Waterfilling Estimation for AWGN MIMO Channel Modeled as a Random Matrix

Victor M. Vergara
University of New Mexico/Dept. of Electrical and Computer Engineering, Albuquerque, NM, USA
Email: vvergara@ece.unm.edu

Silvio E. Barbin
Escola Politécnica da Universidade do São Paulo/Depto. de Engenharia de Telecomunicações e Controle, São Paulo, SP, Brasil and Centro de Tecnologia da Informação Renato Archer, Campinas, SP, Brazil
Email: barbin@usp.br

Ramiro Jordan
University of New Mexico/Dept. of Electrical and Computer Engineering, Albuquerque, NM, USA
Email: rjordan@ece.unm.edu

***Abstract*— Waterfilling solutions provide optimal power distribution in multiple-input multiple-output (MIMO) system design. However, the optimal distribution is usually obtained through costly computational processes, such as the determination of the system eigenvalues. For communication channels in a fast paced environment, the costs are even higher due to the necessity of tracking channel changes. In addition, the computational costs increase with the number of inputs and outputs, i.e. the size of the MIMO channel matrix. A solution for reducing the computational burden is to utilize pre-determined waterfilling based on the channel's statistics. No updates are required unless the channel statistical characteristics change. This work studies waterfilling estimations based on random matrix theory. The results can be applied when the channel coefficients follow a Rayleigh distribution and the noise is additive, white, and Gaussian.**

***Index Terms*— MIMO systems, random matrix, eigenvalues, waterfilling**

I. Introduction

Space diversity techniques have significantly evolved allowing multiple sub-channels to share the same transmission media. A communications design that compiles a set of several sub-channels in a single channel can be classified as a Multiple-Input Multiple-Output (MIMO) system. The system performance depends on the characteristics of the sub-channels, which are generally different from one another. In general, some sub-channels require more power than others, but the total transmission power is a limited resource. The optimal power distribution is known as waterfilling solution [1] and it improves the system performance.

The symbols transmitted through each sub-channel are firstly arranged in a vector s, known as the transmitted symbols vector. Symbols are assumed to be drawn from a random process and thus their power is related to their variance σ_{ss}^2. Power allocation is achieved by multiplying s by a diagonal matrix Φ before transmission. The coefficients ϕ_{ii} in the main diagonal of matrix Φ constitute the waterfilling solution. The transmitted vector z can be expressed as

$$z = V\Phi s, \tag{1}$$

where V is a unitary matrix obtained from the eigenvectors of the system. In addition to finding V, several other computations are required to recover the symbols at the receiving side. Although this is also an interesting problem it is not in the scope of this research. This work is focused on the determination of estimated solutions for Φ based on the statistical characteristics of the MIMO channel.

Optimal waterfilling solutions have been proposed in [2], [3] and [4], based on the channel matrix eigenvalues. As will be shown latter, the solution to a waterfilled MIMO system described by

$$y = Hz + n, \tag{2}$$

can be obtained using the eigenvalues λ_{ii} of the matrix $H^H H$, where H is the MIMO channel matrix, for a noise vector n with Gaussian and white elements. The waterfilling solutions were found assuming exact knowledge of H. However, estimated solutions can be determined for random channel matrices when the exact knowledge of H is not available. Assuming that every matrix element $h_{i,j}$ is an independent and identically distributed (i.i.d.) zero mean complex Gaussian variable with variance σ_H^2, it can be shown that the eigenvalues of $H^H H$ follow a very specific probability density function. In this case a waterfilling solution can be estimated by knowing only the size of the channel matrix and the variance of its elements σ_H^2, which are also the same requirements to the estimation of the eigenvalues.

Random matrix theory has been successfully applied in the determination of MIMO systems channels asymptotic capacity [5]. It has been shown that the channel capacity is a function of the channel matrix eigenvalues and has an asymptotic behavior determined using Girko's Law [6]. In this work the method proposed in [7] is expanded to exploit the asymptotic behavior of system eigenvalues. Instead of the Wigner's Quarter Circle Law (QCL) [8], that is only valid for square matrices, the Marčenko-Pastur Law [9] will be utilized to describe the system behavior for more general rectangular matrices H. The result is an asymptotic waterfilling solution that approximates the real solution as the number of sub-channels of the MIMO system increases.

The asymptotic waterfilling solution can be pre-calculated and stored off line. Then the pre-calculated solution can be applied during the symbol estimation while the system is in operation. The required Channel State Information (CSI) is based on the variance σ_H^2 of elements in H. A waterfilling solution update is not necessary unless a change in the variance σ_H^2 is detected.

The paper is structured as follows: after Section I, Section II presents some relevant results from random matrix theory. Section III applies these results in waterfilling solutions. Section IV presents numerical simulations and discusses the final results. Section V contains the conclusion.

II. Some Useful Results on Random Matrices

A. System Eigenvalues

Significant information on a system behavior can be obtained from the eigenvalues of the matrix $W = H^H H$, where H is the channel matrix with size $M \times N$. If the elements $h_{i,j}$ in H are complex random numbers following Gaussian distributions with zero mean and variance σ_H^2 the eigenvalues of W are continuous random variables that can be asymptotically described by a distribution known as the Marčenko-Pastur law [9].

The matrix eigendecomposition for W is

$$W = H^H H = V \Lambda V^H \tag{3}$$

where V is a unitary matrix composed of eigenvectors, and Λ is the diagonal matrix whose entries λ_{ii} are the eigenvalues of W. For the purpose of finding a waterfilling solution, the eigenvalues are arranged in decreasing order, i.e. $\lambda_{i-1,i-1} \geq \lambda_{i,i}$. The eigenvectors in V are arranged in an order that matches the eigenvalue elements in Λ.

There are two estimation goals that will be pursued: the estimation of each eigenvalue λ_{ii}, which leads to the matrix Λ, and the determination of a closed form for the expectation

$$E\left\{ \lambda^m \middle| \lambda \geq \lambda_{\overline{NN}} \right\} = \int \lambda^m p_{(\lambda)} d\lambda \cong \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \lambda_{ii}^m \,, \tag{4}$$

for the specific cases when $m = \{-2, \ -1, \ 1\}$. The term $p_{(\lambda)}$ refers to the single eigenvalue probability density function (p.d.f.). The single eigenvalue λ can be thought as a random process where the outcome is any eigenvalue indistinctively.

The analysis of the single eigenvalue p.d.f., performed in [10] for square matrices, proof that $p_{(\lambda)}$ converges to Wigner's QCL as the matrix size increases. A similar asymptotic behavior is known to occur for the Marčenko-Pastur law, which describes the case of square and rectangular matrices.

