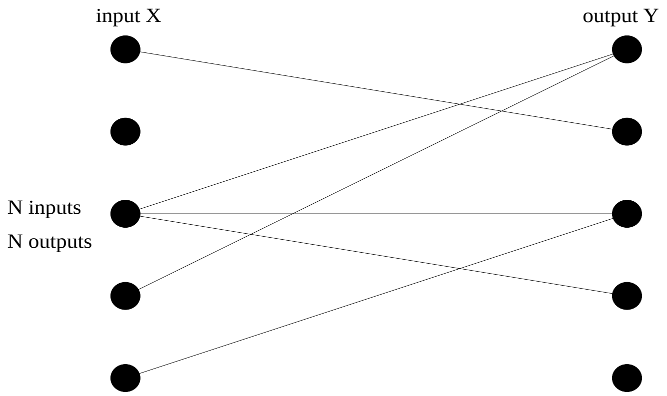

Figure 2: A binary MIMO system.

simulations the correctness of our upper bound on the expected capacity, and the fact that it is reasonably tight for most levels of interference. We conclude with Section 5.

2 Problem Setting

2.1 Channel Matrix

Consider a binary multiple input, multiple output (MIMO) system. For the sake of simplicity, we will consider the problem when the number of inputs and outputs are equal $M = N$, that is the channel matrix T_N is a square matrix of dimension N. The entries of the channel matrix T_N belong to the binary set $\mathcal{B} = \{\mathtt{0}, \mathtt{1}\}$: $T_N \in \mathcal{B}[N \times N]$. These entries specify the connectivity pattern of the channel. Namely, if the input terminal x_i is connected to the output terminal y_j, then $T_N(i, j) = \mathtt{1}$, and otherwise $T_N(i, j) = \mathtt{0}$.

We consider the case of a random channel with interference. We model the randomness of the channel fading as follows. The channel between the input bit x_i and the output bit y_j is ON with probability p (see Figure 2):

$$T_N(i,j) = \begin{cases} \mathtt{1}, & \text{w. p.} \quad p \\ \mathtt{0}, & \text{w. p.} \quad 1-p \end{cases}$$

As p varies between 0 and 1, we explore the full range from no connections and no interference to total connection and complete interference. For simplicity, we assume that channel connections are chosen independently.

2.2 Interference Model

We model interference by a scenario where the output is given by the minimum value of the connected inputs. Namely, we assume that for a given input word $\mathbf{x} = (x_1, \ldots, x_N)$, the output word $\mathbf{y} = (y_1, \ldots, y_N)$ is determined as follows:

$$y_j = \prod_{i=1}^{N} x_i^{T_N(i,j)}, \quad j = 1 \ldots N.$$

If the inputs connected to a given output are identical, then that output is received without change. If the inputs differ, then interference occur, and a $\mathtt{0}$ is received. If there are no connected input terminals, then by convention the output takes the value 1.

Note that the dual model, when interference is produced by an input bit 1 (namely, if at least an input 1 is connected to the output then that output is 1, and otherwise it is 0), yields the same capacity as that achieved by the above stated model.

2.3 Expected Capacity

The goal of this work is to find the expected capacity of the binary MIMO system with interference introduced in the previous sections. Since the channel is deterministic, the capacity equals the maximal entropy at the channel output. As a result, the capacity for a given transition matrix T_N equals the logarithm (in base two) of the maximal number $M(T_N)$ of distinct outputs on the given channel. We here calculate the expected capacity $C(p, N)$, where the expectation is taken over the realizations of the random channel matrix:

$$C(p, N) = E_{T_N}\left[\log M(T_N)\right].$$

For simplicity of notation, we will omit the subscript in the expectation. We illustrate the procedure for finding the capacity with the following example.

Example 1 *We refer to the example in Figure 3, which illustrates a realization of the channel matrix for a system with two inputs and two outputs. The channel matrix in this example is:*

$$T_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

The output words (the columns on the right-hand side) corresponding to the input words (the columns on the left-hand side) are as follows:

x_1	0	0	1	1	$\rightarrow$	0	0	1	1	y_1
x_2	0	1	0	1		0	0	0	1	y_2

For this example, there are $M(T_2) = 3$ possible output words when the input word spans the set $\mathcal{B}^2$. The capacity for channel T_2 equals the logarithm of the number of output words $\log M(T_2) = \log 3$. The expected capacity is $E\left[\log M(T_2)\right]$ over the realizations of the connection matrix T_2. Note that for a given N and $p \in (0, 1)$, the realization of the matrix T_N can take any value in the whole set $\mathcal{B}^{N^2}$. The probability that matrix T_N takes on a particular value depends on p.

