[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Capacity Approximations for a Deterministic MIMO Channel

Ira S. MOSKOWITZ[1], Paul COTAE[2], Myong H. KANG[3], Pedro N. SAFIER[4]
[1]*Center for High Assurance Computer Systems, Code 5540*
Naval Research Laboratory, Washington, DC 20375
moskowitz@itd.nrl.navy.mil
[2]*The University of the District of Columbia,*
School of Engineering and Applied Sciences
Department of Electrical and Computer Engineering, Washington, DC 20008
pcotae@udc.edu
[3]*Center for High Assurance Computer Systems, Code 5540*
Naval Research Laboratory, Washington, DC 20375
mkang@itd.nrl.navy.mil
[4]*S&J Solutions LLC, 107 S. West St. PMB 509, Alexandria, VA 22314*
Pedro-safier@sj-solutions.com

***Abstract*—In this paper, we derive closed form approximations for the capacity of a point-to-point, deterministic Gaussian MIMO communication channel. We focus on the behavior of the inverse eigenvalues of the Gram matrix associated with the gain matrix of the MIMO channel, by considering small variance and large power assumptions. We revisit the concept of deterministic MIMO capacity by pointing out that, under transmitter power constraint, the optimal transmit covariance matrix is not necessarily diagonal. We discuss the water filling algorithm for obtaining the optimal eigenvalues of the transmitter covariance matrix, and the water fill level in conjunction with the Karush-Kuhn-Tucker optimality conditions. We revise the Telatar conjecture for the capacity of a non-ergodic channel. We also provide deterministic examples and numerical simulations of the capacity, which are discussed in terms of our mathematical framework.**

***Index Terms*—MIMO, transmitter optimization, channel capacity, Telatar conjecture, water filling.**

I. Introduction

In this paper we reexamine some of the fundamental concepts of MIMO channel capacity, focusing on deterministic MIMO channels. However, we also consider probabilistic channels, both of the ergodic and non-ergodic type. Our analysis shows that Telatar's conjecture [1] for the capacity of a non-ergodic channel needs a similarity adjustment via unitary matrices. We present evidence of this claim by extrapolating from the deterministic case.

We have found that some of the fundamental results in MIMO capacity have taken on the status of "folk theorems" here, we gather these results and make sure they are on firm mathematical ground.

Furthermore, as in the spirit of [2, 3], we have given approximations for the capacity of deterministic MIMO channels under realistic transmitter power constraints and properties of the gain matrix. These approximations are important for gleaning information about how capacity behaves, without the necessity of numerical calculations at every stage of the analysis.

To assist the reader, we conclude the introduction with a subsection on the notation used in this paper.

A. Notation

All vectors and matrices are complex, unless noted otherwise. Vectors are denoted by bold lower-case letters $\boldsymbol{a}$, matrices are denoted by bold upper-case letters $\mathbf{A}$. The determinant of $\mathbf{A}$ is $\det(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$ is the rank, $\mathrm{tr}(\mathbf{A})$ is the trace, and $\mathbf{A}^*$ denotes the conjugate transpose. We denote the (i, j) entry of $\mathbf{A}$ by $\boldsymbol{A}_{i,j}$. For real x, x^+ denotes max(x, 0), for complex $z, \bar{z}$ denotes the conjugate of z, and $|z|^2 = z\bar{z}$. The $n \times n$ identity matrix is written as $\mathbf{I}_n$. Given an $n \times n$ matrix $\mathbf{A}$, we denote the spectrum of $\mathbf{A}$ as the multiset of eigenvalues ε_i (possibly with repeated values) as $eig(\boldsymbol{A}) \triangleq \{\varepsilon_1,...,\varepsilon_n\}$. We have a similar multiset $eig^+(\boldsymbol{A})$ consisting, with multiplicity, of the positive (if any) eigenvalues of $\mathbf{A}$. If $\mathbf{A}$ has only real eigenvalues ε_i, then we may order them in non-increasing order. We denote the multiset of eigenvalues listed in non-increasing order as $\overleftarrow{eig}(\mathbf{A}) \triangleq \{\varepsilon_{\overline{1}}, \ldots, \varepsilon_{\overline{n}},\}$, where $\varepsilon_{\overline{j}} \geq \varepsilon_{\overline{j+1}}$.

We use the notation $\mathbf{\Lambda}$ to represent a diagonal matrix. If the diagonal matrix is a diagonalization[1] of the matrix $\mathbf{A}$, then we write $\mathbf{\Lambda_A}$. Note that $\mathbf{\Lambda_A}$ is, in general, not unique since there may be more than one diagonalization of $\mathbf{A}$. If $\mathbf{A}$ is diagonalized by a unitary[2] matrix $\mathbf{U}$, that is $\mathbf{A = U\Lambda_A U^*}$, then we have that $\mathbf{\Lambda_A = U^*AU}$, which easily tells us that $\mathrm{eig}(\mathbf{A}) = \mathrm{eig}(\mathbf{\Lambda_A})$. Therefore, the diagonal entries of $\mathbf{\Lambda_A}$ are, with multiplicity, the eigenvalues of $\mathbf{A}$. Note that the spectral theorem [4, Thm. 2.5.6] tells us that any Hermitian[3] matrix is diagonalized by a unitary matrix. The notation $\mathbf{\Lambda}_{(a_1,\ldots,a_n)}$ (often written in other literature as $\mathrm{diag}(a_1, \ldots, a_n)$) denotes the specific $n \times n$ diagonal matrix with a_i in the i,i entry. Note that $\mathbf{\Lambda}_{(a_1,\ldots,a_n)}$ is a specific matrix of the form $\mathbf{\Lambda_A}$.

[1] That is $\mathbf{A}$ and $\mathbf{\Lambda_A}$ are similar.
[2] A square matrix $\mathbf{U}$ is unitary iff $\mathbf{U}^* = \mathbf{U}^{-1}$.
[3] A square matrix $\mathbf{M}$ is Hermitian (self-adjoint) iff $\mathbf{M} = \mathbf{M}^*$.

Digital Object Identifier 10.4316/AECE.2011.03001

1582-7445 © 2011 AECE

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Finally, if the diagonal entries of $\mathbf{M}$ are denoted as $\mathsf{M}_{i,i}$, then we use the notation $\mathbf{\Lambda}_{\mathsf{M}_{i,i}}$ to be the diagonal matrix with (i, i) entry $M_{i,i}$.

II. MIMO Channel Model

We consider a point-to-point Gaussian MIMO communication channel, where the single sender employs T transmitting antennas, and R antennas are used by the sole receiver. The seminal reference for the analysis of MIMO capacity is Telatar [1]. The channel, in normalized[4] form (see [5, Sec.II.A]), between the sender and the receiver is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{x}$ is the transmitted $T \times 1$ input vector of the sender, $\boldsymbol{y}$ is the $R \times 1$ received vector, the *gain matrix* $\boldsymbol{H}$ is $R \times T$, and $\mathbf{n}$ is the $R \times 1$ additive white circularly symmetric complex Gaussian noise random vector with covariance matrix $E[\boldsymbol{n}\boldsymbol{n}^*] = \mathbf{I}_R$. (Such a random vector has zero mean, and is totally described by its covariance matrix [6, Sec A.1.3]). We may also denote such a MIMO channel as a (T, R) MIMO channel.

