NTNU – Trondheim
Norwegian University of
Science and Technology

Non-Hermitian Random Matrix Theory for MIMO Channels

Burak Cakmak

Medical Technology
Submission date: June 2012
Supervisor: Ralf Reiner Müller, IET

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

Acknowledgements

First and foremost, let me start by thanking God for I believe without His willing, I could not write even these sentences. Moreover, I would like to express my sincere gratitude to my family for giving me continuous support and encouragement to study a master of science.

I owe a great appreciation to my advisor Prof Ralf Müller who gave the meaning to my master study. He created an excellent atmosphere therefore I finished this work with a minimum stress. He has shown a high tolerance to me in case when I lacked in knowledge on some topics. Moreover, he introduced me to Random Matrix and Information Theory that plays a major role in my thesis.

Particular gratitude belongs to Prof. Lars Lundheim for his support and introduction to Estimation and Detection Theory that plays an important role in my master study. Furthermore, I would like to express my appreciation to Solomon Tesfamicael and Muhammad Jafar thanks to their friendly support and advices. I would like to thank Prof. Maciej Nowak for insightful discussions and a great tutorial introduction into Non-Hermitian Free Probability while visiting The Norwegian University of Science and Technology. Very special thanks goes to my intimate friends Harun Armağan and Çağrı Kaplan thanks to their intellectual discussions.

Last but not the least, I would like to thank Papy Zefaniya and miss and mister Alic for proof-reading the manuscript.

Contents

List of Figures

Abstract

The propagation mechanism of signals for multiple input multiple output (MIMO) channels can be explained via a random matrix. Random matrix theory is a very powerful tool to understand behaviour of such channels and analyse their performance measure of MIMO systems. In this work we study:

The asymptotic eigenvalue distribution and the mutual information of a multiuser (MU) multiple-input multiple output (MIMO) channel with a certain fraction of users experiencing line-of-sight. It shows that the AED of the channel matrix decomposes into two separate bulks for practically relevant parameter choices and differs very much from the common assumption of independent identically distributed (iid) entries which induces the quarter circle law. This happens even without antenna correlation at either side of the channel. In order to tackle this problem the paper makes use of recent developments in free probability theory which allow to deal with complex-valued eigenvalue distributions of non-Hermitian matrices.

Moreover to understand behaviour of MIMO channels we derived asymptotic complex-valued eigenvalue distributions of practically relevant channels models by means of their respective *square equivalent* and *singular equivalent* of channel matrices.

Finally we derived an explicit mutual information formula which allows us calculate the mutual information (in general) analytically in high signal-to noise-ratio (SNR) regime for numerous practical important scenarios. Furthermore the numerical result shows that, high-SNR approximation draws reliable portrait even for quite moderate SNR level.

Chapter 1

Introduction

Wireless channels are usually not amiable as the wired one. Unlike wired channels that are stationary and predictable, wireless channels are extremely random.

The radio waves propagates through multi-paths in general. During that propagation the radio wave can impinge on an object whose dimension is larger than the signal wavelength λ that causes *reflection*, can impinge on a sharp object that causes *refraction*, can impinge on several objects whose dimensions are smaller than or comparable to λ that causes *scattering* or can propagate directly to receiver antenna without experiencing *reflection, refraction or scattering* that is called *line-of-sight* component [6]. Therefore multi-path channels cause an arbitrary time dispersion, attenuation, and phase shift, know as *fading*, in the received signal.

Fading causes poor signal quality thus bigger bit error probability. Suppose now, the replica of transmitted signals are sent through independent fading channels, it is highly probable to receive one of them that is not severely degraded by fading. If we imagine that, we have T antennas at the transmitter side, R antennas at the receiver side. The first motivation for Multiple Input Multiple Output (MIMO) system, was to investigate whether it is possible to create a number of independent fading channels with the maximum value TR to mitigate fading effects.

Moreover, with regards to technological development, high data rate demands increase rapidly. MIMO systems can be used to increase data rate by creating set of parallel channels. On the other hand, MIMO systems require advanced mathematical and physical frameworks. In order to design efficient and reliable MIMO system we first must understand propagation of signal thorough MIMO channels which are physical framework of MIMO systems. After the modelling of the channels, we must calculate practically relevant parameters or efficiency of algorithms and so on. This is the mathematical framework of MIMO systems.

MIMO systems first studied by Telatar [4] and Foschini & Gans [5] under the rich scattering assumption of the channel. With their pioneering works, it is realized that, with same power or bandwidth constraint, the data rate can be remarkably increased. Moreover the fading effects can be significantly decreased [1].These motivations made MIMO systems one of main research areas of wireless communications.

However with the increasing interest on MIMO system; with more realistic channel model assumptions; it is highly desired to understand more on MIMO channels and respective performance measures or equivalently the singular value distribution of channel matrix.

This is quite non-trivial mathematical problem in finite size analysis. On the other hand today we have a situation that the singular value distribution of a random matrix converges fast which brings much more interest on the asymptotic analysis of MIMO system.

In this work, we will analyse MIMO channels in a large system limit, as the number of antennas grows to infinity with a fixed ratio between the number of transmitting to receiving antennas. We will address a recent mathematical development *free probability* which was initiated by Voiculescu in the 1980s in [11] to study asymptotic singular value distribution of random matrices.

Free probability allows to infer the asymptotic eigenvalue distributions of sums or products of hermitan random matrices with known eigenvalue distribution provided that these Hermitian random matrices are free. This allowed us to deal with a great number of channel models in wireless communications and put the basis for the success of free probability in information theory of wireless channels, see e.g. [25], [26], [27].

In contrast, most of the engineering literature is greatly focused on Non-hermitian matrices, and we derived the empirical eigenvalue distribution of *square* or *singular*[2] equivalent of practically relevant channel matrices. For, we believe, it will bring a new point of view to understanding of behavior of MIMO channels and respective performance measures. To show capability of non-Hermitian approach we solved one open problem: MU-MIMO with line-of-sight channel model.

[1] Note that, there is a trade of between data rate and diversity rate. Both can't be increased simultaneously. We refer an interested reader to Chapter 10.14 in[62]

[2] Square and singular equivalent of rectangular random matrices are defined in fourth chapter.

Chapter 2

On the Capacity of MIMO Channels

In the second chapter we give a quick introduction onto MIMO system and some important facts on information theory are reviewed. Then we will revise Telatar work [4] and finally we review to convergence of random variables to start the asymptotic analysis of random matrices.

2.1 MIMO Channels

Consider a MIMO channel

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{1,t} & \cdots & h_{1,T} \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ h_{r,1} & \cdots & h_{r,t} & \cdots & hr,T \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ h_{R,1} & \cdots & h_{R,t} & \cdots & h_{R,T} \end{bmatrix} \tag{2.1}$$

where $h_{r,t}$ represents a fading gain (coefficient) corresponding to r^{th} receiving and t^{th} transmitting antenna pairs. The r^{th} column of $\mathbf{H}$ is often refered to as a spatial signature that represents the fading profile from the t^{th} transmitting antenna to receiving antennas arrays. The relative geometry of T spatial signatures determine distinguishability of signals at receiver side [3]. Thus, allowing to increase the data rate as sending the independent signals by creating independent set of independent spatial signatures (independent fading channels).

2.1.1 SU-MIMO

In single user (SU)-MIMO systems, the data stream that belongs to single user $\mathbf{x}$ is sent by multiple antennas, and transmitted through the channel $\mathbf{H}$ to the

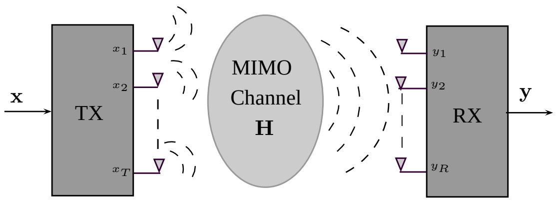


Figure 2.1: Single user MIMO: The signal vectors **x** is transmitted with multiple antennas thorough the **H** and received by multiple antennas.

receiver antennas (shown in Figure 1.1) can be modelled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.2}$$

where $\mathbf{H} \in \mathbb{C}^{R \times T}$, $\mathbf{x} \in \mathbb{C}^{T}$, $\mathbf{y}$, $\mathbf{n} \in \mathbb{C}^{R}$, the channel, channel input, channel output, and the normalized additive white Gaussian noise (AWGN) such that,

$$E\left(\mathbf{n}\mathbf{n}^{\dagger}\right) = \mathbf{I} \tag{2.3}$$

where **I** is identity matrix. Furthermore the transmitted signal has a constraint as

$$E\left(\mathbf{x}^{\dagger}\mathbf{x}\right) \le \gamma \tag{2.4}$$

where γ represents the signal-to-noise ratio (SNR)[1].

2.1.2 Multi-User MIMO

In this subsection we talk about two basic multi-user MIMO channel models: the MIMO multiple access channel (MAC) and the MIMO broadcast channel (BC).

MIMO-MAC Channel

In MIMO-MAC system we have multiple users on the transmitter side and each user have T_k antennas. In the receiver side (Base Station) we have R antennas.