At first sight, it seems possible that the solution of this problem can be found using the theory of random matrices over finite fields [3]. Namely, one could attempt to write the system in a matrix form as $\mathbf{y} = T_N\mathbf{x}$, and find the solution for the expected capacity as the expected value of the matrix rank. However, it is easy to see that our problem is non-linear, since the outputs are not a linear function of the inputs (namely, in our setting, the function that determines the outputs is AND rather than XOR). For instance, in the example in Figure 3, the logarithm of the number of output words is $\log 3$, and different from the rank of the channel matrix, which is 2.

Note that finding the capacity is simple for the two extreme cases $p = 0$ and $p = 1$ (see Figure 5). On one hand, if $p = 0$, then with probability 1 there is no connection between the input and the output, and thus the only possible output word, regardless of the input

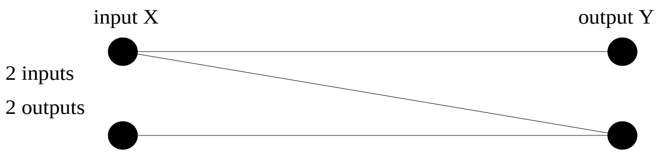

Figure 3: An example with two inputs and two outputs.

word, is $\mathbf{y} = (\mathtt{1}, \ldots, \mathtt{1})^T$. Thus, the capacity is $C(0, N) = E[\log 1] = 0$. On the other hand, if $p = 1$, then all the inputs are connected to all the outputs with probability 1, and thus, the only two possible output words are $\mathbf{y} = (\mathtt{0}, \ldots, \mathtt{0})^T$ corresponding to any input word except $\mathbf{x} = (\mathtt{1}, \ldots, \mathtt{1})^T$, and $\mathbf{y} = (\mathtt{1}, \ldots, \mathtt{1})^T$ corresponding to $\mathbf{x} = (\mathtt{1}, \ldots, \mathtt{1})^T$. Thus, the capacity in this case is $C(1, N) = E[\log 2] = 1$.

For $p \in (0, 1)$, finding the capacity is not trivial. Intuitively, when p is small, few connections are made between the input and the output, and thus the chances of having many possible different words at the output increase with p. Passing past a certain value p_0 though, the number of words for $p > p_0$ starts to decrease due to the increased interference. In general, for a given N, there is an optimal probability $p_0(N)$ that maximizes the capacity $C(p, N)$.

3 Determining an Upper Bound on the Capacity

In this section, we find the expected number of output words $E[M(T_N)]$; $\log E[M(T_N)]$ provides an upper bound on the capacity, since by Jensen's inequality:

$$C(p, N) = E[\log M(T_N)] \leq \log E[M(T_N)].$$

We compute the expectation by considering separately each possible output word $\mathbf{y} \in \mathcal{B}^N$. Note that

$$E[M(T_N)] = E\left[\sum_{\mathbf{y}} I(T_N, \mathbf{y})\right]$$

where $I(T_N, \mathbf{y})$ is 1 if there exists an input $\mathbf{x}$ to channel T_N that yields output $\mathbf{y}$, and 0 otherwise. We thus need to find the expected probability of appearance $P_{\mathbf{y}}(p, N) = E[I(T_N, \mathbf{y})]$ of the output word $\mathbf{y}$, and then obtain the expected number of output words by computing:

$$E[M(T_N)] = \sum_{\mathbf{y}} P_{\mathbf{y}}(p, N). \tag{1}$$

Note that another possible approach would be to attempt to compute the number of output words $M(T_N)$ for each realization T_N of the channel matrix, and then to sum over the possible realizations of the channel matrix, while weighting by the corresponding probability of appearance P_{T_N} of that realization:

$$E[M(T_N)] = \sum_{T_N} P_{T_N} \cdot M(T_N). \tag{2}$$

The reason to choose our approach is that we are able to compute formula (1) by calculating each of its 2^N terms in a structured manner. On the contrary, there are 2^{N^2} terms in formula (2), and there is no clear way to derive a structured strategy to compute those terms. Namely, even the change of a single element in a realization of the channel matrix T_N may radically change the set of output words, which renders the computation of (2) difficult. As we will see, studying instead the probability of appearance of output words simplifies the problem.