The transmit covariance $T \times T$ matrix is defined as the expectation matrix $\mathbf{Q} \triangleq E[\mathbf{x}\mathbf{x}^*]$. The MIMO communication system is assumed to have a total power (for all transmitting antennas) constraint given by the non-negative real number P such that $tr(\mathbf{Q}) \leq P$. Note that since $tr(\mathbf{Q}) = \mathsf{tr}(E[\mathbf{x}\mathbf{x}^*]) = E\mathsf{tr}[\mathbf{x}\mathbf{x}^*] = E[\mathbf{x}^*\mathbf{x}]$ we may also express our power constraint as $E[\boldsymbol{x}^*\boldsymbol{x}] \leq P$.

Telatar discusses two types of MIMO communication channels:

1. Deterministic — The gain matrix $\mathbf{H}$ is deterministic. In this scenario $\mathbf{H}$ is known by both the sender and the receiver.
2. Probabilistic — The gain matrix $\mathbf{H}$ is random and its distribution is known by the sender, and its realization is known by the receiver. Often the condition of *Rayleigh fading* is assumed (which means that the magnitudes of the elements of the random matrix $\mathbf{H}$ are independently Rayleigh distributed), but it does not have to be so. There are two possible cases in this scenario.
    a) Ergodic[5] — The gain matrix $\mathbf{H}$ is probabilistic, and each time the sender transmits, a realization of $\mathbf{H}$ is chosen according to its distribution.
    b) Non-ergodic — The gain matrix $\mathbf{H}$ is probabilistic, but once it is picked it never changes.

Let us start with the first situation.

2.1 Deterministic Channel

We assume that the gain matrix $\mathbf{H}$ (the channel state information, hereafter CSI), is known perfectly by the sender and receiver. Given $\mathbf{Q}$, such that $tr(\mathbf{Q}) \leq P$, the *mutual information*[6] [1, 5] between the sender and the receiver is

$$\mathcal{I}(\mathbf{Q}) \triangleq \log\det(\mathbf{I}_R + \mathbf{H}\mathbf{Q}\mathbf{H}^*). \tag{2}$$

Since the additive noise is normalized to have variance 1, the signal to noise ratio (SNR) is given by $P/1 = P$.

Mutual information is well-defined in the sense that $\det(\mathbf{I}_R + \mathbf{H}\mathbf{Q}\mathbf{H}^*) \geq 1$. This is because $\det(\mathbf{I}_R + \mathbf{H}\mathbf{Q}\mathbf{H}^*) = \prod_i(1 + \varepsilon_i)$, where ε_i are the eigenvalues (with multiplicity) of $\mathbf{H}\mathbf{Q}\mathbf{H}^*$. Thus, for the mutual information to be well-defined, it suffices to show that the eigenvalues $\varepsilon_i \geq 0$. First, since $\mathbf{Q}$ is a covariance (complex) matrix, it is Hermitian. Second, a covariance matrix [7] is positive semidefinite (psd)[7]. Note that since $\mathbf{Q}$ is psd, $\mathbf{v}^*\mathbf{H}\mathbf{Q}\mathbf{H}^*\mathbf{v} = (\mathbf{H}^*\mathbf{v})^*\mathbf{Q}(\mathbf{H}^*\mathbf{v}) \geq 0$. Then, it follows that $\mathbf{H}\mathbf{Q}\mathbf{H}^*$ is also psd. Therefore, as required, we have shown that the eigenvalues of $\mathbf{H}\mathbf{Q}\mathbf{H}^*$ are non-negative.

The very important *determinant identity* [8, Cor. 18.1.2], [7, 1.13.Thm.9] is often used in MIMO papers, yet the proof is hard to find. For the sake of completeness, since some tricks are used, we sketch the proof given in our references above.

Theorem 2.1. *(Determinant Identity) If $\boldsymbol{A}$ is $m \times n$ and $\boldsymbol{B}$ is $n \times m$ then*

$$\det(\mathbf{I}_m + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A}). \tag{3}$$

Proof. Since

$$\begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m + \mathbf{A}\mathbf{B} & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix}\begin{pmatrix} \mathbf{I}_m & 0 \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}\begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n + \mathbf{B}\mathbf{A} \end{pmatrix}$$

we have that

$$\det\begin{pmatrix} \mathbf{I}_m + \mathbf{A}\mathbf{B} & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix}\det\begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix} =$$

$$= \det\begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}\det\begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n + \mathbf{B}\mathbf{A} \end{pmatrix}$$

One can easily show by induction that partitioned matrices of the form $\begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{0} & \mathbf{M}_3 \end{pmatrix}$ or $\begin{pmatrix} \mathbf{M}_1 & \mathbf{0} \\ \mathbf{M}_2 & \mathbf{M}_3 \end{pmatrix}$ have determinant equal to $\det(\mathbf{M}_1)\det(\mathbf{M}_3)$. Thus the result trivially follows. □

Corollary 2.1. *([7]) The scalar λ is a non-zero eigenvalue of $\boldsymbol{AB}$ iff it is a non-zero eigenvalue of $\boldsymbol{BA}$.*

Proof. Say that $\lambda \neq 0$ is an eigenvalue of $\boldsymbol{AB}$. Then we know that $\det(\lambda\mathbf{I}_m - \mathbf{A}\mathbf{B}) = 0$. Since $\det(\lambda\mathbf{I}_m - \mathbf{A}\mathbf{B}) = \det\lambda(\mathbf{I}_m - \lambda^{-1}\mathbf{A}\mathbf{B}) = \lambda^m\det(\mathbf{I}_m - \lambda^{-1}\mathbf{A}\mathbf{B}) = \lambda^m\det(\mathbf{I}_m - \lambda^{-1}\mathbf{B}\mathbf{A}) = \lambda^{m-n}\det(\lambda\mathbf{I}_n - \mathbf{B}\mathbf{A})$. So $\lambda \neq 0$ is also an eigenvalue for $\mathbf{BA}$, the rest follows. □

The determinant identity allows us to express $\mathcal{I}(\mathbf{Q})$ as

[4] The normalization is done, as in [1], by modifying $\mathbf{H}$, so that the noise has unit power.

[5] This term is used in the sense that the temporal average is equivalent to keeping time constant, but averaging over different realizations.

[6] All logarithms are base 2, therefore information is measured in bits.

[7] We say that an $n \times n$ matrix $\mathbf{M}$ is psd iff $\mathbf{v}^*\mathbf{M}\mathbf{v} \geq 0$ for all n vectors $\mathbf{v}$. Note that if $\mathbf{v}$ is an eigenvector of $\mathbf{M}$ with eigenvalue λ, then $0 \leq \mathbf{v}^*\mathbf{M}\mathbf{v} = \lambda|\mathbf{v}|^2$ so any eigenvalue of $\mathbf{M}$ must be non-negative. (Note that the converse is true provided $\mathbf{M}$ is also Hermitian).