Let $\mathbf{x}_k \in \mathbb{C}^{T_k}$ represents the transmitted data stream corresponding to k^{th} user. Moreover let the $R \times T_k$ matrix **H** represents the channels between the k^{th} user and the base station. Moreover in the MAC, each user is subject to an individual power constraint of P_k [3]

$$E\left[\mathbf{x}_k^{\dagger}\mathbf{x}_k\right] \le P_k, \quad \forall k \tag{2.5}$$

Now we define

$$\mathbf{x} = [\mathbf{x}_1, ..., \mathbf{x}_K] \tag{2.6}$$
$$\mathbf{H} = [\mathbf{H}, ..., \mathbf{H}_K] \tag{2.7}$$

[1] Note that, the noise is with normalized variance, thus the power constraint it equal SNR.

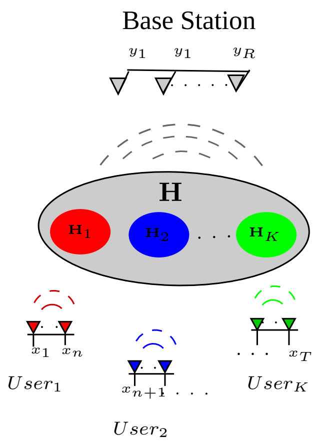


Figure 2.2: K users MIMO-MAC Channel.

Then the received signal can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.8}$$

where $\mathbf{H} \in \mathbb{C}^{R \times T}$, $\mathbf{x} \in \mathbb{C}^{T}$, $\mathbf{y}$, $\mathbf{n} \in \mathbb{C}^{R}$, the channel, channel input, channel output, and the normalized AWGN as in equation (2.3).

MIMO-BC Channel

In MIMO-BC system, we have multiple users on the receiver side and each user have R_k antennas. In the transmitter side (Base Station) we have T antennas.

Let $\mathbf{y}_k \in \mathbb{C}^{R_k}$ represents the received data stream corresponds to kth user. Moreover let the $T \times R_k$matrix $\mathbf{H}_k$ represents the channels between the base station and kth user. Now we define

$$\mathbf{y} = [\mathbf{y}_1, ..., \mathbf{y}_K] \tag{2.9}$$

$$\mathbf{H} = [\mathbf{H}, ..., \mathbf{H}_K] \tag{2.10}$$

Then the received signal can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.11}$$

where $\mathbf{H} \in \mathbb{C}^{R \times T}$, $\mathbf{x} \in \mathbb{C}^{T}$, $\mathbf{y} \in \mathbb{C}^{R}$, $\mathbf{n} \in \mathbb{C}^{R}$, the channel, channel input, channel output, and the normalized AWGN as in (2.3). Moreover the base station subject to an average power constraint

$$E\left[\mathbf{x}^{\dagger}\mathbf{x}\right] \leq P. \tag{2.12}$$

It is immediate that MIMO-BC and MIMO-MAC channels are identical. Obviously one is Hermitian of the other.

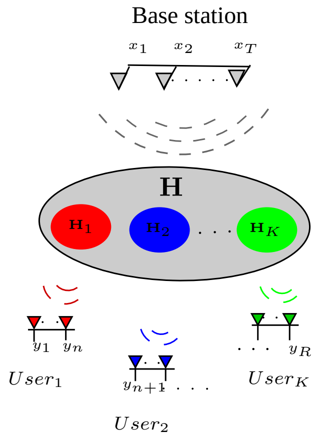


Figure 2.3: K users MIMO-BC channel.

2.2 On the Capacity of MIMO Channel

In the late 1940, Claude Shannon defined the concept of channel capacity as the maximum data rate which can be transmitted reliably over the channel. He proved that there is a coding scheme that performs at rate less then capacity [1]. *Calculating capacity of a channel has been the most fundamental problem of the communication theory since Shannon's result.*

2.2.1 Glancing Over Information Theory

In this subsection we quickly review some important facts from information theory.

Definition 1. *(i) The measure of randomness or the information content associated with a random variable $x \in X$ with a transmitted probability $p(x)$ is defined as*

$$I(x) \triangleq \log \frac{1}{p(x)}$$

(ii) The Shannon entropy can be written as average information content of x as

$$H(x) \triangleq E_x \left(\log \frac{1}{p(x)} \right) \tag{2.13}$$

i.e. average amount of information lost over the channel can be written as

$$H(x|y) \triangleq E_{x,y} \left(\log \frac{1}{p(x|y)} \right)$$

Corollary 1. *Mutual Information: The average amount of information at output of the channel can be written as*

$$I(x, y) \triangleq H(x) - H(x|y) \tag{2.14}$$

The mutual information can be also written as

$$I(x, y) \triangleq H(y) - H(y|x) \tag{2.15}$$

Theorem 1. *[1] Shannon proved that the channel capacity is*

$$C = \max_{p(x)} I(x, y) \tag{2.16}$$

Definition 2. *Relative Entropy: is a measure the distance of two distribution.*

$$\begin{aligned} D(p(x)||q(x)) &= E_p\left(\log \frac{p(x)}{q(x)}\right) \\ &= \int p(x) \log \frac{p(x)}{q(x)} d(x) \end{aligned} \tag{2.17}$$

Remark 1. *Let $p(x)$ and $q(x)$ are two different distribution of x. Then*

- $H(x_p) - H(x_q) = -D(p(x)||q(x))$.
- $D(p(x)||q(x)) \geq 0$.

Proof.

$$\begin{aligned} H(x_p) - H(x_q) &= -\left\{\int p(x) \log[p(x)]dx - \int p(x) \log[q(x)]dx\right\} \\ &= -\int p(x) \log[\frac{p(x)}{q(x)}]dx \\ &= -D(p(x)||q(x)) \end{aligned} \tag{2.18}$$

$$\begin{aligned} D(p(x)||q(x)) &= \int p(x) \log \frac{p(x)}{q(x)} d(x) \\ &\geq \int p(x)\left(1 - \frac{q(x)}{p(x)}\right) d(x) \\ &= 0 \end{aligned} \tag{2.19}$$

where $\ln \frac{1}{u} \geq 1 - u$ for $u > 0$. □

Example 1. *In the continuous case find a distribution $p(x)$ that maximizes the entropy with the constraints $E(x) = 0, E(x^2) = \sigma^2$.*

Solution 1. *We can maximize the entropy by using Lagrange multiplier as:*

$$\begin{aligned} J &= -\int p(x)\log(p(x)) + \lambda_1\left[-1+\int p(x)dx\right] + \lambda_2\left[-\sigma^2+\int x^2p(x)dx\right] \\ \frac{dJ}{dp} &= -\log p(r) - 1 + \lambda_1 + \lambda_2 x^2 \end{aligned}$$

Setting the final term to zero gives,

$$p(x) = e^{\lambda_1+\lambda_2x^2} \tag{2.20}$$

where Gaussian distribution is a solution, hence $p(x) \sim \mathcal{N}(0,\sigma^2)$.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}} \tag{2.21}$$

Example 2. *In the continuous case find a distribution* $p(\mathbf{x})$ *that maximizes the entropy with the constraints* $E(x_ix_j^*) = C_{ij}$ *with zero mean and* $\mathbf{x} \in \mathbb{C}^N$. *Then calculate its entropy.*

Solution 2. *As in the previous example, we can maximize the entropy by using Lagrange multiplier as:*

$$\begin{aligned} J &= -\int p(\mathbf{x})\log[p(\mathbf{x})]\mathbf{x} + \beta\left[-1+\int p(\mathbf{x})d\mathbf{x}\right] \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{i,j}\left[-C_{ij}+\int p(\mathbf{x})|x_{ij}|^2d\mathbf{x}\right] \qquad (2.22) \\ \frac{dJ}{dp} &= -\log p(\mathbf{x}) - 1 + \beta + \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{i,j}|x_{ij}|^2 \\ &= -\log p(\mathbf{x}) - 1 + \beta + \mathbf{x}^\dagger\Lambda\mathbf{x} \qquad (2.23) \end{aligned}$$

Setting the final term to zero gives,

$$p(\mathbf{x}) = exp(\beta + \mathbf{x}^\dagger\Lambda\mathbf{x}) \tag{2.24}$$

It is immediate that, multivariate normal distribution is a solution. Hence $p(\mathbf{x}) \sim \mathcal{N}(0,\mathbf{C})$ *maximizes the entropy given constraint.*

$$p(\mathbf{x}) = \frac{1}{\det \pi\mathbf{C}}e^{-\mathbf{x}^\dagger\mathbf{C}^{-1}\mathbf{x}} \tag{2.25}$$

Finally we can calculate the entropy as

$$\begin{aligned} H(\mathbf{x}) &= E\left(-\log p(\mathbf{x})\right) \\ &= E\left(\log\det[\pi\mathbf{C}] + \mathbf{x}^\dagger\mathbf{C}^{-1}\mathbf{x}\log e\right) \\ &= \log\det\pi\mathbf{C} + tr(\mathbf{C}^{-1}E(\mathbf{x}\mathbf{x}^\dagger))\log e \\ &= \log\det\pi\mathbf{C} + tr(\mathbf{C}^{-1}\mathbf{C})\log e \\ &= \log\det\pi\mathbf{C} + N\log e = \log\left[e^N\det\pi\mathbf{C}\right] \\ &= \log\det\pi e\mathbf{C} \qquad (2.26) \end{aligned}$$

since i.e $a^N\det[\mathbf{A}] = \det[a\mathbf{A}]$ *for a scalar value* a.