Several studies [9]-[13] exploit this asymptotic behavior of the eigenvalues. The upper limit index of the summation in (4) is the integer $\tilde{N}$ where $1 \leq \tilde{N} \leq \Theta$, and Θ is the total number of eigenvalues. Notice that $i = \{1,2,3,\ldots,\Theta\}$ where the first eigenvalue λ_{11} is the largest one and the Θ^{th} eigenvalue $\lambda_{\Theta\Theta}$ is the smallest one.

Expectations that can be expressed by (4) arise when solving waterfilling solutions. The random matrix theory provides methods that help us estimate Λ and (4). The following subsections will describe three approaches: the first uses $p(\lambda_{ii})$ if known, the second employs the roots of an associated Laguerre Polynomial, and the third integrates the Marcenko-Pastur law in a similar way as the QCL was used in [7]. Each approach exhibit advantages and that will be utilized according to the objectives described before.

B. Numerical Estimation by Expectation

The first approach considered here is the estimation of λ_{ii} using its expected value $\overline{\lambda_{ii}} = E\{\lambda_{ii}\}$. The expectation can be obtained analytically if the p.d.f. of the i^{th} eigenvalue $p_{(\lambda_{ii})}$ is available.

$$\overline{\lambda_{ii}} = \int \lambda_{ii} p_{(\lambda_{ii})} d\lambda \tag{5}$$

The exact closed form $p_{(\lambda_{ii})}$ for the i^{th} eigenvalue is yet an unsolved problem, except for the smallest eigenvalue $\lambda_{\Theta\Theta}$. Edelman [11] found a set of equations $p_{(\lambda_{\Theta\Theta})}$ for an arbitrary matrix size $M \times N$. Appendix A describes the formulation of $p_{(\lambda_{\Theta\Theta})}$ in more detail.

Expectation (5) cannot be easily computed. However, the expected value $\overline{\lambda_{ii}}$ can still be obtained by means of numerical Monte Carlo simulation methods. The mean $\overline{\lambda_{ii}}$ and variance $\sigma_{\lambda_{ii}}^2$ can be estimated by randomly selecting a large number of realizations of $H_{M\times N}$, calculating (3) and then averaging the corresponding moment. This method is slow and depends on the performance of random numbers generators.

C. Eigenvalues Estimation by Laguerre Polynomial

Other common estimates of the eigenvalues are obtained from the roots of an associated Laguerre polynomial. The p.d.f. $p_{(\lambda_{ii})}$ can be approximated as a Gaussian distribution centered at a root of an associated Laguerre polynomial. The estimation of the eigenvalue λ_{ii} is a corresponding Laguerre root multiplied by a normalization factor [12].

To apply this method, it is firstly necessary to closely examine the dimensions of $H_{M\times N}$. If $N > M$, then the matrix Λ defined in (3) has $N - M$ zero elements on the main diagonal, which is not a problem. If $M > N$, no zero elements in Λ are expected. In the formulation of Laguerre polynomial roots, the matrix dimensions are used as follows

$$Q = \max\{ M, N\} \tag{6}$$

$$\Theta = \min\{ M, N\}. \tag{7}$$

In summary, the roots $l_1, l_2, ..., l_\Theta$ of the associate Laguerre polynomial $c_0 + c_1 t + c_2 t^2 + \cdots$, which is defined as

$$L_\Theta^{(Q-\Theta)}(Qt) = \sum_{j=0}^{\Theta} \frac{1}{j!} \begin{pmatrix} Q \\ \Theta - j \end{pmatrix} (-Qt)^j . \tag{8}$$

approximate the normalized eigenvalues $\lambda_{ii} / Q\sigma_H^2$. The estimation then becomes

$$\lambda_{ii} = Q\sigma_H^2 l_i . \tag{9}$$

The convergence rate for the error, if $Q(\Theta) \to \infty$ as $\Theta \to \infty$, is

$$\max_{1\le i\le\Theta} \left| \frac{\lambda_{ii}}{Q\sigma_H^2} - l_i \right| = O\left(\left[\frac{\ln(\Theta)}{Q} \right]^{\frac{1}{4}} \right). \tag{10}$$

The estimation error converges to zero as the matrix size increases.

D. Eigenvalues Estimation by p.d.f. Integral

The last two methods directly estimate the eigenvalues in Λ, but the expectation in (4) is left for posterior processing. The waterfilling formulation can be greatly improved by using closed form equations for (4). The Marchenko-Pastur law [9] provides an appropriate method to directly calculate the expectations of eigenvalues as defined in (4). Extending the work presented in [6] for square matrices, this section will follow a similar approach with the exception that we are also considering non-square matrices, i.e. $Q \ge \Theta$. The development that will be presented depends on the parameter

$$\beta = \frac{\Theta}{Q} \Rightarrow \beta \le 1. \tag{11}$$

The eigenvalues requires a normalization expressed by

$$x = \frac{\lambda}{Q\sigma_H^2}. \tag{12}$$

The eigenvalues x are lower and upper limited as

$$a \le x \le b \text{, where} \tag{13}$$

$$a = \left(1 - \sqrt{\beta} \right)^2 \text{, and} \tag{14}$$

$$b = \left(1 + \sqrt{\beta} \right)^2. \tag{15}$$

The limits in (13) are also the convergence of the minimum λ_{NN} and the maximum λ_{11} respectively as $Q \to \infty$ [13]. The probability density function (p.d.f.) $p_{(x)}$ of x is then

$$p_{(x)} = \frac{\sqrt{(b-x)(x-a)}}{2\pi\beta x}. \tag{16}$$

As proposed in [7] the index j of the eigenvalues when ordered increasingly can be obtained by using frequency estimation and the cumulative probability function (c.d.f.)

$$P[a \le x \le x_j] = \int_a^{x_j} p_{(x)} dx \cong \frac{2j-1}{2\Theta}. \tag{17}$$

The integration in (17) leads to a mapping between indexes and normalized eigenvalues as

$$j \cong \frac{1}{2} + \frac{\Theta}{2}\left(1 + \frac{1}{\pi\beta} \Omega_{(x)} \right). \tag{18}$$

where

$$\Omega_{(x)} = \sqrt{R_{(x)}} - (1-\beta)\arcsin\varphi_{(x)} - (1+\beta)\arcsin\nu_{(x)} \tag{19}$$

$$R_{(x)} = -x^2 + 2(1+\beta)x - (1-\beta)^2 \tag{20}$$

$$\varphi_{(x)} = \frac{(1+\beta)x - (1-\beta)^2}{2x\sqrt{\beta}} \tag{21}$$

$$\nu_{(x)} = \frac{1+\beta-x}{2\sqrt{\beta}}. \tag{22}$$

The estimation of eigenvalues requires the inversion of (18) that is yet unknown. However, a numerical inversion is always an option and requires matching the index j with corresponding x_j. So far only numerical inversion is known except for the special case where $\beta = 1$, i.e. square matrices $Q = \Theta$, it is possible to utilize the approximate inversion formula [7]

$$x_j\big|_{Q=\Theta} \approx \sin^2\left(\frac{\pi}{5}\left[\frac{2j-1}{2\Theta}\right] - \frac{\pi}{20}\ln\left(1 - \frac{20\pi}{63}\left[\frac{2j-1}{2\Theta}\right]\right)\right). \quad (23)$$