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

$$\mathcal{I}(\mathbf{Q}) = \log\det(\mathbf{I}_T + \mathbf{Q}\mathbf{H}^*\mathbf{H}). \quad (4)$$

The MIMO deterministic *capacity*, in units of bits per second per Hertz (bps/Hz) [8], is the maximum of the mutual information under a transmitting power constraint

$$C \triangleq \max_{\mathbf{Q}:tr(\mathbf{Q})\le P} \mathcal{I}(\mathbf{Q}). \quad (5)$$

Note that $\mathbf{H}^*\mathbf{H}$ is, by definition, a Gram matrix. What is important, aside from $\mathbf{H}^*\mathbf{H}$ being Hermitian, that it is also positive semi-definite. This is because $\mathbf{v}^*\mathbf{H}^*\mathbf{H}\mathbf{v} = |\mathbf{H}\mathbf{v}|^2 \ge 0$. Some observations are in order.

- The first observation is that the maximum is well-defined. By this we mean a supremum, for the above constrained subset, of $\mathcal{I}(\cdot)$ exists, but one must show that a *maximum* is actually achieved on the subset. The covariance matrices $\mathbf{Q}$ with trace less than or equal to P form the inverse image of a closed set, in the natural matrix topology that $T \times T$ covariance matrices inherit from the topology of all $T \times T$ matrices. Now, we use the Frobenius norm [4] of $\mathbf{Q}$, $||\mathbf{Q}|| \triangleq \sqrt{\mathrm{tr}(\mathbf{Q}\mathbf{Q}^*)} = \sqrt{\mathrm{tr}(\mathbf{Q}^2)}$ which is bounded, since the trace of $\mathbf{Q}$, which is the sum of the eigenvalues, is bounded by P. Thus, we see that $\mathbf{Q} : tr(\mathbf{Q}) \le P$ is a compact set, so a maximum is obtained. Note, one may also make a direct Karush-Kuhn-Tucker (KKT) optimization argument as in [9, Appendix].
- The second observation is that we can replace the maximization constraint $\mathbf{Q} : tr(\mathbf{Q}) \le P$ in (5) with $\mathbf{Q} : tr(\mathbf{Q}) = P$. Let us show this by contradiction. We ignore the logarithm, since it is an increasing function and just concentrate on $\det(\mathbf{I}_T + \mathbf{HQH}^*)$. Say that the maximum of the determinant is obtained for some $\mathbf{Q}' : tr(\mathbf{Q}') = P_1 < P$. We know that $\det(\mathbf{I}_T + \mathbf{H}\mathbf{Q}'\mathbf{H}^*) = \prod_i(1+\varepsilon_i)$, where the ε_i are, as before, the possibly non-distinct eigenvalues (with multiplicity) of $\mathbf{HQ'H}^*$. Consider the matrix $\mathbf{Q}'' = \frac{P}{P_1}\mathbf{Q}'$. If μ_i are the eigenvalues of $\mathbf{Q}'$, then $\frac{P}{P_1}\mu_i$ are the eigenvalues of $\mathbf{Q}''$, so $tr(\mathbf{Q}'') = P$. Note that $\det(\mathbf{I}_T + \mathbf{HQ'H}^*) = \det(\mathbf{I}_T + \frac{P}{P_1}\mathbf{HQ'H}^*) = \prod_i(1 + \frac{P}{P_1}\varepsilon_i)$. Since there must be at least one $\varepsilon_i \ne 0$ (we do not consider cases where $\mathbf{H}$ is the zero matrix, and we know that $\mathbf{Q}'$ is not the zero matrix), we see that $\prod_i(1 + \frac{P}{P_1}\varepsilon_i) > \prod_i(1+\varepsilon_i)$. So, by contradiction, we have[9] that

$$C = \max_{\mathbf{Q}\in\Phi} \mathcal{I}(\mathbf{Q}) \quad (6)$$

where Φ is the set of $T \times T$ covariance matrices with trace P.□

Since the CSI is known, both the sender and receiver know $eig(\mathbf{H}^*\mathbf{H}) = \{\mu_1, \ldots, \mu_T\}$, and $eig^+(\mathbf{H}^*\mathbf{H}) = \{\mu_1, \ldots, \mu_\varsigma\}$, where $\varsigma \le \min\{T, R\}$. Note that we could also use the identical (by Cor. 2.1) multiset $eig^+(\mathbf{HH}^*)$.

Since log is an increasing function, maximizing the mutual information $\mathcal{I}(\mathbf{Q})$ can be done by maximizing $\det(\mathbf{I}_T + \mathbf{QH}^*\mathbf{H})$. We call any $\mathbf{Q}$ that maximizes $\mathcal{I}(\mathbf{Q})$ *optimal* and use the notation $\mathbf{Q}_{op}$ for an optimal $\mathbf{Q}$, since the eigenvalues of $\mathbf{Q}$ are denoted as q_i, we denote the eigenvalues[10] of a $\mathbf{Q}_{op}$ as q_i^{op}. Let $eig(\mathbf{Q}) = \{q_1, \ldots, q_T\}$, we are interested in this eigenvalue spectrum when $\mathbf{Q}$ is optimal.

In terms of historical precedence, Telatar [1] discusses the idea of converting a point-to-point MIMO channel into orthogonal parallel, noninterfering SISO channels[11]. Also, [10] discusses "water filling" on the inverse eigenvalues of $eig^+(\mathbf{H}^*\mathbf{H})$.

We can ignore $\log(\cdot)$ in our discussion and simply concentrate on $\det(\mathbf{I}_T + \mathbf{QH}^*\mathbf{H})$. If $\mathbf{Q}$ commuted with $\mathbf{H}^*\mathbf{H}$, then $\det(\mathbf{I}_T + \mathbf{QH}^*\mathbf{H})$ could be trivially expressed [4, Thm.2.5.5], [11, Thm.3.1]. However, a priori we have no reason to assume such commutativity. Telatar [1, Sec. 3.2] cleverly applies the determinant identity (3) twice and shows that $\det(\mathbf{I}_T + \mathbf{QH}^*\mathbf{H}) \le \det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}}) = \prod_{i=1}^{T}(1 + \mathbf{\Lambda}_{Q_{i,i}}\mathbf{\Lambda}_{H^*H_{i,i}})$. However, there is a slight gap in Telatar's exposition. He has to show that the maximum is obtained for some $\mathbf{Q}'$ that is a covariance matrix with trace P, and then show that $\det(\mathbf{I}_T + \mathbf{Q}'\mathbf{H}^*\mathbf{H}) = \det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}})$. We will bridge that gap.

Since $\mathbf{H}^*\mathbf{H}$ is Hermitian, by the spectral theorem [4, Thm. 2.5.6], there exists a unitary matrix $\hat{\mathbf{U}}$, such that $\mathbf{H}^*\mathbf{H} = \hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{H}^*\mathbf{H}}\hat{\mathbf{U}}$.

Consider $\hat{\mathbf{U}}^*\mathbf{Q}\hat{\mathbf{U}}$, this matrix is Hermitian, and, because $\hat{\mathbf{U}}^* = \hat{\mathbf{U}}^{-1}$, it has the same spectrum as $\mathbf{Q}$. Therefore, $\hat{\mathbf{U}}^*\mathbf{Q}\hat{\mathbf{U}} \in \Phi$. Note that $\hat{\mathbf{U}}^*\mathbf{Q}\hat{\mathbf{U}} \in \Phi$ has real non-negative diagonal values and its trace is P. Therefore, $\mathbf{\Lambda}_{(\hat{U}^*Q\hat{U})i,i} \in \boldsymbol{\Phi}$.