2.2.2 Capacity Formula of MIMO Channel

In the following we will address the wireless MIMO system described by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.27}$$

where $\mathbf{H} \in \mathbb{C}^{R\times T}$, $\mathbf{x} \in \mathbb{C}^{T}$, $\mathbf{y} \in \mathbb{C}^{R}$, $\mathbf{n} \in \mathbb{C}^{R}$, the channel, independent and identically distributed (iid) channel input, channel output, and the normalized additive white Gaussian noise (AWGN) $\mathbf{n} \sim \mathcal{N}(0, \mathbf{I})$. Moreover the covariance of channel input is,

$$E(\mathbf{x}\mathbf{x}^\dagger) = \mathbf{P} \tag{2.28}$$

where $\mathbf{P}$ is a $T \times T$ diagonal matrix. Then the mutual information between channel input $\mathbf{x}$ and channel output $\mathbf{y}$ reads,

$$\begin{aligned} I(\mathbf{x},\mathbf{y}) &= H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}) \\ &= H(\mathbf{y}) - H(\mathbf{H}\mathbf{x} + \mathbf{n}|\mathbf{x}) \\ &= H(\mathbf{y}) - H(\mathbf{n}) \end{aligned} \tag{2.29}$$

since the noise $\mathbf{n}$ is independent of the channel input $\mathbf{x}$. With (2.26) we have

$$\begin{aligned} I(\mathbf{x},\mathbf{y}) &= H(\mathbf{y}) - \log\det \pi e \mathbf{I} \\ &= H(\mathbf{y}) - (\log \pi e)^R \end{aligned} \tag{2.30}$$

Maximizing $I(\mathbf{x},\mathbf{y})$ is equivalent to maximizing $H(\mathbf{y})$. In example 2. we show that, maximum entropy can be achieved with normal distribution. Hence $\mathbf{y} \sim \mathcal{N}(0, \mathbf{C})$ where the covariance matrix can be written as

$$\begin{aligned} \mathbf{C} &= E\left((\mathbf{H}\mathbf{x} + \mathbf{n})(\mathbf{H}\mathbf{x} + \mathbf{n})^\dagger\right) \\ &= \mathbf{H}E(\mathbf{x}\mathbf{x}^\dagger)\mathbf{H}^\dagger + \mathbf{I} + \mathbf{H}E(\mathbf{x})E(\mathbf{n}^\dagger) + E(\mathbf{n})E(\mathbf{x}^\dagger)\mathbf{H}^\dagger \\ &= \mathbf{H}\mathbf{P}\mathbf{H}^\dagger + \mathbf{I} \end{aligned} \tag{2.31}$$

Then the capacity reads,

$$\begin{aligned} C(\mathbf{P}) &= \log\det\left[\pi e(\mathbf{I} + \mathbf{H}\mathbf{P}\mathbf{H}^\dagger)\right] - (\log \pi e)^R \\ &= \log\det\left[\mathbf{I} + \mathbf{H}\mathbf{P}\mathbf{H}^\dagger\right]. \end{aligned} \tag{2.32}$$

$C(\mathbf{P})$ is a random variable. Therefore we define ergodic (mean) capacity which is expectation of C over $\mathbf{H}$ as

$$C_e(\mathbf{P}) = \int C(\mathbf{P}) dP(\mathbf{H}) \tag{2.33}$$

$$= \int \log\det\left[\mathbf{I} + \mathbf{H}\mathbf{P}\mathbf{H}^\dagger\right] dP(\mathbf{H}). \tag{2.34}$$

With eigenvalue decomposition

$$\mathbf{H}\mathbf{P}\mathbf{H}^\dagger = \mathbf{U}\mathbf{X}\mathbf{U}^\dagger \tag{2.35}$$

where $\mathbf{U}$ is a unitary matrix and consist of eigenvector of $\mathbf{HPH}$ which is $\mathbf{X} = diag(x_1, ..., x_R)$. Then the capacity (ergodic) expression becomes,

$$C_e(\mathbf{P}) = \int \log\det\left[\mathbf{I} + \mathbf{X}\right] dP_{\mathbf{HPH}^\dagger}(\mathbf{X}) \tag{2.36}$$

$$= \sum_{r=1}^{R} \int \log(1 + x_r) dP_{\mathbf{HPH}^\dagger}(x_r) \tag{2.37}$$

$$= R \int \log(1 + x) dP_{\mathbf{HPH}^\dagger}(x) \tag{2.38}$$

where x denotes an arbitrary eigenvalue of $\mathbf{HPH}^\dagger$. For notational convenience we express ergodic capacity as

$$\frac{C_e(\mathbf{P})}{R} = \int \log(1 + x) dP_{\mathbf{HPH}^\dagger}(x) \tag{2.39}$$

Moreover in this work we address on the equal power case such that

$$\mathbf{P} = \gamma\mathbf{I} \tag{2.40}$$

known as *mutual information* can be expressed

$$\mathcal{I}(\gamma) = \int \log(1 + \gamma x) dP_{\mathbf{HH}^\dagger}(x) \tag{2.41}$$

The non-trivial question is how to find the eigenvalue distribution of $\mathbf{HPH}^\dagger$. Even in the simplest channel model it is still quite non-trivial question. In the following we will address calculation of the ergodic capacity of Rich scattering MIMO channel by assuming no channel state information at the transmitter side.

2.2.3 Capacity of Rich Scattering MIMO Channel

Suppose, there is a channel state information at the receiver but no channel state information at the transmitter. Then we will have equal power strategy, thus

$$E(\mathbf{x}\mathbf{x}^\dagger) = \gamma \tag{2.42}$$

where γ represents SNR. Note that, we assume the noise $\mathbf{n}$ is normalized AWGN such that, $\mathbf{n} \sim \mathcal{N}(0, \mathbf{I})$. Then the capacity formula can be written as in (2.41):

$$\frac{C_e(\gamma)}{R} = \int \log(1 + \gamma x) dP_{\mathbf{HH}^\dagger}(x). \tag{2.43}$$

Theorem 2. *[2]* Wishart Matrices: *Let the entries of $R \times T$ matrix $\mathbf{H}$ be iid with zero mean and variance $1/R$. Then the unordered joint eigenvalue distribution of $\mathbf{HH}^\dagger$ is given by*

$$p_{\mathbf{HH}^\dagger}(\mathbf{x}) = \frac{Z}{R!} \exp\left[-\sum_{i=1}^{R} x_i\right] \prod_{i=1}^{R} x_i^{T-R} \prod_{l=i+1}^{R} (x_i - x_l)^2 \tag{2.44}$$

where $R!$ term in the dominator is due to allowing to all possible ordering.

In the following we will revise Talatar works who calculate the capacity of MIMO rich scattering MIMO channels where the entries of channel $\mathbf{H}$ be assumed iid with zero mean variance $1/R$ as defined in Theorem 2.

We start with Theorem 2: By using a column operation on determinant (or row operation since $\det \mathbf{A} = \det \mathbf{A}^\dagger$): multiplication of a column by a constant multiplies the determinant by that constant, and we can simply calculate marginal eigenvalue distribution. Note that during calculation we do not take into account the normalization constant, we will compensate it at the final step:

$$\begin{aligned}
p_{\mathbf{HH}^\dagger}(\mathbf{x}) &\sim \exp\left[-\sum_{i=1}^{R} x_i\right] \prod_{i=1}^{R} x^{T-R} \prod_{l=i+1}^{R} (x_i - x_l)^2 \\
&= \left\{ \prod_{i=1}^{R} \exp\left[-x_i/2\right] x_i^{(T-R)/2} \prod_{l=i+1}^{R} (x_i - x_l) \right\}^2 \\
&= \left\{ \prod_{i=1}^{R} \exp\left[-x_i/2\right] x^{(T-R)/2} \begin{vmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_R \\ \vdots & & \vdots \\ x_1^{R-1} & \cdots & x_R^{R-1} \end{vmatrix} \right\}^2 \\
&= \begin{vmatrix} x_1^{(T-R)/2} e^{-x_1/2} & \cdots & x_R^{(T-R)/2} e^{-x_1/2} \\ \vdots & & \vdots \\ x_1^{(T+R)/2-1} e^{-x_1/2} & \cdots & x_R^{(T+R)/2-1} e^{-x_1/2} \end{vmatrix}^2 \\
&= \begin{vmatrix} \phi_1(x_1) & \cdots & \phi_R(x_1) \\ \vdots & & \vdots \\ \phi_R(x_1) & \cdots & \phi_R(x_K) \end{vmatrix}^2
\end{aligned} \tag{2.45}$$

Now we use the combinatoric definition of determinant i.e. consider the 3×3 matrix $\mathbf{A}$ then,

$$\det \mathbf{A} = \mathbf{A}_{1,1}\mathbf{A}_{2,2}, \mathbf{A}_{3,3} + \sum_{\sigma \in P_3; \sigma \neq \{1,2,3\}} (-1)^{|\sigma|} \prod_i \mathbf{A}_{i,\sigma(i)}$$

where P_N is the set of all permutation of the set $\{1, 2...N\}$. Moreover if the reordering $\sigma \to (1, 2...N)$ can be done with even number of pairwise switching then, $(-1)^{|\sigma|} = 1$ else $(-1)^{|\sigma|} = -1$[8].
With the definition of determinant, (2.45) reads

$$p_{\mathbf{HH}^\dagger}(\mathbf{x}) \sim \sum_{\sigma \in P_N} \sum_{\omega \in P_N} (-1)^{|\sigma|} (-1)^{|\omega|} \prod_i \phi_{\sigma(i)}(x_i) \phi_{\omega(i)}(x_i) \tag{2.46}$$