The discrepancy between increasing order index j and decreasing i is corrected using the equation

$$i = \Theta - j + 1. \quad (24)$$

The results of (4) for $\tilde{N} < \Theta$ is a complex and do not offer substantial advantages. Only the case $\tilde{N} = \Theta$ has a simple and convenient closed form that depends on $Q - \Theta$.

$$E\{\lambda^{-1}\} = \int_a^b \frac{p_{(x)}dx}{Q\sigma_H^2 x} = \frac{1}{\sigma_H^2(Q-\Theta)} \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii}^{-1} \quad (25)$$

$$E\{\lambda^{-2}\} = \int_a^b \frac{p_{(x)}dx}{\left(Q\sigma_H^2\right)^2 x^2} = \frac{Q}{\sigma_H^4(Q-\Theta)^3} \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii}^{-2} \quad (26)$$

$$E\{\lambda\} = \int_a^b \left(Q\sigma_H^2 x\right) p_{(x)}dx = Q\sigma_H^2 \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii} \quad (27)$$

III. WATERFILLING USING ESTIMATED EIGENVALUES

A. Linear Precoder Solutions

The waterfilling solutions were originally formulated with linear pre-coders and decoders [2] and [3]. The channel model for linear pre-coders and decoders can be with or without memory. However, the channel model with memory requires the addition of zero samples between transmissions. The channel models are described in great detail in [3].

The waterfilling matrix Φ is part of the pre-coder and requires the knowledge of the noise covariance matrix R_{nn}. The system eigendecomposition rely on the product $H^H R_{nn}^{-1} H$. The AWGN case reduces the covariance to $R_{nn} = \sigma_n^2 I$ which leads to the consequent eigendecomposition

$$H^H R_{nn}^{-1} H = H^H\left(\frac{1}{\sigma_n^2}\right)H = V\left(\frac{1}{\sigma_n^2}\Lambda\right)V^H. \quad (28)$$

However, the expression (3) will be utilized instead of (28) because of its simplicity. The noise variance σ_n^2 becomes a scaling factor that can be rather included in the pre-coder and decoder solutions.

The recovery $\hat{s}$ of the originally transmitted symbols s can be performed multiplying the received vector y in (2) by the linear matrix $G = \Gamma\Lambda^{-1}V^H H^H$. The matrix Γ is the receiver counterpart and a function of the waterfilling matrix Φ. Estimated symbols can be described by the equation

$$\hat{s} = \Gamma\Phi s + Gn. \quad (29)$$

Several solutions for Φ will be discussed, which depend on the eigenvalues of $H^H H$. The function $\Gamma(\Phi)$, required to complete the symbols estimation, is described in [3] and it is calculated once Φ is determined. For this reason it will not be analyzed it in this work.

B. Zero Forcing Solutions

Cases when $\Gamma = \Phi^\dagger$ are called Zero Forcing (ZF) solutions. The symbol † here is used to represent the pseudo inverse of a matrix. The solution for each diagonal element ϕ_{ii} in Φ can be generalized as [7]

$$\left|\phi_{ii}\right|^2 = \alpha\lambda_{ii}^{-1}. \quad (30)$$

Due to its simplicity, ZF solutions will be explained firstly clarifying some concepts of using estimated eigenvalues.

A common constraint used in finding waterfilling solutions is the total power P_0 that can be expressed as

$$P_0 = E\{\mathrm{tr}(z^H z)\} = \sigma_{ss}^2\,\mathrm{tr}(\Phi^H\Phi) \quad (31)$$

where z is defined in (1) and tr( ) is the trace function. The unitary matrix V does not contribute to the total power P_0. The solution for α, obtained from (30) and (31), under power constraint is

$$\alpha = \frac{P_0}{\sigma_{ss}^2\sum_{k=1}^{\Theta}\lambda_{kk}^{-1}} \quad (32)$$

and thus from (30)

$$\left|\phi_{ii}\right|^2 = \frac{P_0\lambda_{ii}^{-1}}{\sigma_{ss}^2\sum_{k=1}^{\Theta}\lambda_{kk}^{-1}}. \quad (33)$$

If H is known then $\hat{\phi}_{ii}$ can be calculated from (33). If H is not perfectly known but is an estimate, (33) will also result in estimated values $\hat{\phi}_{ii}$ with a precision that depends on the channel estimation. Moreover, if the elements in H follow a complex Gaussian distribution, then a method to estimate $\hat{\phi}_{ii}$ is to use the estimated eigenvalues obtained from random matrix theory, as described in Section II. The requirements to estimate the eigenvalues are the knowledge of the channel variance σ_H^2 and the size of H. The precision increases asymptotically with the size of H.

For square matrices $H_{N\times N}$ the estimation $\overline{\lambda}_{ii}$ can be calculated using (23) and scaling relation (12). Rectangular matrices $H_{M\times N}$ will require the use of (18) instead of (23) and some algorithm to find the inverse of (18). The Laguerre roots method (9) is also an option to estimate the eigenvalues.

Closed form estimation can be obtained for $\hat{\alpha}$ in the case of rectangular matrices, i.e. matrices where $|M-N|=Q-\Theta>0$. The eigenvalues summation in (32) can be considered the scaled expectation $\sum\lambda^{-1}\cong\Theta E\{\lambda^{-1}\}$ allowing us to find estimation for $\hat{\alpha}$ based on (25) as

$$\hat{\alpha}=\frac{\sigma_n^2}{\sigma_{ss}^2}\gamma\left(\frac{1}{\beta}-1\right), \quad (34)$$

where β was defined in (11). The variable

$$\gamma=\frac{\sigma_H^2 P_0}{\sigma_n^2} \quad (35)$$

reflects the relationship between transmitted power, channel and noise variances.