Let $\mathbf{Q}' = \hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{Q}}\hat{\mathbf{U}}$, then $\mathbf{Q}' \in \boldsymbol{\Phi}$ also[12].

Consider $\det(\mathbf{I}_T + \mathbf{Q}'\mathbf{H}^*\mathbf{H}) = \det(\mathbf{I}_T + \mathbf{Q}'\hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{H}^*\mathbf{H}}\hat{\mathbf{U}}) = \det(\mathbf{I}_T + \hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{Q}}\hat{\mathbf{U}}\hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{H}^*\mathbf{H}}\hat{\mathbf{U}}) = \det(\mathbf{I}_T + \hat{\mathbf{U}}^*\mathbf{\Lambda}_{\mathbf{Q}}{}^*\mathbf{\Lambda}_{\mathbf{H}^*\mathbf{H}}\hat{\mathbf{U}})$ and, by the determinant identity
$\det(\mathbf{I}_T + \hat{\mathbf{U}}\hat{\mathbf{U}} * \mathbf{\Lambda_Q\Lambda_{H^*H}}) = \det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}})$.

Thus, we have shown that $\det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}})$ is not only an upper limit, but it is actually an achievable value for some $\mathbf{Q}' \in \boldsymbol{\Phi}$. Therefore it suffices to maximize $\det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}})$ for $\mathbf{Q} \in \boldsymbol{\Phi}$.

Trivially, we have that

$$\det(\mathbf{I}_T + \mathbf{\Lambda_Q\Lambda_{H^*H}}) = \prod_i(1 + q_i\mu_i) \quad (7)$$

[8] We have initially factored the bandwidth, in Hz, out of the capacity equation.

[9] See footnote 7.

[10] A priori there may be many such multisets of optimal eigenvalues.

[11] SISO stands for Single Input Single Output — the classical Shannon-type channel.

[12] A priori there is no reason why Q' should be diagonal; this is contrast to Telatar's confusing statement that the maximizing $\mathbf{Q}$ is diagonal. Actually, his statement applies to Hadamard's inequality—which, in fact, is maximized by some diagonal $\mathbf{Q}$ –not to the maximization of the mutual information.

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Since the μ_i are fixed, we must determine the corresponding q_i that maximize (7). We denote these maximizing eigenvalues as q_i^{op}. Note that

$$\max_{\mathbf{Q}\in\mathbf{\Phi}} \det(\mathbf{I}_T + \mathbf{\Lambda}_{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{H}^*\mathbf{H}})$$

is well-defined since we may let $\mathbf{Q}$ be $\mathbf{\Lambda}_{(q_i^{op},\ldots,q_T^{op})} \in \mathbf{\Phi}$.

The optimal eigenvalues are obtained by water filling and applying the Karush-Kuhn-Tucker (KKT) [12, 13] optimality conditions:

$$q_i^{op} = \left(\omega - \frac{1}{\mu_i}\right)^+. \tag{8}$$

where ω is the *water fill level.* If it were not for the $(\cdot)^+$ operation, finding the water level ω would be trivial we would just sum the non-zero q_i^{op}, set the sum to P and solve for ω. We have the multisets (in non-increasing) order $\overleftarrow{eig}$ (**H*H**) and $\overleftarrow{eig^+}$ (**H*H**), which gives us the unique multiset (in non-increasing order) $\overleftarrow{eig^+}$ ($\mathbf{Q}_{\text{op}}$). This gives us the unique multiset of $\{q_{\overleftarrow{1}}^{op}, \cdots, q_{\overleftarrow{T}}^{op}\}$, where

$$q_{\overleftarrow{i}}^{op} = \begin{cases} \omega - \frac{1}{\mu_{\overleftarrow{i}}} & \text{for } 1 \le i \le \varsigma \\ 0 & \text{for } \varsigma \le i \le T. \end{cases} \tag{9}$$

Note that the q_i^{op} correspond to the power allocated on the orthogonal parallel channels discussed above. This is the approach shown in [1, Sec. 3.2] where it is stated that when we have perfect CSI (knowledge of a fixed **H**)[13], then

$$C = \log\det(\mathbf{I}_T + (\widehat{\mathbf{U}}^*\mathbf{\Lambda}_{(q_1^{op},\ldots,q_T^{op})}\widehat{\mathbf{U}})\mathbf{H}^*\mathbf{H}). \tag{10}$$

Thus,

$$C = \log\prod_{i=1}^{T}(1 + q_i^{op}\mu_i) = \sum_{t=1}^{T}\log(1 + q_i^{op}\mu_i).^{14} \tag{11}$$

One must perform an iterative algorithm to find first the q_i^{op}, and then obtain ω by summing the non-zero q_i^{op} and setting that to P. It is an issue of how much power P is available.

The water filling analogy of the solution algorithm is as follows: We have a water tank with infinitely high sides that is T units across and one unit deep. We place T bricks in the bottom of the tank. The ith brick is 1 unit long and deep and $\frac{1}{\mu_{\overleftarrow{i}}}$ units high, and we consider them in ascending order of height $\frac{1}{\mu_{\overleftarrow{i}}}$ from brick 1 to brick T. The key to determining ω is to keep in mind that $tr(\mathbf{Q}^{op}) = \sum_i q_i^{op} = P$. The procedure is as follows:

1. The heights of the bricks are non-decreasing as we go from left to right, i.e., the first brick that we consider is the shortest one. There is always enough water (i.e., power P) to cover the first brick, and therefore $q_{\overleftarrow{1}}^{op} = \omega - \frac{1}{\mu_{\overleftarrow{1}}} > 0$. We initially set $\omega = P + \frac{1}{\mu_{\overleftarrow{i}}}$. The height of $\frac{1}{\mu_{\overleftarrow{2}}}$ determines if there is residual power to cover the second brick; that is if $P + \frac{1}{\mu_{\overleftarrow{1}}} < \frac{1}{\mu_{\overleftarrow{2}}}$ we cannot cover the second brick and we stop here, else:
2. We have enough power to cover at least the first two bricks and we now also have that $q_{\overleftarrow{2}}^{op} = \omega - \frac{1}{\mu_{\overleftarrow{2}}} > 0$. We know the eigenvalues of Q^{op} and adjust the water fill level such that $\omega = \frac{1}{2}\left(P + \frac{1}{\mu_{\overleftarrow{1}}} + \frac{1}{\mu_{\overleftarrow{2}}}\right)$. We now determine if we have enough power to cover the third brick, that is, if $P \le \frac{2}{\mu_3} - \frac{1}{2}\left(\frac{1}{\mu_{\overleftarrow{1}}} + \frac{1}{\mu_{\overleftarrow{2}}}\right)$ we are done, else:
3. We keep iterating this process until all the bricks are covered or we run out of power. The index of the last brick that is covered is $\varsigma \le T$. This gives us a water fill level of

$$\omega = \frac{P}{\varsigma} + \frac{1}{\varsigma}\sum_{i=1}^{\varsigma}\frac{1}{\mu_{\overleftarrow{i}}}. \tag{12}$$