In addition, ϕ_i are orthonormal polynomials such that [2]:

$$\int \phi_i(x) \phi_j(x) d\lambda = \delta_{ij} \tag{2.47}$$

Now we can easily calculate the marginal eigenvalue distribution with the orthogonality principle (2.47)

$$
\begin{aligned}
p_{\mathbf{HH}^\dagger}(x_1) &= \int_{\mathcal{R}^{R-1}} p_{\mathbf{HH}^\dagger}(\mathbf{x}) dx_2..dx_R \\
&\sim \sum_{\sigma \in P_R} \sum_{\omega \in P_R} (-1)^{|\sigma|+|\omega|} \int_{\mathcal{R}^{R-1}} \prod_i \phi_{\sigma(i)}(x_i)\phi_{\omega(i)}(x_i) dx_2..dx_K \\
&\sim \sum_{\sigma(1)=1}^{R} \sum_{\omega(1)=1}^{R} (-1)^{|\sigma(1)|+|\omega(1)|} \phi_{\sigma(1)}\phi_{\omega(1)}(x_1)\delta(\sigma(1)-\omega(1)) \\
&= \sum_{i=1}^{R} \phi_i^2(x_1). \qquad (2.48)
\end{aligned}
$$

Finally the normalization constant can be computed as,

$$
\begin{aligned}
p_{\mathbf{HH}^\dagger}(x_1) &= \frac{1}{\sum_{i=1}^{R} \int \phi_i^2(x_1) dx_1} \sum_{i=1}^{R} \phi_i^2(x_1) \\
&= \frac{1}{R} \sum_{i=1}^{R} \phi_i^2(x_1). \qquad (2.49)
\end{aligned}
$$

To evaluate (2.49), one can use the following relation [2],

$$
\begin{aligned}
\phi_{k+1}(x) &= x^{(T-R)/2} \left[\frac{k!}{(k+R-T)!} \right]^{1/2} \frac{1}{k!} x^{T-R} e^x \frac{d^k}{dx^k} [e^{-\lambda} \lambda^{T-R+k}] \\
&= x^{(T-R)/2} \left[\frac{k!}{(k+T-R)!} \right]^{1/2} L_k^{T-R}(x), \quad k = 0, 1..R-1. \qquad (2.50)
\end{aligned}
$$

where $L_k^{T-K}(x)$ is the generalized Laguerre polynomial of order k. Thus plugging (2.50) in 2.43 we have:

Theorem 3. *[4] Let entries of the $R \times T$ matrix $\mathbf{H}$ be iid with zero mean with variance $1/R$. With assuming channel state information at the receiver side but no channel state information at transmitter the capacity of a channel $\mathbf{H}$ is*

$$
\frac{C_e(\gamma)}{R} = \int \log \left[1 + \frac{\gamma}{R} x \right] \sum_{k=1}^{R} \frac{k!}{(k+T-R)!} \left[L_k^{(T-R)} \right]^2 x^{T-R} e^{-x}(x) dx. \qquad (2.51)
$$

2.3 Asymptotic Analysis

Assuming MIMO channels with iid entries is very limited case in practice. We may need to assign a variance profile to channel matrix, or the channel may have a line-of sight component, or we might take into account interference, we may need to generalize the scattering richness of channel, the channel might be correlated in a certain level, or strongly correlated by means of Kronecker channel model, or any combination of these situation and the others.

There are very few mathematical result even the joint probability density function (pdf) form of eigenvalue distribution. Moreover we need to marginalize joint pdf where in iid case we used the machinery of orthogonal polynomials (2.47). However the orthogonality principle does not work for such as scenarios mentioned above.

On the other hand, the asymptotic analysis of random matrices quite trivial comparing to finite size analysis allowing us to work in many practically relevant channel models. Moreover the eigenvalue distribution of most of random matrices converges very fast, thus asymptotic analysis draws reliable portrait even for quite moderate number of antennas such as 4×4 or sometimes 2×2 [9].

2.3.1 Converges of Random Variable

Suppose you are interested in a sequence of random variables $\{X_1, X_2.....\}$ defined in a probability space Ω and want to determine statistic of the sequence. In the following, we want to address the statistic of X_N when number of sequence goes infinity called convergence of random variables:

Definition 3. *Let $\{X_1, X_2\}$ be a sequence of random variables. Let X be a random variable. Then $\{X_N\}$ is said to converge to X in distribution if*

$$\lim_{N\to\infty} Pr(X_N \leq x) = Pr(X \leq x). \tag{2.52}$$

A generic example of converging in distribution is the central limit theorem:

Example 3. *Suppose that $\{X_1, X_2, ..\}$ are iid random variables with the mean μ and variance σ^2. Define,*

$$Y_N = \frac{1}{\sqrt{\sigma^2 n}} \sum_{i=1}^{n} (X_i - \mu) \tag{2.53}$$

Then we have,

$$\lim_{N\to\infty} Pr(Y_N \leq y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{t^2}{2}\right) dt \tag{2.54}$$

Definition 4. *Let $\{X_1, X_2\}$ be a sequence of random variables. Let X be a random variable. Then $\{X_N\}$ is said to converge to X in probability if*

$$\lim_{N\to\infty} Pr(|X_N - X| > \epsilon) = 0 \tag{2.55}$$

for any $\epsilon > 0$.

Example 4. *[10]Let* $\{X_1, ...X_n\}$ *are iid Uniform random variables on the interval* $[0, 1]$*. Define*

$$Y_N = \max(X_1, ...X_N) \tag{2.56}$$

such that,

$$Pr(Y_n \leq y) \leq y^N \tag{2.57}$$

Then for any $0 < \epsilon < 1$ *we have*

$$\lim_{N\to\infty} Pr(|Y_N - 1| > \epsilon) = \lim_{N\to\infty} Pr(Y_n < 1 - \epsilon) \tag{2.58}$$

$$= \lim_{N\to\infty} (1 - \epsilon)^N = 0 \tag{2.59}$$

Definition 5. *Let* $\{X_1, X_2\}$ *be a sequence of random variables. Let* X *be a random variable. Then* $\{X_N\}$ *is said to converge to* X *almost surely if*

$$Pr\left(\lim_{N\to\infty} X_N = X\right) = 1. \tag{2.60}$$

Let us give an example regarding to the thesis topic:

Example 5. *Consider a* $N \times N$ *random matrix* $\mathbf{H}$*. Recall that,*

$$tr(\mathbf{H}) = \sum_{n=1}^{N} x_n \tag{2.61}$$

where x_n *represents eigenvalues of* $\mathbf{H}$*. Then define*

$$Y_N = \frac{1}{N} tr(\mathbf{H}^n) \tag{2.62}$$

Thus we have,

$$Pr\left(\lim_{N\to\infty} Y_N = m_{n,\mathbf{X}}\right) = 1. \tag{2.63}$$

where $m_{n,\mathbf{H}}$ *is* $n^t h$ *order moment of the asymptotic eigenvalue distribution of* $\mathbf{H}$*.*

Note that, almost sure convergence implies convergence in probability, convergence in probability implies convergence in distribution. But the converse is not true in general.

Chapter 3

Random Matrix & Free Probability Theory

In this chapter we will address how to find the empirical eigenvalue distribution of sum and product of *Hermitian* random matrices as the size goes to infinity. First, fundamental results on random matrix theory are introduced by means of *Stieltjes Transform.* Then *Free Probability* as proposed by Voiculesco in [11] which allows us to find sum and product of *free* random matrices is introduced.

Due to our main concern is the asymptotic eigenvalue distribution of a Hermitian random matrix, the moments of eigenvalue distribution will have a key role. When we talk about moments of asymptotic eigenvalue distribution we always use the following definition of normalized trace operator:

Definition 6. *Consider a $R \times R$ hermitian matrix $\mathbf{A}$ and define*

$$\phi(\mathbf{A}) = \lim_{R \to \infty} \frac{1}{R} tr(\mathbf{A}). \tag{3.1}$$

Remark that, nth moment of $\mathbf{A}$ can be expressed as $\phi(\mathbf{A}^n)$.

3.1 Random Matrix Theory

In this section, we first introduce practically relevant random matrices and their asymptotic empirical eigenvalue distribution to analyse the statistic of random matrices.

The moments of most of important random matrices can be expressed in terms of *Narayana-Catalan* numbers defined as follows [7] :

Definition 7 (Narayana-Catalan Numbers)**.** *Narayana numbers is defined as*

$$N_{n,k} = \frac{1}{n} \binom{n}{k} \binom{n}{k-1}. \tag{3.2}$$

The definition of Catalan Number is,

$$C_n = \sum_{k=1}^{n} N_{n,k}. \tag{3.3}$$

Moreover these numbers have a key meaning in combinatoric which will be pointed out in the next section.

The empirical eigenvalue distribution of important types of random matrices are examined in the following theorems:

Theorem 4 (Girko Full Circle Law [12])**.** *Let the entries of the $R \times R$ matrix $\mathbf{H}$ be independent identically distributed entries with zero mean and variance $1/R$. Then the empirical eigenvalue distribution of $\mathbf{H}$ converges almost surely to the limit given by*

$$p_{\mathbf{H}}(z) = \begin{cases} \frac{1}{\pi} & |z| < 1 \\ 0 & elsewhere \end{cases} \tag{3.4}$$

as $R \to \infty$.