The main result of this paper is stated in the following theorem:

Theorem 1 *The expected number of output codewords for an N-dimensional binary MIMO system in which each element of the channel matrix T_N is generated i.i.d. as $\mathtt{1}$ with probability p and $\mathtt{0}$ with probability $1-p$, is:*

$$E\left[M(T_N)\right] = \sum_{k=0}^{N} \left(\begin{array}{c} N \\ k \end{array} \right) \sum_{s=0}^{N} \left(\begin{array}{c} N \\ s \end{array} \right) (1-(1-p)^k)^s (1-p)^{k(N-s)} (1-(1-p)^{N-s})^{N-k}. \quad (3)$$

Proof: First, we separate the terms in (1) into $N+1$ terms, grouping together all output words that contain exactly k $\mathtt{1}$-s and $N-k$ $\mathtt{0}$-s, where $k = 0 \ldots N$. Note that $P_{\mathbf{y}}(p, N)$ is constant for all $\mathbf{y}$ with k $\mathtt{1}$-s and $N-k$ $\mathtt{0}$-s, due to the symmetry of the problem. We denote the probability of each such string by P_k and, without loss of generality, compute P_k by analyzing the probability of the string $\mathbf{y} = (\mathtt{1}, \ldots, \mathtt{1}, \mathtt{0}, \ldots, \mathtt{0})^T$ that contains $\mathtt{1}$-s in the first k positions and $\mathtt{0}$-s in the last $N-k$ positions. There are $\left(\begin{array}{c} N \\ k \end{array} \right)$ possible words with k $\mathtt{1}$-s; thus we can rewrite (1) as:

$$E\left[M(T_N)\right] = \sum_{k=0}^{N} \left(\begin{array}{c} N \\ k \end{array} \right) P_k. \quad (4)$$

We state now the main idea behind computing P_k. Given the output word $\mathbf{y} = (\mathtt{1}, \ldots, \mathtt{1}, \mathtt{0}, \ldots, \mathtt{0})^T$ with k $\mathtt{1}$-s in its first k positions and $\mathtt{0}$-s in its remaining positions, let D denote the set of indices j such that column j of matrix T_N contains a $\mathtt{1}$ in one or more of its k first positions. The set D describes all positions in the input vector that are connected to one or more of the first k positions in the output vector. Therefore achieving output $\mathbf{y}$ requires $x_j = 1$ for all $j \in D$. Let $D^C = \{1, \ldots, N\} - D$. For the given $\mathbf{y}$ to be achievable, each of the last $N-k$ output terminals must be connected to at least one of the inputs in D^C. Setting $x_j = 0$ for all $j \in D^C$ then achieves the desired result.

We provide the following example to clarify this reasoning.

Example 2 *Suppose $N = 4$, and the realization of the channel matrix is:*

$$T_4 = \left(\begin{array}{cccc} \mathtt{1} & \mathtt{1} & \mathtt{0} & \mathtt{1} \\ \mathtt{0} & \mathtt{1} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{1} & \mathtt{0} \\ \mathtt{1} & \mathtt{0} & \mathtt{1} & \mathtt{0} \end{array} \right).$$

Suppose the desired output word is $\mathbf{y} = (\mathtt{1}, \mathtt{1}, \mathtt{0}, \mathtt{0})^T$ (so $k = 2$). Set D contains the indices of all columns that contain $\mathtt{1}$-s in their first $k = 2$ rows; thus $D = \{1, 2, 4\}$). The remaining column indices correspond to input values that can be set to $\mathtt{0}$. Each of the $\mathtt{0}$ outputs (in positions 3 and 4) must be connected to at least one of the inputs in $D^C = \{3\}$. Thus achieving output $\mathbf{y}$ requires a $\mathtt{1}$ in column 3 of rows 3 and 4. Since this constraint is met, output $\mathbf{y}$ is achievable.