What is interesting about the above equation is that the water fill level is given by the total power P normalized by ς added to the average heights of the bricks that are covered. Furthermore, this tells us that the non-zero eigenvalues of $\mathbf{Q}^{op}$ are

$$q_{\overleftarrow{i}}^{op} = \begin{cases} \frac{P}{\varsigma} + \left(\frac{1}{\varsigma}\sum_{k=1}^{\varsigma}\frac{1}{\mu_{\overleftarrow{k}}}\right) - \frac{1}{\mu_{\overleftarrow{i}}} & \text{for } 1 \le i \le \varsigma \\ 0 & \text{for } \varsigma < i \le T. \end{cases} \tag{13}$$

Since $\omega \ge (<)\frac{1}{\mu_{\overleftarrow{k}}}\ for\ i \le (>)\varsigma$, we have:

$$\omega\mu_{\overleftarrow{i}} = \begin{cases} \ge 1 & \text{for } 1 \le i \le \varsigma \\ < 1 & \text{for } \varsigma < i \le T. \end{cases} \tag{14}$$

From (11) we obtain Telatar's result [1, Sec. 3.2]:

$$C = \sum_{i=1}^{\varsigma}\log\left(1 + \left(\omega - \frac{1}{\mu_{\overleftarrow{i}}}\right)\mu_{\overleftarrow{i}}\right) + \sum_{i=\varsigma+1}^{T}\log(1+0) \tag{15}$$

$$= \sum_{i=1}^{\zeta}\log(\omega\mu_{\overleftarrow{i}}) + \sum_{i=\zeta+1}^{T}0 \tag{16}$$

$$= \sum_{i=1}^{T}\left(\log\left(\omega\mu_{\overleftarrow{i}}\right)\right)^+ = \sum_{i=1}^{T}(\log(\omega\mu_i))^+ \tag{17}$$

[13] Keep in mind that need not equal $q_{\overleftarrow{i}}^{op}$, and similarly for the q_i^{op} and similarly for the $\mu_{\overleftarrow{i}}$ and μ_i.

[14] Which is also equal to $\sum_{i=1}^{T}\log(1 + q_{\overleftarrow{i}}^{op}\mu_{\overleftarrow{i}})$.

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Notice that if the $\frac{1}{\mu_{\overleftarrow{i}}}, i < \zeta$ are clustered about their mean $1/\mu$ we can approximate

$$\frac{1}{\zeta}\sum_{k=1}^{\zeta}\frac{1}{\mu_{\overleftarrow{k}}} = \frac{1}{\mu}. \tag{18}$$

We call this the *Small Variance Assumption* (SVA); using it in (12) and (13), we obtain the SVA value of the water-fill level:

$$\text{SVA:}\quad \omega = \frac{P}{\zeta} + \frac{1}{\mu}, q_{\overleftarrow{i}}^{op} \approx \begin{cases} \frac{P}{\zeta} & \text{for } 1 \le i \le \zeta \\ 0 & \text{for } \zeta < i \le T. \end{cases} \tag{19}$$

2.1.1 SVA Capacity Approximation

If the SVA is assumed then we approximate the capacity as follows

$$C_{SVA} = \sum_{i=1}^{\zeta}\log\left(1 + \frac{P}{\zeta}\mu_{\overleftarrow{i}}\right) \tag{20}$$

where ζ is, as before, the index of the last $\mu_{\overleftarrow{i}}$ to be "covered with water." This lets us also express C_{SVA} as

$$C_{SVA} = \log\det\left(\mathbf{I}_T + \widehat{\mathbf{U}}^*\frac{P}{\zeta}\Lambda\left(\underbrace{*,\ldots\ldots\ldots\ldots\ldots\ldots,*}_{\zeta \text{ ones and } T-\zeta \text{ zeros}}\right)\widehat{\mathbf{U}}\cdot\mathbf{H}^*\mathbf{H}\right) \tag{21}$$

where as before $\mathbf{H^*H} = \widehat{\mathbf{U}}^*\mathbf{\Lambda_{H^*H}}\widehat{\mathbf{U}}$[15]. We express the diagonal matrix as ζ ones, followed by T - ζ zeros because we do not have control over the ordering of how the eigenvalues of $\mathbf{H^*H}$ are expressed in the diagonal form $\widehat{\mathbf{U}}^*\mathbf{\Lambda_{H^*H}}\widehat{\mathbf{U}}$.

If *all* of the $\frac{1}{\mu_{\overleftarrow{i}}}$ are approximately equal, then the index ζ is set equal to T because there is enough power to flow over all of the $\frac{1}{\mu_{\overleftarrow{i}}}$, since they are all at the "same" height $\frac{1}{\mu}$. Furthermore we can drop the index on the optimal eigenvalues of $\mathbf{Q}^{op}$. Thus we obtain the *Strong Small Variance Assumption* (SSVA; we use the word strong since it involves the maximal value $\zeta = T$)

$$SSVA:\quad \omega \approx \frac{P}{T} + \frac{1}{\mu}, q_i^{op} \approx \begin{cases} \frac{P}{T} & \text{for } 1 \le i \le \zeta \\ 0 & \text{for } \zeta < i \le T. \end{cases} \tag{22}$$

2.1.2 SSVA Capacity Approximation

If SSVA (which is a special case of the SVA) is assumed, then we approximate the capacity as follows

$$C_{SSVA} = \log\det\left(\mathbf{I}_T + \frac{P}{T}\mathbf{H}*\mathbf{H}\right) = \sum_{i=1}^{T}\log\left(1 + \frac{P}{T}\mu_{\overleftarrow{i}}\right)$$

$$= \sum_{i=1}^{T}\log\left(1 + \frac{P}{T}\mu_i\right). \tag{23}$$

Note that under the SVA or SSVA the capacity approximation involves an, at worst, suboptimal choice of eigenvalues for $\mathbf{Q}$, therefore $C_{SVA}, C_{SSVA} \le C$.

Two points must be stressed. First, the fact that the $1/\mu_i$ are clustered around their mean $\overline{1/\mu_i}$ does not guarantee that the μ_i are close to their mean $\overline{\mu_i}$ and vice versa. Second, it is worth nothing that the deterministic MIMO capacity naturally involves the *inverse* of the eigenvalues of $\mathbf{H^*H}$, not the eigenvalues *per se*.

2.2 Ergodic Channel

Recall that in this case $\mathbf{H}$ is probabilistic (it is usually assumed that $\mathbf{H}$ represents Rayleigh fading), and every time the channel transmits, a new realization of $\mathbf{H}$ is drawn. In this situation expected values of mutual information and capacity are used. Of course, one must be cautious with such terms because Shannon's [14] coding results were not originally given for such concepts, and new thoughts in coding and throughput must be considered. Following, the discussion as in [1], we define the ergodic capacity $\mathfrak{C}$, also in units of bps/Hz as an expected value:

$$\mathfrak{C} \triangleq \mathcal{E}[\log\det(\mathbf{I}_T + (P/T)\mathbf{H}^*\mathbf{H})]. \tag{24}$$

Note that multiplying the matrix $\mathbf{H^*H}$ on the left by the scalar P/T is equivalent, to the matrix multiplication $\underbrace{\mathbf{\Lambda}_{(P/T,\ldots,P/T)}}_{\text{T terms}}\mathbf{H}^*\mathbf{H}$.