Theorem 5 (Wigner Semi Circle Law [13])**.** *Let the entries of the $R \times R$ hermitian matrix $\mathbf{H}$ be independent identically distributed entries with zero mean and variance $1/R$. Then the empirical eigenvalue distribution of $\mathbf{H}$ converges almost surely to the limit given by*

$$p_{\mathbf{H}}(x) = \frac{1}{2\pi}\sqrt{4 - x^2} I_{\{-2,2\}} \tag{3.5}$$

as $R \to \infty$.

In free probability semicircle element distribution is analogous with normal distribution in classical probability theory. Furthermore, the even moments of semicircle distribution can be expressed with *Catalan number* as [43]

$$\phi\left(\mathbf{H}^{2n}\right) \quad = \quad C_n. \tag{3.6}$$

Theorem 6 (Deformed Quarter Circle Law [14])**.** *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be independent identically distributed entries with zero mean and variance $1/R$. Then the empirical singular value distribution of $\mathbf{H}$ converges almost surely to the limit given by*

$$p_{\sqrt{\mathbf{H}\mathbf{H}^{\dagger}}}(x) = \max(0, 1-\beta)\delta(x) + \frac{\sqrt{4\beta - (x^2 - 1 - \beta)^2}}{x\pi} I_{\{|1-\sqrt{\beta}|, 1+\sqrt{\beta}\}} \tag{3.7}$$

as $R, T \to \infty$ with $\beta = T/R$ fixed.
Moreover the transformation random variable X as $Y = X^2$ reads

$$p_Y(y) = \frac{1}{2\sqrt{y}} p_X(\sqrt{y}) \tag{3.8}$$

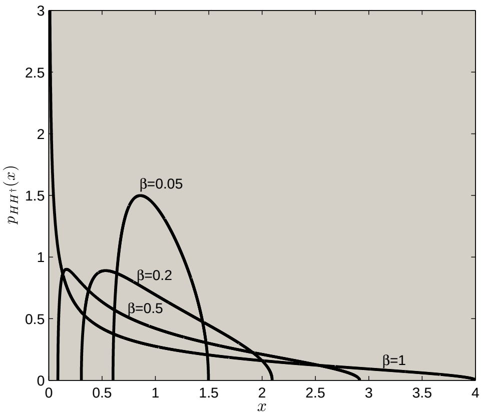


Figure 3.1: Deformed Quarter Circle Law (eigenvalues) with the fixed ratio $\beta = T/R$. Note in case $\beta \leq 1$, then the pdfs have some zero measure that are skipped.

This gives

$$p_{\mathbf{HH}^\dagger}(x) = \max(0, 1-\beta)\delta(x) + \frac{\sqrt{4\beta - (x-1-\beta)^2}}{2x\pi} I_{\{(1-\sqrt{\beta})^2, (1+\sqrt{\beta})^2\}}. \quad (3.9)$$

which is known as the Marchenko-Pastur distribution and its moments are given by

$$\phi\left((\mathbf{HH}^\dagger)^n\right) = \sum_{k=1}^{n} N_{n,k}\beta^k. \quad (3.10)$$

.

Deformed quarter circle law has very important role in many practical fields. As an example in MIMO system, Rayleigh i.i.d. channel $\mathbf{H} \in \mathbb{C}^{R\times T}$ is a simple application of the deformed quarter circle law and the mutual information reads the following expression:

$$\lim_{R\to\infty} \frac{1}{R} \log\det\left(\mathbf{I} + \gamma\mathbf{HH}^\dagger\right) = \lim_{R\to\infty} \frac{1}{R} tr\log(\mathbf{I} + \gamma\mathbf{HH}^\dagger) \quad (3.11)$$

$$= \phi(\mathbf{I} + \frac{1}{\sigma^2}\mathbf{HH}^\dagger) = \int_{(1-\sqrt{\beta})^2}^{(1+\sqrt{\beta})^2} \log(1+\gamma x) \frac{\sqrt{4\beta - (x-1-\beta)^2}}{2x\pi} dx. \quad (3.12)$$

as $T, R \to \infty$ with the ratio $\beta = T/R$ fixed [1].

[1] A closed form expression for (3.12) will be given at the end of the section in (3.182).

On the Unitary Matrices

Recall, $R \times R$ matrix $\mathbf{U}$ is called unitary if

$$\mathbf{U}\mathbf{U}^{\dagger} = \mathbf{U}^{\dagger}\mathbf{U} = I \tag{3.13}$$

Unitary matrices has central role in free probability to define set of free matrices[2]. Here we present some important class of random matrices in term of their unitarily property.

Theorem 7 (Haar Distribution [15])**.** *Let the entries of the $R \times R$ matrix $\mathbf{H}$ be independent identically complex distributed entries with zero mean and finite positive variance. Define*

$$\mathbf{U} = \mathbf{H}(\mathbf{H}^{\dagger}\mathbf{H})^{-\frac{1}{2}} \tag{3.14}$$

Then the empirical eigenvalue distribution of $\mathbf{T}$ converges almost surely to the limit given by

$$p_{\mathbf{U}}(z) = \frac{1}{2\pi}\delta(|z| - 1) \tag{3.15}$$

as $R \to \infty$.

Note that, the eigenvalues of a unitary matrix lie on the complex unit circle. A Haar matrix is special class of unitary matrix, where its eigenvalues are uniformly distributed on the complex unit circle.

Remark 2. *Let the entries of the $R \times R$ matrix $\mathbf{H}$ be independent identically complex distributed entries with zero mean and finite positive variance. Then $\mathbf{H}$ can be decomposed as*

$$\mathbf{H} = \mathbf{U}\mathbf{Q} \tag{3.16}$$

where $\mathbf{U}$ is a Haar matrix and $\mathbf{Q}$ fulfils same conditions as needed for the quarter circle law.

Definition 8. *If a hermitian random matrix $\mathbf{H}$ has same distributed with*

$$\mathbf{U}\mathbf{H}\mathbf{U}^{\dagger} \tag{3.17}$$

for any unitary matrix $\mathbf{U}$ independent of $\mathbf{H}$, then the matrix $\mathbf{H}$ is called unitarily invariant.

Lemma 1. *[31] A unitarily invariant $\mathbf{X}$ can decomposed*

$$\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{U} \tag{3.18}$$

with $\mathbf{U}$ is a Haar matrix independent of the diagonal matrix $\mathbf{\Lambda}$.

Lemma 2. *[31],[33] Consider a function*

$$\mathbf{Y} = g(\mathbf{X}) \tag{3.19}$$

with unitarily invariant matrix $\mathbf{X}$ as an input and a hermitian matrix $\mathbf{Y}$ as an output. Then the matrix $\mathbf{Y}$ is also unitarily invariant.

[2] The definition of freeness will be given in the next section.

Example 6. *[26] A matrix fullfills same conditions as needed for the semi circle law or deformed quarter circle law, or Haar distribution is unitarily invariant.*

Definition 9. *[26] If the joint distribution of the entries of a $R \times T$ matrix $\mathbf{X}$ has equal to the joint distribution of the entries of a matrix $\mathbf{Y}$ such that*

$$\mathbf{Y} = \mathbf{U}\mathbf{H}\mathbf{V}^{\dagger} \tag{3.20}$$

then the matrix $\mathbf{X}$ is called bi-unitarily invariant random matrix.

Note that, an identity matrix is also a unitary matrix. Then one can consider bi-unitarily invariant $R \times T$ random matrix $\mathbf{H}$ such that whose singular value distribution is invariant both by left and right a unitary matrix product.

Example 7. *Let the set $\{\mathbf{H}_1, \cdots, \mathbf{H}_N\}$ consists of independent standard Gaussian matrices with the size of $\mathbf{H}_n$ is $T_n \times T_{n-1}$. Moreover define a matrix*

$$\mathbf{H} = \prod_{n=1}^{N} \mathbf{H}_n. \tag{3.21}$$

Then $\mathbf{H}$ is bi-unitarily invariant.

Theorem 8. *[31] A square random matrix $\mathbf{H}$ is a bi-unitarily-invariant, if it can be decomposed as*

$$\mathbf{H} = \mathbf{U}\mathbf{Y} \tag{3.22}$$

where$\mathbf{U}$ is Haar matrix and independent of unitarily invariant positive definite matrix $\mathbf{Y}$.

3.1.1 Stieltjes Transform

There is no doubt that, the most useful transform in Random matrix theory is Stieltjes transform. One can propose that, fully understanding the capability of Stieltjes transform is equivalent to understanding half of random matrix theory.