In contrast, if the desired output word is $\mathbf{y}' = (\mathtt{0}, \mathtt{1}, \mathtt{0}, \mathtt{1})^T$, the index set corresponding to row 3 is completely covered by the union of index sets corresponding to rows 2 and

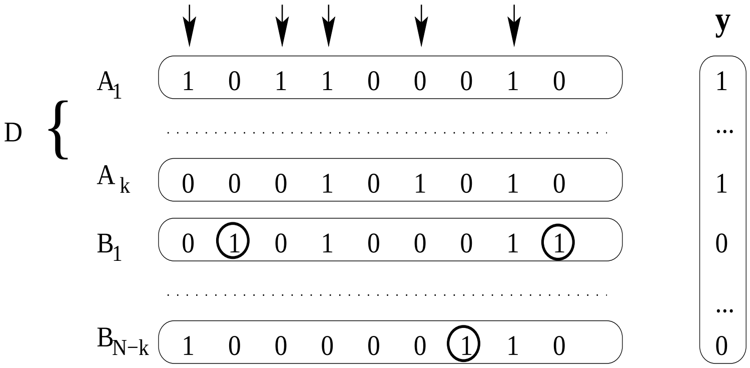

Figure 4: Consider an output word $\mathbf{y}$ with k 1-s (left). The channel matrix T_N (right) is represented by its rows. The sets of column indices $\{A_l\}_{l=1}^{k}$ correspond to the 1-s in the output word. The union set $D = \cup_{l=1}^{k} A_l$ contains the column indices highlighted by arrows. The sets of column indices $\{B_m\}_{m=1}^{N-k}$ correspond to the 0-s in the output word. The condition for $\mathbf{y}$ to be a valid output word is that each set $\{B_m\}_{m=1}^{N-k}$ should contain at least one element in D^C (those elements are indicated by circles).

4, *and thus there is no input word that can generate the desired* $\mathbf{y}'$. *However, given the symmetry of the problem, there exists another realization of the matrix* T_4, *obtained by swapping rows* 1 *and* 4, *that happens with the same probability, and for which the word* $\mathbf{y}'$ *can be produced at the output, while the word* $\mathbf{y}$ *can not; it is thus enough to study the probability of appearance for only one of the two words.*

We now formalize this idea. Fix k, and denote the following sets of output indices corresponding to a realization of the matrix T_N (see Figure 4):

- $\{A_l\}_{l=1}^{k}$ the sets of *column* indices $A_l = \{j : T_N(l, j) = 1\}$, each set corresponding to one of the first k *rows* of the matrix T_N.
- $\{B_m\}_{m=1}^{N-k}$ the sets of *column* indices $B_m = \{j : T_N(m + k, j) = 1\}$, each set corresponding to one of last $N - k$ *rows* of the matrix T_N.

Note that the actual values of $\{A_l\}_{l=1}^{k}$ and $\{B_m\}_{m=1}^{N-k}$ are statistically independent, since the elements of the matrix T_N are drawn independently.

Using this notation $D = \cup_{l=1}^{N}\{A_l\}$, then D is a random set of indices, of maximum size N, and each of the indices in the set $\{1, \ldots, N\}$ belongs to the set D with i.i.d. Bernoulli distributions generated by biased coin tosses of probability q. In order to compute q, we observe that the probability that a given column index does *not* belong to any of the sets $\{A_l\}_{l=1}^{k}$ is $(1-p)^k$, and thus:

$$q = 1 - (1-p)^k. \tag{5}$$

Further, we rewrite P_k, by summing over all the possible sizes $s = 0 \ldots N$ of the set D, weighted by the probability that the set D has size s:

$$P_k = \sum_{s=0}^{N} \left(\begin{array}{c} N \\ s \end{array} \right) \cdot q^s (1-q)^{N-s} \cdot P_s(B \nsubseteq D, \forall B), \tag{6}$$

where we denote by $P_s(B \nsubseteq D, \forall B)$ the probability that none of the sets $\{B_m\}_{m=1}^{N-k}$ is included in the set D, given that the size of D is s. Thus $P_s(B \nsubseteq D, \forall B)$ is the probability that each of the `0` outputs is connected to one or more of the `0` inputs (that is, the columns in D^C).