In general, the Signal to Noise Ratio (SNR) of the i^{th} sub-channel for linear pre-coders and decoders is known to be [3]

$$SNR_i=\frac{\sigma_{ss}^2}{\sigma_n^2}\left|\phi_{ii}^2\right|\lambda_{ii}. \quad (36)$$

The ZF solution has equal SNR for every sub-channel. An estimated sub-channel $S\hat{N}R^{ZF}$ for ZF solutions under total power constraint can be calculated from (30), $\hat{\alpha}$ in (34), and (36)

$$S\hat{N}R^{ZF}=\frac{\sigma_{ss}^2}{\sigma_n^2}\hat{\alpha}=\gamma\left(\frac{1}{\beta}-1\right). \quad (37)$$

A new design concept can now be introduced. According to (37), the SNR depends on β, or the relation between matrix dimensions Q and Θ. If the size of H can be freely chosen during system design (for example, by selecting the number of antennas), the dimensions Q and Θ can be adjusted to achieve a desired SNR. The design equation is

$$\frac{1}{\beta}=\frac{Q}{\Theta}=\left(\frac{S\hat{N}R^{ZF}}{\gamma}+1\right). \quad (38)$$

C. Waterfilling for Maximum Mutual Information on Parallel Gaussian Channels

The second case considered is a little more complicated. However, it is known to achieve a better performance for low SNR. The waterfilling solution for each element in Φ that maximizes the mutual information is

$$|\phi_{ii}|^2=\frac{\sigma_n^2}{\sigma_{ss}^2}\left(\left[\frac{P_0}{\tilde{N}\sigma_n^2}+\frac{1}{\tilde{N}}\sum_{k=1}^{\tilde{N}}\lambda_{kk}^{-1}\right]-\lambda_{ii}^{-1}\right)^+. \quad (39)$$

The eigenvalues in (39) as well as their summation can be obtained using any of the methods presented on Section II. Due to the Karush-Kuhn-Tucker Condition (KKTC) [1], settled by the function $(g)^+=\max[0,g]$, $\Theta-\tilde{N}$ elements ϕ_{ii} are zero. The determination of $\tilde{N}$ is made through an iterative process.

Analyzing (1), it can be seen that having zeroed values ϕ_{ii} will nullify corresponding symbols in s. In order to avoid this problem, all the eigenvalues need to comply with

$$\lambda_{ii}>\left[\frac{P_0}{\Theta\sigma_n^2}+\frac{1}{\Theta}\sum_{k=1}^{\Theta}\lambda_{kk}^{-1}\right]^{-1}. \quad (40)$$

Notice that if the minimum eigenvalue $\lambda_{\Theta\Theta}$ complies with (40), then the other eigenvalues will also do. A way to avoid zero elements in Φ is to allow enough transmission power P_0. Another possibility, which will be better described in the following paragraphs, is to guarantee enough difference $Q-\Theta$. This second method requires that the dimensions of $H_{M\times N}$ can be freely selected. It is known that the minimum eigenvalue is almost sure (a.s.) larger than a [13], as defined in (14) and scaled by (12). Using (25), (12), (14) and the inequality $x_{\Theta\Theta}\geq a$ to solve the inequality (40), results in

$$\frac{Q}{\Theta}\geq\left(\frac{1}{9}\left|1+\frac{4\sqrt[3]{\gamma}}{\xi_1}+\frac{\xi_1}{\sqrt[3]{\gamma}}\right|^2\right) \quad (41)$$

where

$$\xi_1=\sqrt[3]{27-8\gamma+3\sqrt{3(27-16\gamma)}} \quad (42)$$

The result obtained from (39) establishes the inequality relationship between Q and Θ that complies with KKTC.

Differently from the ZF solutions, the actual waterfilling produce a different SNR_i for each i^{th} sub-channel. Using (36) and (39) the SNR for Maximum Mutual Information (MMI) is

$$SNR_i^{MMI} = \left[\frac{P_0}{\tilde{N}\sigma_n^2} + \frac{1}{\tilde{N}}\sum_{k=1}^{\tilde{N}}\lambda_{kk}^{-1}\right]\lambda_{ii} - 1. \quad (43)$$

For the cases of rectangular matrices, $Q > \Theta$, and if all the eigenvalues comply with (40), the closed form equation for the estimation $S\hat{N}R_i^{MMI}$, using (25), is

$$S\hat{N}R_i^{MMI} = \left[\frac{\gamma}{\beta} + \frac{1}{1-\beta}\right]\frac{\lambda_i}{Q\sigma_H^2} - 1. \quad (44)$$

Furthermore, defining the average estimated SNR as

$$\begin{aligned} S\bar{N}R^{MMI} &= \frac{1}{\Theta}\sum_{i=1}^{\Theta} S\hat{N}R_i^{MMI} \\ &= \left[\frac{\gamma}{\beta} + \frac{1}{1-\beta}\right]\left(\frac{1}{Q\sigma_H^2}\right)\left(\frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_i\right) - 1 \end{aligned} \quad (45)$$

and considering (27), the closed form for the average SNR becomes

$$S\bar{N}R^{MMI} = \frac{\gamma}{\beta} + \frac{1}{1-\beta} - 1. \quad (46)$$

The design solution to find the relation between Q and Θ can be obtained from (46) and results in a quadratic equation with solution

$$\frac{1}{\beta} = \frac{Q}{\Theta} = \frac{1}{2}\left(1 + \frac{S\bar{N}R^{MMI}}{\gamma} + \sqrt{\left(1 - \frac{S\bar{N}R^{MMI}}{\gamma}\right)^2 - \frac{4}{\gamma}}\right). \quad (47)$$

D. Waterfilling for Maximum Information Rate

This last case is the most complex of the three. The waterfilling solution for each element in Φ is

$$|\phi_{ii}|^2 = \frac{\sigma_n^2}{\sigma_{ss}^2}\left(\left[\frac{\frac{P_0}{\tilde{N}\sigma_n^2} + \frac{1}{\tilde{N}}\sum_{k=1}^{\tilde{N}}\lambda_{kk}^{-1}}{\frac{1}{\tilde{N}}\sum_{k=1}^{\tilde{N}}\lambda_{kk}^{-\frac{1}{2}}}\right]\lambda_{ii}^{-\frac{1}{2}} - \lambda_{ii}^{-1}\right)^+. \quad (48)$$

Estimated eigenvalues can be used to solve (48) using the methods of Section II. In order to avoid zero elements in Φ it is required that every eigenvalue complies with

$$\lambda_{ii} > \left[\frac{\frac{P_0}{\Theta\sigma_n^2} + \frac{1}{\Theta}\sum_{k=1}^{\Theta}\lambda_{kk}^{-1}}{\frac{1}{\Theta}\sum_{k=1}^{\Theta}\lambda_{kk}^{-\frac{1}{2}}}\right]^{-\frac{1}{2}}. \quad (49)$$

The problem to find a minimum for the matrix dimension Q from (49) is the solution for $E\{\lambda^{-\frac{1}{2}}\}$ which includes elliptical integrals, as shown in Appendix B. Instead of the original solution, two approximations will be used

$$E\left\{\lambda^{-\frac{1}{2}}\right\} \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii}^{-\frac{1}{2}} \cong \frac{1}{\sqrt{\sigma_H^2 Q}\sqrt[3]{1-\beta}}, \quad (50)$$

$$E\left\{\lambda^{-\frac{1}{2}}\right\} \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii}^{-\frac{1}{2}} \cong \frac{1}{\sqrt{\sigma_H^2 Q}}. \quad (51)$$