We have that (24) can be expressed as:

$$\mathfrak{C} = \mathcal{E}[log\prod_i^T\left(1 + \frac{P}{T}\mu_i\right)] = \mathcal{E}\left(\sum_i^T\log\left(1 + \frac{P}{T}\mu_i\right)\right). \tag{25}$$

where the μ_i are the random eigenvalues, with multiplicity, of $\mathbf{H^*H}$. Thus, in the ergodic case the optimal $\mathbf{Q}$ is also of the form

$$\Lambda_{(P/T,\ldots,P/T)} = (P/T)\mathbf{I}_T. \tag{26}$$

2.3 Non-ergodic Channel

2.3.1 Telatar's conjecture

As above, we consider a probabilistic $\mathbf{H}$. However, in the non-ergodic case, once $\mathbf{H}$ has been chosen it is constant. Attempting to maximize mutual information will fail, because there is a non-zero probability that the chosen $\mathbf{H}$ will not support a given capacity value. Nevertheless, if we incorporate the concept of outage probabilities, then one can attempt to find a $\mathbf{Q}$ that optimizes the throughput. The details for this are in [1, Sec. 5.1]. We now have the famous Telatar conjecture [1, Sec. 5.1]:

[15] Note that if $T = \zeta$ the diagonalizing matrices $\widehat{\mathbf{U}}$, $\widehat{\mathbf{U}}^*$ cancel out in (21) because the diagonal matrix has all ones down the diagonal and hence commutes with all matrices of the proper dimensions.

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Conjecture 2.1. *(*Telatar*) The optimal* **Q** *is of the form* $\frac{P}{k}\boldsymbol{\Lambda}_{(\underbrace{1,\ldots,1}_{k},\underbrace{0,\ldots,1}_{T-k})}$. *The value of* k *is inversely related to the outage probability.*

2.3.2 Adjustment of the Telatar Conjecture

Equation (21) mimics the Telatar conjecture for the non-ergodic channel. Note that (21) calls into question Telatar's choice of the optimal **Q** in his conjecture. We present a modified version of the conjecture below.

Conjecture 2.2. *For a non-ergodic channel, the optimal* **Q** *is of the form* $\boldsymbol{U}^{*}\frac{P}{k}\boldsymbol{\Lambda}_{\left(\underbrace{*,\ldots,*}_{k\ ones\ and\ T-k\ zeros}\right)}\boldsymbol{U}$. *The value of* k *is inversely related to the outage probability and* **U** *is unitary.*

2.4 Discussion

We see in all three situations that the optimal Q, under either the SVA for the deterministic case or, in general, for the other two cases, is of the form $\boldsymbol{\Lambda}_{\left(\frac{P}{T},\ldots,\frac{P}{T}\right)}$ or $\frac{P}{k}\boldsymbol{\Lambda}_{\left(\underbrace{1,\ldots,1}_{k},\underbrace{0,\ldots,0}_{T-k}\right)}$.

It remains to be seen how good the SVA approximation is, i.e., we wish to evaluate how far off from the actual capacity, the SVA capacity is. This approach was taken for binary input discrete memoryless channels in [15, 16, 2, 3]. We turn our attention to this issue in the next section.

III. Quality of the Approximation of Deterministic Capacity

3.1 SSVA revisited

We will assume that we are in the deterministic case and analyze the SSVA a bit further. We assume that:

1. There are T transmitting antennas,
2. Both the sender and the receiver know **H**.
3. The inverse eigenvalues $1/\mu_i$ of **H*H** are all approximately equal, and
4. The value of the total transmission power $P = tr(\mathbf{Q})$.

From our previous results (23) we know that we can approximate the capacity as

$$C_{SSVA} = \sum_{i=1}^{T}\log\left(1+\frac{P}{T}\mu_i\right)$$

3.2 The Case of Large Power P

Now let us examine the situation where the total power $P = \mathrm{tr}(\mathbf{Q})$ satisfies the inequality

$$P \gg \sum_{i=1}^{T}\frac{1}{\mu_i}.$$

This assures us that there is enough water to cover all of the inverse eigenvalues $\frac{1}{\mu_i}$. We find that the water fill level is

$$\omega = \frac{P}{T}+\frac{1}{T}\sum_{i=1}^{T}\frac{1}{\mu_i}$$

and that

$$q_i^{op} = \frac{P}{T}+\left(\frac{1}{T}\sum_{k=1}^{T}\frac{1}{\mu_k}\right)-\frac{1}{\mu_i}$$

with

$$C = \sum_{i=1}^{T}(\log\omega\mu_i) = \sum_{i=1}^{T}\log\left(\left(\frac{P}{T}+\frac{1}{T}\sum_{k=1}^{T}\frac{1}{\mu_k}\right)\mu_i\right) =$$

$$= \sum_{i=1}^{T}\left[(\log\mu_i)+\log\left(\frac{P}{T}+\frac{1}{T}\sum_{k=1}^{T}\frac{1}{\mu_k}\right)\right].$$

Since $P \gg \sum_{i=1}^{T}\frac{1}{\mu_i}$ we may approximate the capacity as

$$C \approx \sum_{i=1}^{T}\left[(\log\mu_i)+\log\left(\frac{P}{T}\right)\right] = \sum_{i=1}^{T}\left[\log\left(\frac{P\mu_i}{T}\right)\right]$$

Thus we have the *Large Power Assumption* (LPA), which has an approximate capacity of

$$C \approx C_{LPA} = T\log\left(\frac{P}{T}\right)+\sum_{i=1}^{T}\log\mu_i$$

$$= T\log\left(\frac{P}{T}\right)+\log\prod_{i=1}^{T}\mu_i \tag{27}$$

Furthermore [16], if the μ_i are large enough we can approximate $\log\left(\frac{P\mu_i}{T}\right)$ as $\log\left(1+\frac{P\mu_i}{T}\right)$, which gives us the *Large Power and Moderate Eigenvalues Assumption* (LPMEA), which has a capacity approximation similar to the SSVA

$$C_{LPMEA} = \sum_{i=1}^{T}\log(1+\frac{P}{T}\mu_i) \tag{28}$$

Thus, whenever the conditions for the SSVA are met (that is, all the eigenvalues of **H*H** are approximately the same), or the conditions for the LPMEA are met (that is, P is much greater than the sum of the inverse eigenvalues of **H*H** and $\forall i, P\mu_i \gg T$), the same form of the approximation can be used:

$$C_{\aleph} = \sum_{i=1}^{T}\log\left(1+\frac{P\mu_i}{T}\right) \tag{29}$$

Therefore, for the remainder of the paper we use the notation $C_{\aleph}$, when the conditions of either SSVA or LPMEA are assumed to have been met.