Denote by $P_s(B \nsubseteq D)$ the probability that the set B is not included in the set D, given that the size of D is s. Then we can rewrite equation (6) as:

$$\begin{aligned} P_k &= \sum_{s=0}^{N} \binom{N}{s} q^s (1-q)^{N-s} \prod_{m=1}^{N-k} P_s(B_m \nsubseteq D) \\ &= \sum_{s=0}^{N} \binom{N}{s} q^s (1-q)^{N-s} P_s(B_1 \nsubseteq D)^{N-k}. \end{aligned} \tag{7}$$

Now, the random set B_1 is *not* included in the set D if there exists at least one element in B_1 that is not in the set D of size s. The probability that set B_1 does not contain any of the $N-s$ elements not belonging to D is $(1-p)^{N-s}$, and thus:

$$P(B_1 \nsubseteq D) = 1 - (1-p)^{N-s}. \tag{8}$$

By combining equations (7) and (8) we obtain that:

$$P_k = \sum_{s=0}^{N} \binom{N}{s} q^s (1-q)^{N-s} (1-(1-p)^{N-s})^{N-k}. \tag{9}$$

Finally, by combining (5) and (9) into (4) we get the result of the theorem. ■

4 Numerical Simulations

In Figure 5 we plot the simulated capacity curve and its upper bound obtained by both simulations and our theoretical formula (3), where $N = 10$ and the simulation results are obtained by averaging over 40 instances. Note that the upper bound is reasonably tight for most regions of the capacity curve, the reason being that the distribution of the number of words is narrow for most values of p, and thus the Jensen inequality is rather tight as well. From the results in Figure 6, we conjecture that the maximum value of the upper bound of the capacity for a given dimension N is a linear function[1] of N, $\max_p(\log E[M(T_N)]) = aN$, where the linear coefficient is $a \approx 0.64$. Figure 7 shows how the optimal value for the probability $p_0(N)$, for which the maximum normalized value of the upper bound is attained, moves towards zero as N increases. This suggests that the optimal p is inversely proportional with a sub-unitary power of N.

5 Conclusions

We studied a binary MIMO system with interference, where the level of interference is given by the number of links present between the input and the output terminals. We derived an upper bound on the capacity, valid for all the interference intensity levels, and illustrated our results with numerical experiments. Future work includes the study of more complex networks of terminals.

---

[1] The irregularity of the plotted curve is due to the maximal resolution for the probability p chosen in the simulations.

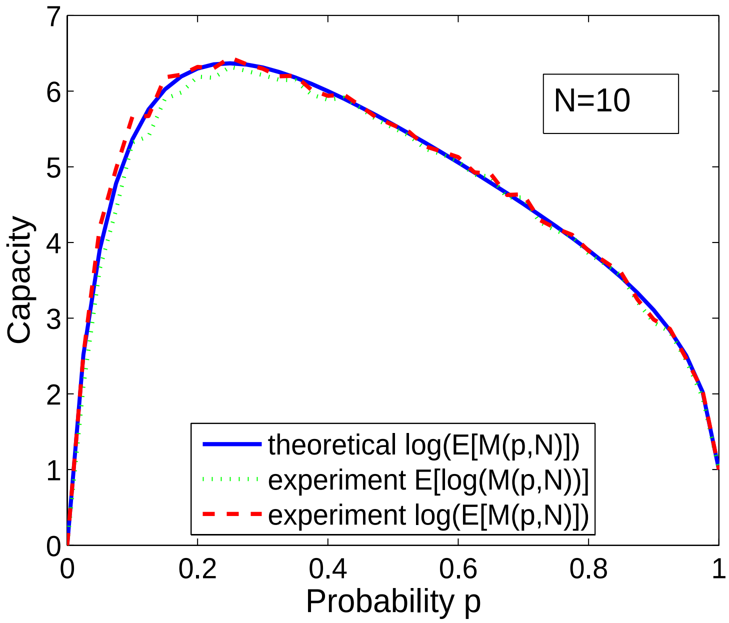

Figure 5: The theoretical and simulated upper bound, and the simulated value of the capacity, for $N = 10$.