The reason for (50) and (51) is that Q becomes the predominant value as the difference $Q - \Theta$ increases, i.e. $\beta \to 0$ as $Q \to \infty$. We have found that $(1-\beta)^{-1/3}$ is a good approximation for $E\{x^{-\frac{1}{2}}\}$, as shown in Appendix B, which then applying (12) results in (50). Approximation (51) is less accurate but also simpler than (50). An approximate solution for the inequality (49) can now be made following the procedure used in Section III-C and (51)

$$\frac{Q}{\Theta} \geq \left|\frac{\sqrt[3]{\frac{2}{3}\gamma}}{\xi_2} + \frac{\xi_2}{\sqrt[3]{18\gamma}}\right|^2 \quad (52)$$

where

$$\xi_2 = \sqrt[3]{9 + \sqrt{3\left(27 - 4\gamma^2\right)}} \quad (53)$$

and the variable γ is defined as in (35).

The SNR_i for Maximum Information Rate (MIR) for each i^{th} sub-channel, from (36) and (48) is

$$SNR_i^{MIR} = \left[\frac{\frac{P_0}{\Theta\sigma_n^2} + \frac{1}{\Theta}\sum_{k=1}^{\Theta}\lambda_{kk}^{-1}}{\frac{1}{\Theta}\sum_{k=1}^{N}\lambda_{kk}^{-\frac{1}{2}}}\right]\lambda_{ii}^{\frac{1}{2}} - 1. \quad (54)$$

In the case of rectangular matrices $Q > \Theta$, and if all the eigenvalues comply with (49), the closed form equation for the estimation $S\hat{N}R_i^{MIR}$ using (50) is

$$S\hat{N}R_i^{MIR} = \frac{\sqrt[3]{1-\beta}}{\sqrt{\sigma_H^2 Q}}\left[\frac{\gamma}{\beta} + \frac{1}{1-\beta}\right]\lambda_{ii}^{\frac{1}{2}} - 1. \quad (55)$$

The average value $S\bar{N}R^{MIR}$ requires now finding $E\{\lambda^{\frac{1}{2}}\}$ which solution also includes elliptical integrals. Proceeding similar to (50), see also Appendix B, the approximation for $E\{\lambda^{\frac{1}{2}}\}$ will be set as

$$E\left\{\lambda^{\frac{1}{2}}\right\} \cong \frac{1}{\Theta}\sum_{i=1}^{\Theta}\lambda_{ii}^{\frac{1}{2}} \cong \sqrt{\sigma_H^2 Q}\sqrt[9]{1-\beta}. \quad (56)$$

Using (55) and (56), an approximate closed form for $S\overline{N}R^{MIR}$ can be written as

$$S\overline{N}R^{MIR} = (1-\beta)^{4/9}\left[\frac{\gamma}{\beta}+\frac{1}{1-\beta}\right]-1. \qquad (57)$$

The inversion of (57) is very complex. For this reason, numerical inversion methods are used if it is necessary to find Q/Θ based on $S\overline{N}R^{MIR}$ and γ.

IV. Numerical Results and Discussion

The conditions chosen for the simulations were extracted from [14], where measurements tell us that a typical value for the channel variance is $\sigma_{HH}^2 = 1$. Symbols and noise are assumed to be independent, white, and Gaussian distributed with variances $\sigma_{ss}^2 = 1$ and σ_{nn}^2, respectively. Several other parameters are set up according to the specific type of simulation. This section shows three different numerical simulations that illustrate the results presented in this work.

A. The Waterlevel

The waterfilling cases MMI and MRI allocate the transmission power P_0 to achieve optimal system performance. The term waterfilling was chosen in literature because power is allocated in a way similar to filling a recipient with water [1]. The water reaches a certain level (the waterlevel) leaving the rest of the recipient unfilled. A nice explanation of the waterlevel problem is presented in [4]. In the case of (39) and (48) some of the elements are zeroed because the waterlevel is not high enough to cover all ϕ_{ii}. The problem of finding the number $\tilde{N}$ of non-zero elements ϕ_{ii} is referred as the waterlevel problem.

A closed form estimation for the waterlevel is very difficult to be obtained because it requires modifying (25), (26) and (27) into an integral with arbitrary upper limit, similarly to (17). Until now, a closed form solution for the waterlevel is not available in the literature.

Figures 1 and 2 compare the average waterlevel of square random matrix $H_{M\times M}$ against the case where the matrix size can be changed by decreasing the smallest dimension Θ. The curve labeled "average Θ" shows the average size for the smaller dimension Θ of $H_{M\times N}$, $\Theta = \min(M,N)$. In both cases, MMI and MIR, the plots show that (41) and (52) are valid inequalities for the waterlevel problem, i.e. the KKTC is met. Notice that the waterlevel has a better sub-channel availability than changing the size of $H_{M\times N}$. Figure 3 shows a comparison between the two lower waterlevel limits. Notice that the MRI allows more non-zero elements ϕ_{ii} than the MMI. Figure 4 illustrates the case when it is desired to increase the size of H by increasing Q. The MMI solution increases faster than MIR as γ decreases.

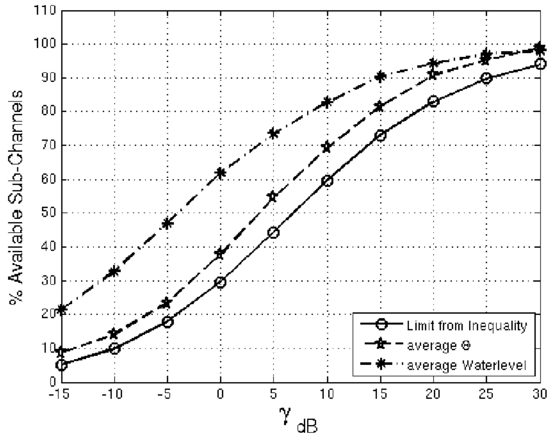

Figure 1. Sub-Channel Availability for Maximum Mutual Information.

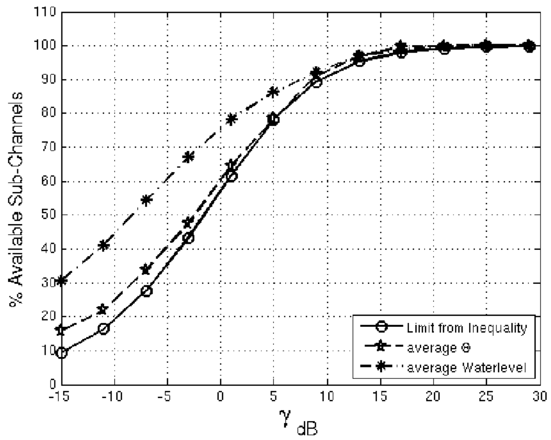

Figure 2. Sub-Channel Availability for Maximum Information Rate.