3.3 Deterministic Examples

We use a (2,2) MIMO channel. Keep in mind that $\mu_{\leftarrow_1} \geq \mu_{\leftarrow_2}$. We assume that we have enough power to cover with water both $1/\mu_{\leftarrow_1}$ and $1/\mu_{\leftarrow_2} \geq 1/\mu_{\leftarrow_1}$, that is $P \geq 1/\mu_{\leftarrow_2} - 1/\mu_{\leftarrow_1}$, so $\omega = \frac{P}{2}+\frac{1}{2}\left(\frac{1}{\mu_{\leftarrow_1}}+\frac{1}{\mu_{\leftarrow_2}}\right)$, and

[16] For large x we have that $\lim_{x\to\infty}(\log(1+x)-\log(x)) = 0$.

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

$$C = -2 + \log\left(1 + \left(P + \frac{1}{\mu_{\overleftarrow{2}}}\right)\mu_{\overleftarrow{1}}\right) + \log\left(1 + \left(P + \frac{1}{\mu_{\overleftarrow{1}}}\right)\mu_{\overleftarrow{2}}\right)$$
$$= -2 + \log\left(1 + \left(P + \frac{1}{\mu_2}\right)\mu_1\right) + \log\left(1 + \left(P + \frac{1}{\mu_1}\right)\mu_2\right).$$

In our example, the water filling conditions are satisfied for $P = 10$, and $2 \le \mu_1, \mu_2 \le 7$. Therefore, we can view (Figure 1) the capacity as a function of $\{(\mu_1, \mu_2)\}$, with a natural symmetry. Or, we can view it (Figure 2) as a function of the "fundamental domain" of $\left\{(\mu_{\overleftarrow{1}}, \mu_{\overleftarrow{2}}) : \mu_{\overleftarrow{1}} \ge \mu_{\overleftarrow{2}}\right\}$, under the "action" that swaps μ_1 with μ_2.

If the inverse eigenvalue variance of $\mathbf{H}^*\mathbf{H}$ is small then we have

$$C_{\aleph} = \log\left(1 + \frac{P}{2}\mu_{\overleftarrow{1}}\right) + \log\left(1 + \frac{P}{2}\mu_{\overleftarrow{2}}\right)$$
$$= \log\left(1 + \frac{P}{2}\mu_1\right)\left(1 + \frac{P}{2}\mu_2\right)$$
$$= \log\left(1 + P\frac{\mu_1 + \mu_2}{2} + P^2\frac{\mu_1\mu_2}{4}\right).$$

This is illustrated in Figure 3 with P =10 and the $\mu_i \in [2,7]$.

As we see, and not surprisingly, the analysis of approximations is directly related to the amount of power and the perturbations in inverse eigenvalues of $\mathbf{H}^*\mathbf{H}$.

If all T inverse eigenvalues are equal, any non-zero P will suffice to give us a water fill level of $P + 1/\mu$. In this case, C and C_{SSVA} are identical. This is well-illustrated in Figure 4. As the differences in inverse eigenvalues grows, so does the error. However, one must keep in mind that the difference in the inverse eigenvalues is *inversely* related to the difference in the actual eigenvalues. Therefore, if the eigenvalues are large, changing them does not have much of an effect upon the validity of the SSVA approximation. However, if the eigenvalues are small, then a slight change in them can result in a large error using the SSVA approximation, unless the power suitably grows.

In Figure 4 we illustrate how slight the approximation error is if P = 10 and the eigenvalues of **H*H** are constrained to the interval [2, 7]. Note that in this scenario, the maximum difference between inverse eigenvalues is 1/2 - 1/7≈.36. If we keep the eigenvalues in the range in question, and force ourselves to have enough power to cover all the inverse eigenvalues with water, then this is about as bad as the approximation will get — which is not very bad at all. However, if we consider very small eigenvalues, then the situation changes.

We let the eigenvalues be in [.0001, .0100]. We need a minimal power greater than 900, to cover both inverse eigenvalues. If we choose P = 1000, we see that we have a very large error (Figures 5 and 6). However, if we up the power to P = 10,000 we see in Figures 7 and 8 that we substantially reduce the error, and we can continue this process. In fact if P = 100,000 the error is $O(10^{-3})$.

In conclusion, we see that if the inverse eigenvalues are close to their mean value, we can approximate the capacity, and the same is true if we have large power.

Capacity MIMO 2,2 with P=10

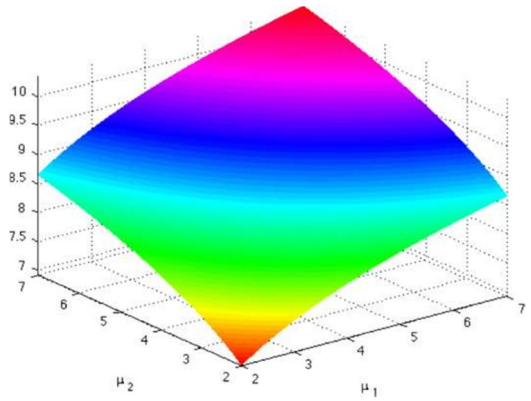


Figure 1. $C = -2 + \log\left(1 + \left(P + \frac{1}{\mu_2}\right)\mu_1\right) + \log\left(1 + \left(P + \frac{1}{\mu_1}\right)\mu_2\right)$

Capacity MIMO 2,2 with P=10

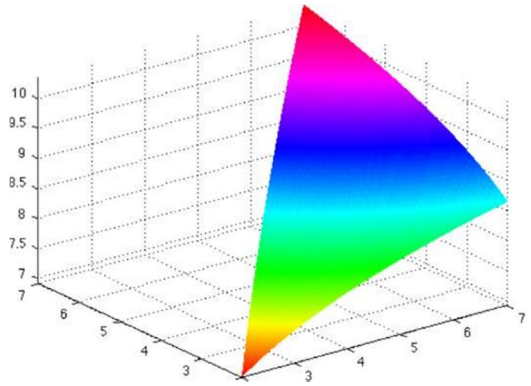


Figure 2. $C = -2 + \log\left(1 + \left(P + \frac{1}{\mu_{\overleftarrow{2}}}\right)\mu_{\overleftarrow{1}}\right) + \log\left(1 + \left(P + \frac{1}{\mu_{\overleftarrow{1}}}\right)\mu_{\overleftarrow{2}}\right)$

Capacity MIMO 2,2 with P=10

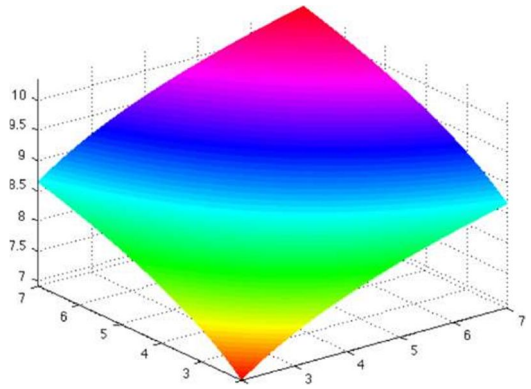


Figure 3. SSVA: $C_{\aleph}$, P =10.