In classical probability theory we take the Fourier transform (or characteristic function) to determine moments of random variable X or vice versa:

$$\begin{aligned} \mathcal{F}_X(s) &\triangleq \int e^{sx} dP_X(x) & (3.23)\\ &= \int \sum_{n=0}^{\infty} \frac{(sx)^n}{n!} dP_X(x) = \sum_{n=0}^{\infty} \frac{s^n}{n!} \int x^n dP(x) & (3.24)\\ &= \sum_{n=0}^{\infty} \frac{m_{X,n}}{n!} s^n. & (3.25)\end{aligned}$$

where $m_{X,n}$ nth order moment of the random variable X. It is reasonable to apply the same method for a hermitian random matrix $\mathbf{A}$

$$\mathcal{F}_{\mathbf{A}}(s) = \int e^{sx} dP_{\mathbf{A}}(x) = \sum_{n=0}^{\infty} \frac{\phi(\mathbf{A}^n)}{n!} s^n. \tag{3.26}$$

Beside we can determine the nth moment of $\mathbf{A}$ as

$$\phi(\mathbf{A}^n) = \frac{d^n}{ds^n}\mathcal{F}_{\mathbf{A}}(s)\Big|_{s=0} \tag{3.27}$$

However in contrast to the classical probability theory, Fourier transform method is not fruitful in random matrix theory. On the other hand consider a hermitian matrix $\mathbf{A}$, define a function

$$\begin{aligned} G_{\mathbf{A}}(s) &= \sum_{n=0}^{\infty} \frac{\phi(\mathbf{A}^n)}{s^{n+1}} && (3.28)\\ &= \sum_{n=0}^{\infty} \int \frac{dP_{\mathbf{A}}(x)}{x^{-n}s^{n+1}} = \int \sum_{n=0}^{\infty} \frac{dP_{\mathbf{A}}(x)}{x^{-n}s^{n+1}} && (3.29)\\ &= \int \frac{dP_{\mathbf{A}}(x)}{s-x}. && (3.30) \end{aligned}$$

This method was proposed by Stieltjes in 1984 [34] to determine a unknown probability distribution of random variable given its moments

Definition 10. *Consider the hermitian random matrix* $\mathbf{A}$*, then the Stieltjes transform of the matrix* $\mathbf{A}$ *is defined as*

$$G_{\mathbf{A}}(s) \triangleq \int \frac{dP_{\mathbf{A}}(x)}{s-x} \tag{3.31}$$

with $\Im s > 0$.

In random matrix theory the Stieltjes transform is analogous Fourier transform (characteristic function) in classical probability theory. In fact, it is immediate to see the following relation between Stieltjes transform and Fourier transform

$$\frac{d^n}{ds^n}\frac{G_{\mathbf{A}}(s)}{s}\Big|_{s=0} = n!\, \frac{d^n}{ds^n}\mathcal{F}_{\mathbf{A}}(s)\Big|_{s=0} \tag{3.32}$$

Moreover, recall the scaling property of Fourier transform,

$$\mathcal{F}_{c\mathbf{A}}(s) = \frac{1}{c}\mathcal{F}_{\mathbf{A}}\left(\frac{s}{c}\right) \tag{3.33}$$

which is identical for Stieltjes transform as

$$\begin{aligned} G_{c\mathbf{A}}(s) &= \int \frac{dP_{\mathbf{A}}(x)}{s-cx} && (3.34)\\ &= \frac{1}{c}\int \frac{dP_{\mathbf{A}}(x)}{\frac{s}{c}-x} && (3.35)\\ &= \frac{1}{c}G_{\mathbf{A}}\left(\frac{s}{c}\right) && (3.36) \end{aligned}$$

with $c \in \mathbb{R}$.

Theorem 9 (Stieltjes inversion formula [35]). *If $a \leq x \leq b$ is a continuity points of $Pr(a \leq x \leq b)$, then we have*

$$Pr(a \leq x \leq b) = -\frac{1}{\pi}\lim_{y\to 0}\int_a^b \Im G(x+jy)dx \tag{3.37}$$

Obviously the probability density function can be simply obtained from the Stieltjes transform as

$$p(x) = -\frac{1}{\pi}\lim_{y\to 0}\Im G(x+jy) \tag{3.38}$$

Lemma 3. *Consider the $R \times T$ matrix $\mathbf{X}$. Then we have the following relation:*

$$G_{\mathbf{X}\mathbf{X}^\dagger}(s) = \beta G_{\mathbf{X}^\dagger\mathbf{X}}(s) + \frac{\beta - 1}{s} \tag{3.39}$$

with $\beta = T/R$.

Sum and Product of Random Matrices

In this subsection, we present some important results regarding to sum and product of random matrices in term of Stieltjes transform:

Theorem 10 (Sum of Random Matrices I [37]). *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be independent identically distributed with zero mean and variance $1/R$. Let the $R \times R$ hermitian matrix $\mathbf{X}$ with an asymptotic empirical eigenvalue distribution converges almost surely to the limit $P_{\mathbf{X}}(x)$. Furthermore let $\mathbf{Y} = diag(y_1, ...y_T)$ be a $T \times T$ diagonal matrix the empirical eigenvalue distribution converges almost surely to a limit distribution $P_{\mathbf{Y}}(y)$. Moreover, let the matrices $\mathbf{H}, \mathbf{X}, \mathbf{Y}$ be jointly independent. Define,*

$$\mathbf{S} = \mathbf{X} + \mathbf{H}\mathbf{Y}\mathbf{H}^\dagger. \tag{3.40}$$

Then, the empirical eigenvalue distribution of $\mathbf{S}$ converges almost surely to a limit distribution whose Stieltjes transform satisfies

$$G_{\mathbf{S}}(s) = G_{\mathbf{X}}\left(s + \beta\int\frac{ydP_{\mathbf{Y}}(y)}{yG_{\mathbf{S}}(s) - 1}\right) \tag{3.41}$$

as $R, T \to \infty$ with $\beta = T/R$ fixed.

In the case of the entries of the matrix $\mathbf{H}$ independent identically complex Gaussian distributed with zero mean and variance $1/R$, the matrix $\mathbf{Y}$ not necessarily diagonal, it can be a Hermitian matrix and independent of $\mathbf{H}$. The reason why a hermitian matrix $\mathbf{Y}$ can be decomposed as

$$\mathbf{Y} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger \tag{3.42}$$

where $\mathbf{U}$ is a unitary matrix. $\mathbf{H}$ and its right-unitary product $\mathbf{HU}$ have the same distribution. Hence, $\mathbf{Y}$ can be replaced by any hermitian matrix $\mathbf{UYU}$ for $\mathbf{U}$ unitary matrix. [27].

With $\mathbf{X} = \mathbf{0}$, theorem 10 has crucial role in MIMO system determining the mutual information with power profile for the users even for a complicated model. As well in Kronecker channel model is one of the known application of theorem 10.

Theorem 11 (Products of Random Matrices I[36]). *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be independent identically distributed entries with zero mean and variance $1/R$. Moreover, let the $R \times R$ hermitian matrix X be independent of $\mathbf{H}$, with an asymptotic empirical eigenvalue distribution converges almost surely to the limit $P_{\mathbf{X}}(x)$. Furthermore, let*

$$\mathbf{P} = \mathbf{H}\mathbf{H}^{\dagger}\mathbf{X}. \tag{3.43}$$

Then, the empirical eigenvalue distribution of $\mathbf{P}$ converges almost surely to a limit distribution whose Stieltjes transform satisfies

$$G_{\mathbf{P}}(s) = \int \frac{dP_{\mathbf{X}}(x)}{s - x(1 - \beta + \beta s G_{\mathbf{P}}(s))} \tag{3.44}$$

as $R, T \to \infty$ with $\beta = T/R$ fixed.

Theorem 12 (Product of Random Matrices II [18]). . *Let the random matrices $\mathbf{H}_n, \forall n$ be size of $T_n \times T_{n-1}$ with independent identically complex Gaussian distributed entries with zero mean and variance $1/K_n$. Let*

$$\mathbf{H} = \prod_{n=1}^{N} \mathbf{H}_n. \tag{3.45}$$

Furthermore define,

$$\rho_n = \frac{T_n}{T_N}. \tag{3.46}$$

Then, the eigenvalue distribution distributions of $\mathbf{H}\mathbf{H}^{\dagger}$ converges almost surely to a limit distribution whose Stieltjes transform satisfies

$$G_{\mathbf{H}\mathbf{H}^{\dagger}}(s) \prod_{n=1}^{N} \frac{sG_{\mathbf{H}\mathbf{H}^{\dagger}}(s) + 1 - \rho_{n-1}}{\rho_n} - sG_{\mathbf{H}\mathbf{H}^{\dagger}}(s) = 1 \tag{3.47}$$

as $T_n \to \infty$ but the ration ρ_n fixed for all $0 \le n \le N$.

Theorem 12 is very relevant in many practical field. i.e. it is immediate to see, it is key formula to find mutual information for a certain number of successive scattering level in MIMO channels. Moreover it can be directly plug into *Layered Relay Network* [16] to explore its performance measure such a Capacity.

Lemma 4. *[Mutual Information Lemma [19]] Consider a $R \times T$ matrix $\mathbf{H}$, then we have*

$$\frac{\mathcal{I}(\gamma)}{R} = \int_{\frac{1}{\gamma}}^{\infty} G_{\mathbf{H}\mathbf{H}^{\dagger}}(-s) + \frac{1}{s} ds. \tag{3.48}$$

as $T, R \to \infty$, with the ratio $\beta = T/R$ fixed.

Proof. Let

$$\mathcal{I}(\gamma) \triangleq \lim_{R\to\infty} \frac{1}{R} \log\det\left(I + \gamma \mathbf{H}\mathbf{H}^{\dagger}\right) = \int \log(1 + \frac{x}{s}) dP_{\mathbf{H}\mathbf{H}^{\dagger}}(x) \Big|_{s=\frac{1}{\gamma}} \tag{3.49}$$

Thus,

$$\frac{d\mathcal{I}(\gamma)}{ds} = \int \frac{x dP_{\mathbf{H}\mathbf{H}^{\dagger}}(x)}{s(s+x)}\Big|_{s=\frac{1}{\gamma}} = \int \frac{dP_{\mathbf{H}\mathbf{H}^{\dagger}}(x)}{(s+x)} - \frac{1}{s}\Big|_{s=\frac{1}{\gamma}} \tag{3.50}$$

$$= -G_{\mathbf{H}\mathbf{H}^{\dagger}}(-s) - \frac{1}{s}\Big|_{s=\frac{1}{\gamma}} \tag{3.51}$$

Then we find,

$$\mathcal{I}(\gamma) = \int_{\frac{1}{\gamma}}^{\infty} G_{\mathbf{H}\mathbf{H}^{\dagger}}(-s) + \frac{1}{s} ds. \tag{3.52}$$

□

It is often to get an analytical expression for the asymptotic eigenvalue distribution of singular values of MIMO Channels is not possible. This is why, in order find the mutual information in MIMO system, the mutual information lemma is quite useful.