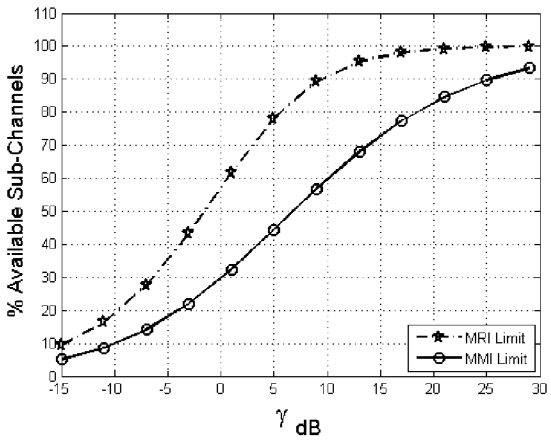

Figure 3. Waterlevel Comparison MRI and MMI.

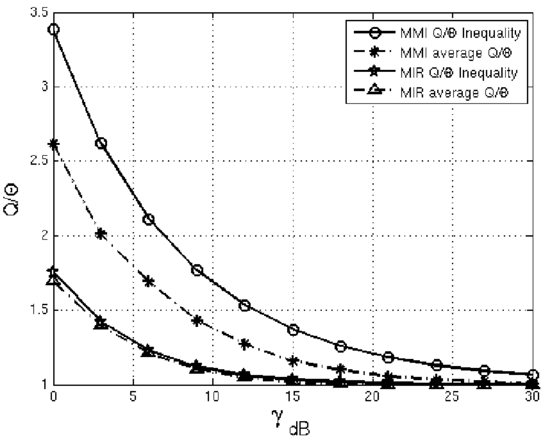

Figure 4. Ratio Q/Θ compared to γ.

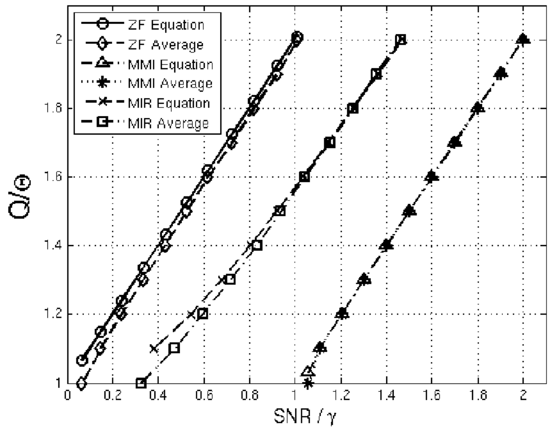

Figure 5. Ratio Q/Θ required for an objective SNR/γ.

B. The Signal to Noise Ratio

The SNR for ZF solutions SNR^{ZF} is the same for each sub-channel and is equal to its average $\overline{SNR}^{ZF}$. There is a linear relationship between SNR^{ZF}/γ a

nd Q/Θ which is valid except for $Q=\Theta$. Since the design objective is to find Q/Θ based on an average SNR, it was more useful to plot Q/Θ vs. SNR/γ, shown in Figure. The MMI and MIR required a value for γ large enough so that the KKTC is met; the value chosen was $\gamma = 30dB$. The theoretical curve for MIR was obtained by numerical inversion of (57).

As it can be seen in Figure 5, as Q/Θ approaches one, the numerical curve labeled as "Average" deviates from the theoretical prediction curve labeled as "Equation". The prediction (37) is that $SNR^{ZF}=0$ for $Q=\Theta$, but the square matrix is a special case and it is not well described by the Marčenko-Pastur law. However, the numerical average of SNR^{ZF} does approach zero for $Q=\Theta$. The MMI and MIR exhibit an undetermined $\overline{SNR}$ for $Q=\Theta$ because of the factor $(1-\beta)^{-1}$ present in (46) and (57).

The numerical and the theoretical curves are very close to each other for all cases ZF, MMI and MIR. This validates the theoretical approximations developed in this work, especially in obtaining (57) from (50), (51) and (56).

C. Symbol Error Rate Simulations

The Symbol Error Rate (SER) was simulated for a 64QAM modulation scheme. Since $\sigma_{HH}^2=1$, the variable γ defined in (35) is $\gamma = P_0/\sigma_n^2$ and can be seen as the system SNR, but different from the SNR_i experienced by each i^{th} sub-channel. The relation between γ and the sub-channel average $\overline{SNR}$ was given in the previous subsection.

The SER curves in Figure 6 are plotted against γ_{dB} because this quantity is a characteristic of the system, contrary to SNR_i that depends on the algorithm that is used. The numbers of transmission and reception antennas for this simulation are 4 and 6, respectively. This arrangement provides a rectangular channel matrix $H_{6\times4}$ that allows utilizing the closed form equations requiring $Q>\Theta$. Even when the MMI method provides a better relation $\overline{SNR}/\gamma$ than the MIR, as shown in Figure 5, it is clear in Figure 6 that this apparent advantage does not hold for SER.

There are two curves for each waterfilling case in Figure 6. The curves labeled as "Montecarlo" belong to the average SER calculated with exact match between MIMO $H_{6\times4}$ channel and its corresponding matrix Φ. The curves labeled "Equation" were calculated with Φ obtained from (39) and (48) but utilizing eigenvalues estimated with (18). Notice that both curves show similar performance. This indicates that once Φ has been estimated it is possible to use the same waterfilling coefficients ϕ_{ii} for a randomly changing MIMO channel.

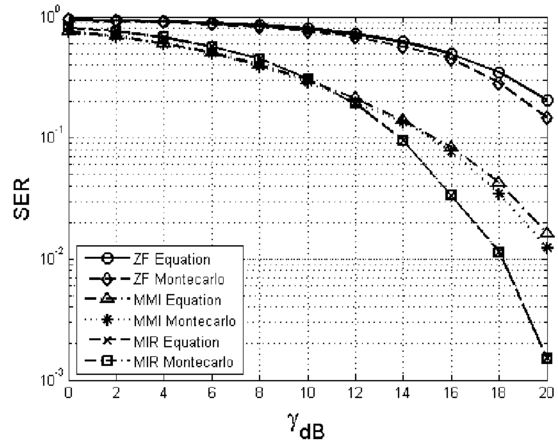

Figure 6. Symbol Error Rate (SER) vs. γ.