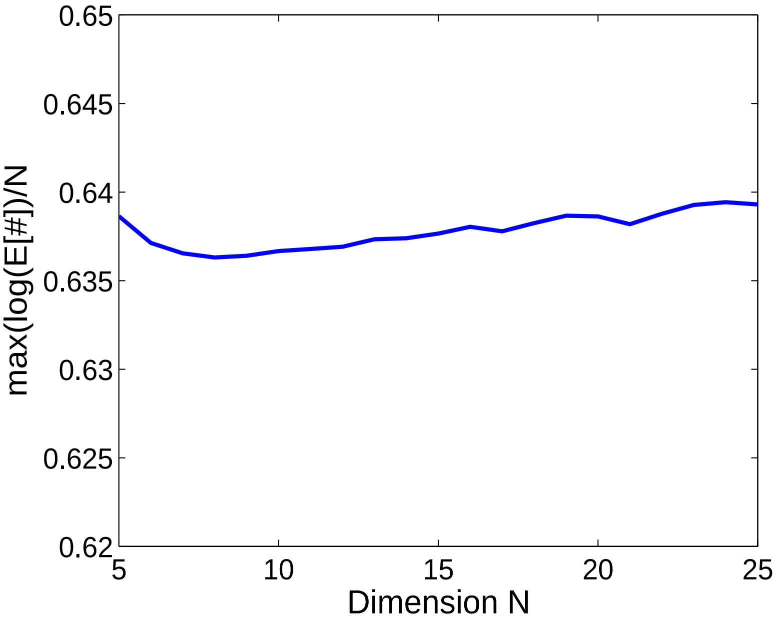

Figure 6: The normalized maximum value of the upper bound, for $N = 5..25$.

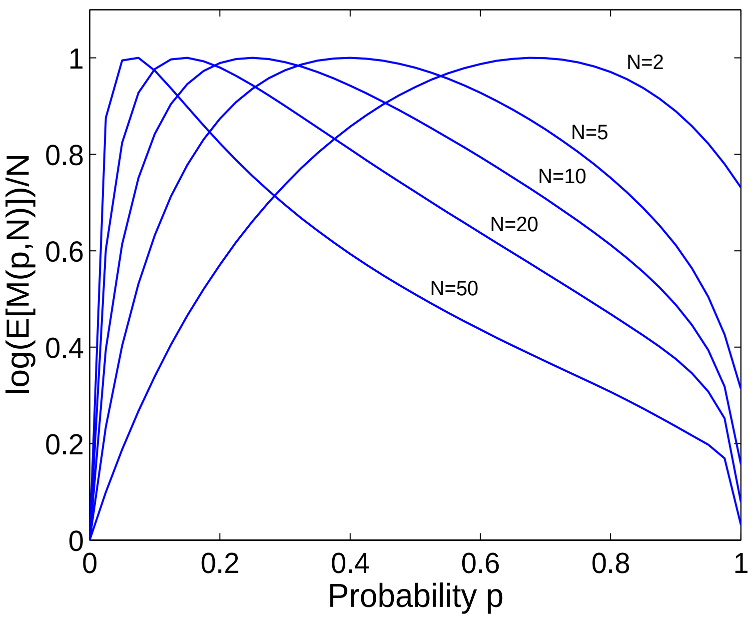

Figure 7: The normalized (by division with N) capacity upper bound, for various values of N.

Acknowledgment

We acknowledge Ralph Koetter for useful discussions that stimulated the study of the problem in this paper.

References

[1] J. Blomer, R. Karp and E. Welzl: *The Rank of Sparse Random Matrices over Finite Fields*, Random Structures and Algorithms 10, 1997.

[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, 1998.

[3] J. Fulman: *Random Matrix Theory over Finite Fields*, Bulletin of the AMS, vol. 39, no. 1, 2001.

[4] A. Goldsmith, S.A. Jafar, N. Jindal and S. Viswanath: *Capacity Limits of MIMO Channels*, IEEE JSAC, vol. 21, 2003.

[5] I. E. Telatar: *Capacity of Multiple-Antenna Gaussian Channels*, European Transactions on Telecommunications, vol. 10, 1999.

[6] D. Tse: *Multiuser Receivers, Random Matrices and Free Probability*, in Proc. Allerton 1999.

[7] J. Winters: *On the Capacity of Radio Communications Systems with Diversity in a Rayleigh Fading Environment*, IEEE JSAC, vol. 5, 1987.