3.2 Free Probability Theory

Consider Random matrices as a non-commutative random variable in general. Then, in contrast to probability theory, we must define the variables in a matrix-valued probability space or a non-commutative probability space which changes whole frame of (classical) probability theory.

In this section a new mathematical field free probability theory which was initiated by Voiculescu in the 1980s [39] will be introduced. The theory is a magic of infinite dimension. i.e. definition of freeness. It applies for non-commutative random variables. Therefore, the main feature of the theory is for random matrix theory..

From Classical probability to Free probability

In (3.25), we have already mentioned the Fourier transform of a distribution can be expressed in term of power series of moments as

$$\mathcal{F}_X(s) = \sum_{n=0}^{\infty} \frac{m_{X,n}}{n!} s^n. \tag{3.53}$$

Moreover let define a function

$$r_X(\omega) \triangleq \log \mathcal{F}_X(\omega) \tag{3.54}$$

where $r(\omega)$ is called *the cumulant generating function* for a corresponding distribution, which can either be expressed as [45]

$$r_X(\omega) = \sum_{n=1}^{\infty} \frac{k_{X,n}}{n!} \omega^n \tag{3.55}$$

where $k_{X,n}$ is called nth order cumulant of the random variable X.

Remark 3. *Let X and Y are independent random variables. Further define a random variable $X + Y$. Then the cumulant generating function of $X + Y$ is,*

$$r_{X+Y}(\omega) = r_X(\omega) + r_Y(\omega), \tag{3.56}$$

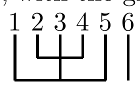
or equivalently

$$k_{X+Y,n} = k_{X,n} + k_{Y,n}. \tag{3.57}$$

The relation between cumulants and moments can be expressed via combinatorics. Let $P(n)$ is set of all partition (permutation) of $\{1, 2, \cdots n\}$. Let π is a partition of this set as

$$\pi = \{B_1, \cdots B_r\} \tag{3.58}$$

where B_i is the block of π connects some elements in the partition π. i.e. $\pi \in P(6)$, $\pi = \{(1, 3, 5), (2, 4), (6)\}$, with the graphical representation



We denote the size of $B_i \in \pi$ as $|B_i|$,(i.e $B_1 = (1, 3, 5) \Rightarrow |B_1| = 3$.).

Definition 11. *Consider a random variables X. Then nth order moment of X can be expressed as*

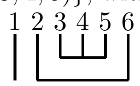
$$m_n = \sum_{\pi \in P(n)} \prod_{B_i \in \pi} k_{|B_i|}. \tag{3.59}$$

where k_m is called mth order cumulant.

On the other hand consider a non-commutative probability space where the random variables do not commute in general. As example let the space be a matrix probability space and let the non-commutative random variable be a hermitian random matrix $\mathbf{X}$. Moreover, in a similar way, let $N(n)$ is set of all non-crossing permutation of $\{1, 2, \cdots n\}$. Let π is a non-crossing partition of this set

$$\pi = \{B_1, \cdots B_r\} \tag{3.60}$$

where B_i is the block of π connects some elements in the(non-crossing) partition π. i.e. $\pi \in N(6)$, $\pi = \{(1), (2, 6), (3, 4, 5)\}$, with the graphical representation



Definition 12. *Consider a random matrix* $\mathbf{X}$*. Then the moment of asymptotic eigenvalue distribution can be expressed*

$$\phi(\mathbf{X}^n) = \sum_{\pi \in N(n)} \prod_{B_i \in \pi} \kappa_{|B_i|}. \tag{3.61}$$

where κ_n *called nth order* free cumulant.

The single difference between two summation is: *(3.59) is over all possible partition* $P(n)$, on the other side (3.61) is over all possible non-crossing partition $N(n)$, this single difference changes the full frame and brings a new theory called free probability theory.

Furthermore, even though we haven't yet defined the concept of *freeness*, it is reasonable conjecture that: If $\mathbf{A}$ and $\mathbf{B}$ are free random variables or random matrices, then their *free cumulants* are additive as

$$\kappa_{\mathbf{A}+\mathbf{B},n} = \kappa_{\mathbf{A},n} + \kappa_{\mathbf{B},n} \tag{3.62}$$

Indeed this conjecture was proven by Voiculescu [39]. However we must warn the reader, having an intuition about free provability by making analogy with probability theory is wrong. We will come this issue later.

As a consequence of this subsection, let's mention the combinatoric meaning of *Narayana and Catalan*:

- The number of a-non-crossing partition of π_r of the set $\{1, 2, c \ldots, n\}$ where π_r includes r blocks can be expressed in term of *Narayana number* as

$$N_{n,r} = \frac{1}{n}\binom{n}{r}\binom{n}{r-1} = \sum_{\pi_r \in NC(n)} \tag{3.63}$$

- the number of non-crossing partition of the set $\{1, 2, \cdots, n\}$ can be expressed in terms of *Catalan Number* as

$$C_n = \sum_{\pi \in NC(n)} = \sum_{r=1}^{n} N_{r,n} \tag{3.64}$$

3.2.1 Freeness

The counterpart of independence in classical probability theory is freeness in free probability. But the concept of freeness comes from infinite dimension. Therefore it is not easy to imagine freeness intuitively. Beside, later on we will show one generic example to understand freeness in terms of eigenvectors of matrices. Now, let's focus on definition of freeness in polynomial approach.

Consider two random matrices $\mathbf{A}$ and $\mathbf{B}$. Then focus on finding the eigenvalue distribution of $\mathbf{A} + \mathbf{B}$ and $\mathbf{AB}$. To find eigenvalue distribution we must know for all $n \in \mathbb{N}$ moments

$$\phi((\mathbf{A}+\mathbf{B})^n) \quad \textbf{and} \quad \phi((\mathbf{AB})^n) \tag{3.65}$$

Thus we need to know the non-commutative polynomial of ϕ.

Non-commutative Polynomials

To see the difference between non-commutative space and commutative space, we will show two examples in the sense of a possible polynomial term of both two commutative variable such as real numbers and two non-commutative variable such as matrices. First let x,y be real numbers, then we define $P_n(x, y)$ as a sum of all possible n^{th} order polynomials in two variables x and y is given by [18]

$$P_n(x, y) \triangleq \sum_{i=1}^{\infty} \alpha_i x^{l_i} y^{m_i} : \quad l_i, m_i \in \{0, 1...n\} \wedge \alpha_i \in \mathbb{R} \forall i. \tag{3.66}$$

i.e. for $P_2(x, y)$, it is sum of nine terms

$$P_2(x, y) = \alpha_1 x^2 y^2 + \alpha_2 x^2 y + \alpha_3 x y^2 + \alpha_4 x^2 + \alpha_5 xy + \alpha_6 y^2 + \alpha_7 x + \alpha_8 y + \alpha_9$$

On the other hand, let $\mathbf{A}$, and $\mathbf{B}$ be real matrices. Then we define $N_n(x, y)$ as a sum of all possible n^{th} order polynomials in two non-commutative variables $\mathbf{A}$ and $\mathbf{B}$ is given by

$$\begin{aligned} N_n(\mathbf{A}, \mathbf{B}) \quad \triangleq \quad & \left\{ \sum_{i=1}^{\infty} \alpha_i \prod_{k=1}^{n} \mathbf{A}^{l_{i,k}} \mathbf{B}^{m_{i,k}} : \right. \\ & \left. \sum_{i=1}^{\infty} l_{i,k}, \sum_{i=1}^{\infty} m_{i,k} \in \{0, 1...n\} \wedge \alpha_i \in \mathbb{R} \forall i. \right\} \end{aligned} \tag{3.67}$$

i.e. for $N_2(\mathbf{A}, \mathbf{B})$, it is sum of nineteen terms

$$\begin{aligned} N_2(\mathbf{A}, \mathbf{B}) \quad = \quad & \alpha_1 \mathbf{A^2 B^2} + \alpha_2 \mathbf{A B^2 A} + \alpha_3 \mathbf{ABAB} + \alpha_4 \mathbf{BABA} + \\ & \alpha_5 \mathbf{B A^2 B} + \alpha_6 \mathbf{B^2 A^2} + \alpha_7 \mathbf{A^2 B} + \alpha_8 \mathbf{ABA} + \alpha_9 \mathbf{AB}^2 + \\ & \alpha_{10} \mathbf{BA}^2 + \alpha_{11} \mathbf{BAB} + \alpha_{12} \mathbf{B^2 A} + \alpha_{13} \mathbf{A^2} + \alpha_{14} \mathbf{BA} + \\ & \alpha_{15} \mathbf{BA} + \alpha_{16} \mathbf{B^2} + \alpha_{17} \mathbf{A} + \alpha_{18} \mathbf{B} + \alpha_{19} \mathbf{I} \end{aligned} \tag{3.68}$$