CONCLUSION

Several methods to estimate waterfilling solutions were developed. These methods take advantage of theoretical probability density functions known to be asymptotically accurate as the dimensions of the MIMO channel increases. However, even for a relatively small channel size, such as 6×4, a numerical example showed that estimated and exact SER are very close to each other. This result is especially important because it indicates that the channel coefficient $h_{i,j}$ may vary at random without extremely changing the system performance.

A method to select the size of a MIMO channel, or to predict its performance, was also presented. It was necessary to calculate the system performance in an average SNR sense, i.e. $\overline{SNR}$. Similar calculations allow us to estimate the channel size so that the KKTC is met for a random channel environment.

APPENDIX A EIGENVALUES' PROBABILITY DISTRIBUTION

The p.d.f. of the i^{th} single eigenvalue λ_{ii} is still an open research topic with the exception of the smallest eigenvalue $\lambda_{\Theta\Theta}$. Figure 7 shows the p.d.f. of the minimum eigenvalue for matrices of size 4×4, 4×5, and 4×6. The shape of the p.d.f. changes as the difference $Q-\Theta$ increases, which, in this case, is 0, 1 and 2, respectively. Exact equations for the p.d.f. of $\lambda_{\Theta\Theta}$ for rectangular random matrices can be found in [11].

An example, preferred for its simplicity, is the p.d.f. of square random matrices of size $\Theta\times\Theta$ with complex Gaussian entries. The p.d.f. is given by the exponential distribution [15]

$$p(\lambda_{\Theta\Theta})\big|_{H_{\Theta\times\Theta}} = \frac{\Theta}{\sigma_H^2} e^{-\lambda\Theta/\sigma_H^2} \tag{58}$$

with a mean value of

$$\overline{\lambda}_{\Theta\Theta} = \frac{\sigma_H^2}{\Theta} \tag{59}$$

and a variance

$$\sigma_{\lambda_{\Theta\Theta}}^2 = \left(\frac{\sigma_H^2}{\Theta}\right)^2 . \tag{60}$$

Notice that the square matrix case indicates a high probability of obtaining a zero valued eigenvalue. This is a problem for waterfilling calculations since they use the sum of the inverse eigenvalues $\sum_i \lambda_{ii}^{-m}$.

The inclusion of $\lambda_{\Theta\Theta}$ in the sum may result in a very large number leading to an ill posed problem. Unless the transmission power is high enough to compensate for the size of $\lambda_{\Theta\Theta}$, it is most likely that at least one coefficient ϕ_{ii} will be zeroed when obtaining a waterfilling solution.

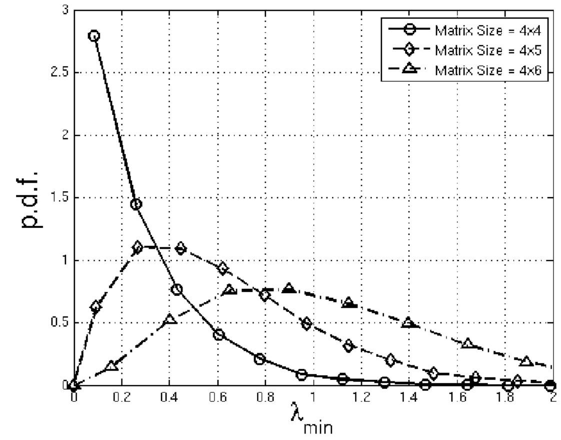

Figure 7. p.d.f. of $\lambda_{\Theta\Theta}$ ($\lambda_{\min}$) for several matrix sizes.

As it can be seen in Figure 7, it is unlikely for rectangular matrices $Q-\Theta>0$ to have a small probability of having a small $\lambda_{\Theta\Theta}$, i.e. an ill pose system will be rarely observed. In fact, the probability $p(\lambda_{\Theta\Theta}=0)$ is zero, indicating the advantage of using rectangular channel matrices.

Approximated p.d.f. for the other eigenvalues were proposed in [13]. In general, the p.d.f. of any eigenvalue can be approximated by a Gaussian distribution with its mean value centered at one of the associated Laguerre polynomial roots. Numerically obtained p.d.f. and their variance relation with the size of the system channel matrix were presented in [7]. The exact p.d.f. for an arbitrary eigenvalue is not yet available in the literature.

APPENDIX B APPROXIMATION OF SOME EXPECTATIONS OF EIGENVALUES

The calculation of (4) for integer exponents m, results in simple equations that depend on $1-\beta$, which is related to $Q(1-\beta)=Q-\Theta$. In (25), (26) and (27) this calculation simplifies to the integration of the normalized eigenvalue x^m multiplied by the p.d.f. (16). The simplicity of the expectations of (4) allows obtaining the closed form equations presented in this work.

A problem occurs when m is not an integer. For example, let us calculate the expectation of $x^{-\frac{1}{2}}$ as

$$E\left\{x^{-\frac{1}{2}}\right\} = \frac{1}{2\pi\beta}\int_a^b \frac{\sqrt{(b-x)(x-a)}}{x^{\frac{3}{2}}}dx . \tag{61}$$

The solution is

$$E\left\{x^{-\frac{1}{2}}\right\} = \frac{1}{\pi\beta}\frac{\sqrt{(b-x)(x-a)}}{x^{\frac{1}{2}}} \left[\frac{\sqrt{-1}\sqrt{(b-a)(x-a)\frac{x}{b}}\left[2\mathrm{E}\{\vartheta|\varpi\}-F\{\vartheta|\varpi\}\right]}{\left(\sqrt{\frac{x}{b}-1}(x-a)\right)}-1\right] \tag{62}$$

where E and F are elliptical integrals of the second and first kind,

$$\vartheta = \sqrt{-1}\,\text{arcsinh}\sqrt{\tfrac{x}{b}-1} \quad \text{and} \qquad \varpi = \frac{b}{b-a}. \qquad (63)$$

The obtained result is too complex to be applied in further analysis. Following the results for integer values of m, it was searched for an approximation of (61) based on the factor $1/(1-\beta)^r$, with the exponent r as the approximation parameter. Error plots for several values of r are presented in Figure 8. The best approximation is obtained for r = 1/3, i.e.

$$E\left\{ x^{-\frac{1}{2}} \right\} \approx \frac{1}{(1-\Theta/Q)^{\frac{1}{3}}}. \qquad (64)$$

After applying the normalization in (12), (50) was obtained. Yet, the result obtained with (64) is also hard to manipulate. The approximation in (51) is valid only because the number $(1-\beta) \to 1$ as $Q \to \infty$.

Figure 9 shows the error for the expectation $E\left\{ x^{\frac{1}{2}} \right\}$ approximated by $(1-\beta)^r$. The best value for the parameter is $r = 1/9$.