DIFFERENCE MIMO 2,2 with P=10

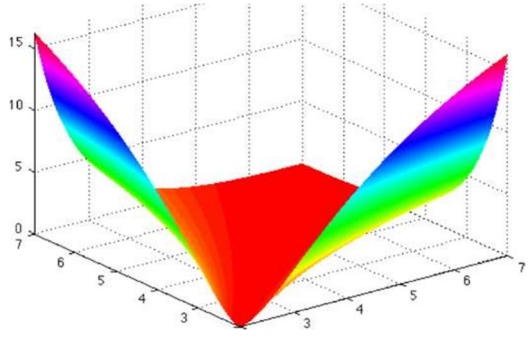


Figure 4. Difference: $O(10^{-3})$, C - $C_{\aleph}$.

[Downloaded from www.aece.ro on Thursday, March 28, 2024 at 22:16:56 (UTC) by 44.200.179.138. Redistribution subject to AECE license or copyright.]

Capacity MIMO 2,2 with P=1000

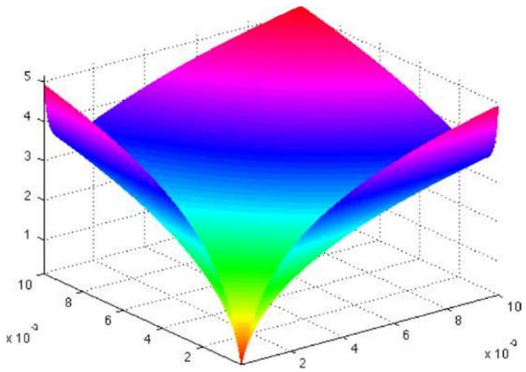


Figure 5. Capacity, P = 1000.

DIFFERENCE MIMO 2,2 with P=1000

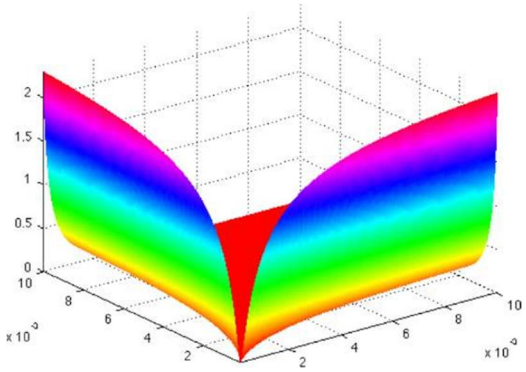


Figure 6. C - C_R , P = 1000.

Capacity MIMO 2,2 with P=10000

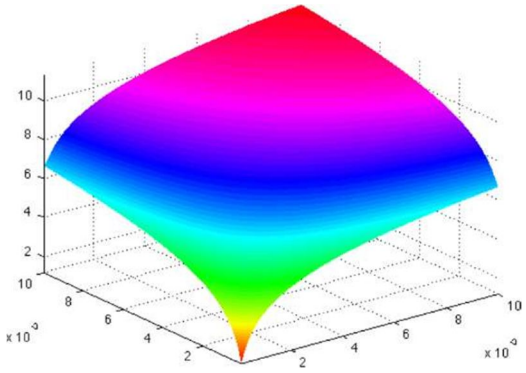


Figure 7. Capacity, P = 10,000.

DIFFERENCE MIMO 2,2 with P=10000

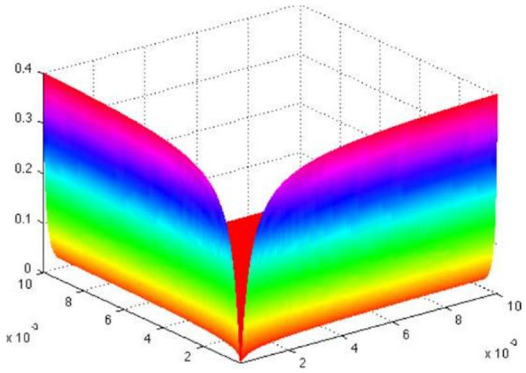


Figure 8. C - C_R P =10,000.

IV. Conclusion

We have examined the capacity formula for deterministic MIMO channels. We have put the analysis on a firm theoretical foundation and we have developed simple, closed form approximations for the capacity of the form

$$\sum_{i=1}^{T} \log\left(1 + \frac{P}{T}\mu_i\right).$$

We have discussed the Telatar conjecture, and have restructured it. Note that an overall theme of this paper has been to show how capacity, in both the deterministic and probabilistic cases, is the natural study of the behavior of inverse eigenvalues.

V. Acknowledgments

We thank Ruth Shulamite for her helpful comments. The second author wishes to acknowledge the support of the Office of Naval Research 2010 ONR-ASEE Fellowship Summer Program, the University of the District of Columbia 2010 Seed Grant Program and, Department of Defense Grant W911NF-11-1-0144.

References

[1] Emre Telatar. Capacity of multi-antenna Gaussian channels. European Transactions on Telecommunications, 10(6):585-596, 1999.

[2] Ira S. Moskowitz. An approximation of the capacity of a simple channel. In Proc. 43rd Annual Conf. on Information Sciences and Systems, CISS 2009, pages 164-169, Baltimore, MD, USA, March 2009.

[3] Ira S. Moskowitz. Approximations for the capacity of binary input discrete memoryless channels. In Proc. 44th Annual Conf. on Information Sciences and Systems, CISS 2010, Princeton, NJ, USA, March 2010.

[4] Roger A. Horn and Charles R. Johnson. Matrix Analysis. Cambridge University Press, 1985.

[5] Andrea Goldsmith, Syed Ali Jafar, Nihar Jindal, and Sriram Vishwanath. Capacity limits of mimo channels. IEEE Journal on Selected Areas in Communication, 21(5):684-702, 2003.

[6] David Tse and Pramod Viswanath. Fundamentals of Wireless Communication. Cambridge University Press, Cambridge, UK, 2005.

[7] Jan R. Magnus and Heinz Neudecker. Matrix Differential Calculus with Applications in statistics and Econometrics, Revised Ed.. *Wiley, Chichester, UK, 1999.*

[8] David A. Harville. Matrix Algebra From a Statistician's Perspective. Springer, New York, 2008/1997.

[9] Anke Feiten, Rudolf Mathar, and Stephen Hanly. Eigenvalue-based optimimum-power allocation for Gaussian vector channels. IEEE Transactions on Information Theory, 53(6):2305-2309, 2007.

[10] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. Wireless Personal Communications, 6:311-335, 1998.

[11] Paul Cotae. On the optimal sequences and total weighted square correlation of synchronous CDMA systems in multipath channels. IEEE Transactions on Vehicular Technology, 56(4):2063-2072, 2007.

[12] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In Proc. 2nd Berkeley Symposium on Mathematical Statistics & Probability, pages 481-492, Berkeley, CA, July, 1950. Univ. of California Press, 1951.

[13] Thomas M. Cover and Joy A. Thomas. Elements of Information Theory. Wiley Interscience, 2006.

[14] Claude E. Shannon. A mathematical theory of communication. Bell Systems Technical Journal, 27:379423,623-656, 1948.

[15] E. E. Majani. A Model for the Study of Very Noisy Channels & Applications, 1988. *PhD thesis, Cal Tech.*

[16] E. E. Majani and H. Rumsey. Two results on binary-input discrete memoryless channels. In Proceedings Int. Symp. on Information Theory, page 104, 1991.