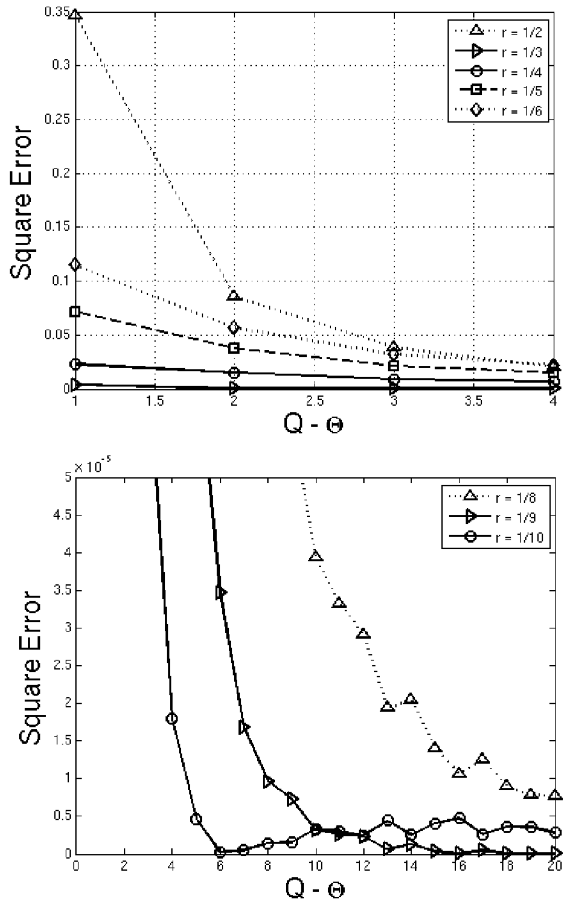

Figure 9. Approximation errors of the exponent r when approximating the expectation E{$x^{1/2}$} with $(1-\beta)^r$.

ACKNOWLEDGMENT

The authors wish to thank Dr. Edelman for his advice.

REFERENCES

[1] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: John Wiley & Sons, Inc., 1991.

[2] A. Scaglione, G. B. Giannakis, and S. Barbarossa, “Redundant Filter Bank Precoders and Equalizers part I: Unification and Optimal Designs,” IEEE Trans. Signal Process., vol. 47, pp. 1988–2006, Jul. 1999.

[3] A. Scaglione, P. Stoica, S. Barbarossa, G.B. Giannakis, and H. Sampath, “Optimal Designs for Space-Time Linear Precoders and Decoders,” IEEE Trans. Signal Process., vol. 50, no. 5, pp. 1051-1063, May 2002.

[4] D. Palomar and J. Fonollosa, “Practical Algorithms for a Family of Waterfilling Solutions,” IEEE Trans. Signal Process., vol. 53, no.2, pp. 686-697, Feb. 2005.

[5] Ü. Sakoğlu and A. Scaglione, “Asymptotic Capacity of Space-Time Coding for Arbitrary Fading Channels: a Closed Form Expression using Girko’s Law,” in Proc. IEEE Int. Acoustics, Speech, and Signal Processing, vol. 4, 7-11 May 2001, pp. 2509 - 2511.

[6] V. L. Girko, Circular Law: Theory of Probability and its Applications, Vol. 29, pp. 649-706, 1985.

[7] Vergara, Victor M.; Jordan, Ramiro; Barbin, Silvio E.; “An Asymptotic Waterfilling Solution for AWGN MIMO System Channels,” Vehicular Technology Conference, 2007. VTC-2007 Fall. 2007 IEEE 66th, Sept. 30 2007-Oct. 3 2007 Page(s): 620 – 624.

[8] H.F. Trotter, Eigenvalue distribution of large Hermitian matrices; Wigner’s semi-circle law and a theorem of Kac, Murdock, and Szegö, Adv. Math. 54 (1984) 67–82.

[9] V.A. Marčenko and L.A. Pastur. “Distribution of eigenvalues for some sets of random matrices,” Math USSR Sbornik, 1:457–483, 1967.

[10] N. Ullah, “On the expansion of the single eigenvalue probability density function”, Journal of Statistical Physics, Springer Netherlands, Volume 24, Number 3, March 1981, Pages 413-418.

[11] A. Edelman, “The Distribution and Moments of the Smallest Eigenvalue of a Random Matrix of Wishart Type” Linear Algebra and Its Applications, 1991, Volume: 159, Page(s): 55-80.

[12] H. Dette and L.A. Imhof, “Uniform approximation of eigenvalues in Laguerre and Hermite β-ensembles by roots of orthogonal polynomials”, Transactions of American Mathematical Society, Volume 359, Number 10, October 2007, Pages 4999–5018.

[13] J.W. Silverstein, “The Smallest Eigenvalue of a large Dimensional Wishart Matrix”, Ann. Prob. Volume 13, Number 4 (1985), Pages: 1364-1368

[14] Wallace, J.W. and Jensen, M.A., “Characteristics of measured 4×4 and 10×10 MIMO wireless channel data at 2.4-GHz”, Antennas and Propagation Society International Symposium, 2001. IEEE, Volume 3, 2001 Page(s): 96 – 99.

[15] A. Edelman. “Eigenvalues and Condition Numbers of Random Matrices” MIT PhD Dissertation, 1989.

Victor M. Vergara was born in Panamá, Republic of Panamá on December 18, 1972. He received the BS degree in Electrical Engineering from the Universidad de Panamá (UP), Republic of Panamá in 1997. In he completed the MS and PhD at the University of New Mexico, Albuquerque, NM, USA, in 2003 and 2006 respectively.

Currently he continues research work at the University of New Mexico, Albuquerque, NM, USA. Member of IEEE, he keeps participating in international conferences and publications.

Silvio E. Barbin was born in Campinas, SP Brasil on May 4, 1952. He received the BS degree in Electrical Engineering from Escola Politécnica da Universidade de São Paulo (USP), Brazil in 1974 and the MS degree and the PhD from the same institution. He was a visiting scholar at University of California at Los Angeles (UCLA) and a research professor at University of New Mexico at Albuquerque, NM, USA.

He worked for AEG-Telefunken in Germany and Brazil.and was the Technical Director of Microline MRF. He is presently a Professor at the Telecommunications and Control Engineering Department of EPUSP and a General Coordinator and Deputy Director in CTI – Centro de Tecnologia da Informação Renato Archer from the Ministry of Science and Technology in Brazil.

Prof Barbin is presently the Chapter Chair of COMSOC in the IEEE South Brazil Section, and the MTTS Coordinator and the Technical Activities Committee Chair for the IEEE –R9.

Ramiro Jordan is an Associate Professor at the University of New Mexico, Albuquerque, NM, USA. He received the PhD from Kansas State University. He is the Vice President of the Ibero American Science and Technology Consortium (ISTEC).