A non-commutative polynomial in p variables of order n can be defined as [20]

$$\begin{aligned} N_n(\mathbf{A}, ...\mathbf{A_p}) \quad \triangleq \quad & \left\{ \sum_{i=1}^{\infty} \alpha_i \prod_{k=1}^{n} \prod_{q=1}^{p} \mathbf{A}_q^{l_{i,k,q}} : \right. \\ & \left. \sum_{i=1}^{\infty} l_{i,k,q} \in \{0, 1...n\} \wedge \alpha_i \in \mathbb{R} \forall i, q. \right\} \end{aligned} \tag{3.69}$$

Freeness between two matrices or non-commutative variables can be defined:

Definition 13. *[40]$\mathbf{A}$ and $\mathbf{B}$ are **free** (with respect to ϕ) if all polynomials $\mathbf{Q_i} \in N_\infty(\mathbf{A})$ and $\mathbf{R_i} \in N_\infty(\mathbf{B})$ such that,*

$$\phi(\mathbf{Q_i}) = \phi(\mathbf{R_j}) = 0 \tag{3.70}$$

Then we have,

$$\phi(\mathbf{Q_1 R_1} \cdots \mathbf{Q_n R_n}) = 0 \tag{3.71}$$

Freeness for the most general form of non-commutative variables can be defined:

Definition 14. *[20],[40] The sets* $\mathcal{Q}_1 \triangleq \{\mathbf{A_1} \cdots \mathbf{A_a}\}, \mathcal{Q}_2 \triangleq \{\mathbf{B_1} \cdots \mathbf{B_b}\} \cdots \mathcal{Q}_r$ *form free a free family* $(\mathcal{Q}_1 \cdots \mathcal{Q}_r)$ *if, for ever sequence* $(s_1 \cdots, s_k \cdots)$ *with* $s_k \in \{1, 2 \cdots r\} \forall k$ *and*

$$s_{k+1} \neq s_k \quad \forall k \tag{3.72}$$

and every sequence of polynomials $(\mathbf{Q_1} \cdots \mathbf{Q_k} \cdots)$ *with* $\mathbf{Q_k} \in N_\infty(\mathcal{Q}_{s_k})$ *and every positive integer* n

$$\phi(\mathbf{Q_1}) = \cdots = \phi(\mathbf{Q_n}) \Longrightarrow \phi(\mathcal{Q}_1 \cdots \mathcal{Q}_n) = 0 \tag{3.73}$$

"Due to the constraint on the sequence s_k adjacent factor in the product $\mathbf{Q_1} \cdots \mathbf{Q_n}$ must be polynomials of different sets of the family. This reflects the non-commutative nature in the definition of freeness[20]."

Example 8. *Using the definition of freeness, calculate the* $\phi(\mathbf{AB})$ *where* $\mathbf{A}$ *and* $\mathbf{B}$ *are free.*

Solution 8. *If we chose non-commutative polynomials*

$$\begin{aligned} \mathbf{Q_1} &= \mathbf{A} - \phi(\mathbf{A}) \\ \mathbf{Q_1} &= \mathbf{B} - \phi(\mathbf{B}) \end{aligned} \tag{3.74}$$

By using the definition

$$\phi(\mathbf{Q_1 Q_1}) = 0$$

$$\phi((\mathbf{A} - \phi(\mathbf{A}))(\mathbf{B} - \phi(\mathbf{B}))) = 0 \tag{3.75}$$

$$\Rightarrow \quad \phi(\mathbf{AB} - \phi(\mathbf{B})\mathbf{A} - \phi(\mathbf{A})\mathbf{B} + \phi(\mathbf{A})\phi(\mathbf{B})) = 0$$

$$\phi(\mathbf{AB}) = \phi(\mathbf{A})\mathbf{B} + \phi(\mathbf{B})\mathbf{A} - \phi(\mathbf{A})\phi(\mathbf{B}) \tag{3.76}$$

Free Random Matrices

Free probability allows to infer the asymptotic eigenvalue distribution of sums and or products of free random matrices. This powerful tool works if and if only if the matrices are free. In this subsection, sets of free family matrices is presented. Most of them were found by Voiculesco and strengthened and extended by Thorbjϕrnsen [38] and Hiai and Pets [31, 32].

Theorem 13. *[38, 31] Let the entries of independent* $R \times T$ *matrices* $\mathbf{H}_i$ *be i.i.d. complex Gaussian distributed with zero mean and variance* $1/R$*. Let the independent* $R \times R$ *matrices* $\mathbf{X}_i$ *have upper bounded norm and a limit distribution as* $R \to \infty$*. Moreover let the independent* $R \times R$ *matrices* $\mathbf{U}_i$ *be Haar-unitary. Assume* $(\mathbf{H_i}, \mathbf{X_i}, \mathbf{U_i})$ *are jointly independent for all* n*.*

(i) *Gaussian square random matrices(*$R = T$*): then the family*

$$\left(\{\mathbf{X_1}, \mathbf{X_1^\dagger}, \mathbf{X_2}, \mathbf{X_2^\dagger} \cdots\}, \{\mathbf{H_1}, \mathbf{H_1^\dagger}\}, \{\mathbf{H_2}, \mathbf{H_2^\dagger}\}, \cdots\right) \tag{3.77}$$

is almost surely asymptotically free as $R \to \infty$*.*

(ii) *Hermitian random matrices: Let* $\mathbf{S_i} \triangleq \mathbf{H_i}\mathbf{H_i}^\dagger$, *then the family*

$$\Big(\{\mathbf{X_1}, \mathbf{X_1}^\dagger, \mathbf{X_2}, \mathbf{X_2}^\dagger \cdots\}, \{\mathbf{S_1}\}, \{\mathbf{S_2}\}, \cdots\Big) \tag{3.78}$$

is almost surely asymptotically free as $R, T \to \infty$ *with ratio* $\beta = T/T$ *fixed.*

(iii) *Unitary random Matrices: the family*

$$\Big(\{\mathbf{X_1}, \mathbf{X_1}^\dagger, \mathbf{X_2}, \mathbf{X_2}^\dagger \cdots\}, \{\mathbf{U_1}, \mathbf{U_1}^\dagger\}, \{\mathbf{U_2}, \mathbf{U_2}^\dagger\}, \cdots\Big) \tag{3.79}$$

and,

$$\Big(\{\mathbf{X_i}\}, \{\mathbf{T_1}\mathbf{X_i}\mathbf{T_1}^\dagger\}, \{\mathbf{U_2}\mathbf{X_i}\mathbf{U_2}^\dagger\}, \cdots\Big) \tag{3.80}$$

are almost surely asymptotically free as $R \to \infty$.

Theorem 14. *[32] Let* $\mathbf{H}_n$ *be an independent family of* $R \times R$ *bi-unitarily invariant matrices for all* $1 \le n \le N$. *Let* $\mathbf{D}_m$ *be an independent family of* $R \times R$ *non-random diagonal matrices for all* $1 \le m \le M$. *Moreover, let* $\mathbf{H}_n\mathbf{H}_n^\dagger$, $\mathbf{D}_m\mathbf{D}_m^\dagger$ *have an upper bounded norm and a limit distribution as* $R \to \infty$. *Then the family,*

$$\Big(\{\mathbf{H_1}, \mathbf{H}_1^\dagger\}, \{\mathbf{H_2}, \mathbf{H}_2^\dagger\} \cdots \{\mathbf{D_1}, \mathbf{D}_1^\dagger\}, \{\mathbf{D_2}, \mathbf{D}_2^\dagger\} \cdots\cdots\Big) \tag{3.81}$$

us almost surely asymptotically free as $R \to \infty$.

3.2.2 Additive Free Convolution

Consider the free random matrices $\mathbf{A}$ and $\mathbf{B}$ and assume that, their asymptotic eigenvalue distributions are known. Now we want to address how to infer the asymptotic eigenvalue distribution of $\mathbf{A} + \mathbf{B}$.

Theorem 15. *[40]Let the hermitian matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are free. Then we have*

$$\kappa_{\mathbf{A}+\mathbf{B},n} = \kappa_{\mathbf{A},n} + \kappa_{\mathbf{B},n} \tag{3.82}$$

where $\kappa_{..}, n$ *is free cumulants as defined in (3.61).*

Definition 15. *Consider an hermitian random matrix* $\mathbf{X}$. *Then the definition of* R-transform *is*

$$R_{\mathbf{X}}(\omega) = \sum_{n=1}^{\infty} \kappa_{\mathbf{X},n}\omega^{n-1} \tag{3.83}$$

Let the matrices $\mathbf{A}$ and $\mathbf{B}$ are free. Then, with theorem 15 we have

$$\begin{aligned} R_{\mathbf{A}+\mathbf{B}}(\omega) &= \sum_{n=1}^{\infty} (\kappa_{\mathbf{A},n} + \kappa_{\mathbf{B},n})\omega^{n-1} && (3.84) \\ &= \sum_{n=0}^{\infty} \kappa_{\mathbf{A},n}\omega^{n-1} + \sum_{n=0}^{\infty} \kappa_{\mathbf{A},n}\omega^{n-1} && (3.85) \\ &= R_{\mathbf{A}}(\omega) + R_{\mathbf{B}}(\omega). && (3.86) \end{aligned}$$

Example 9. *Let the entries of the* $R \times T$ *matrix* $\mathbf{H}$ *be i.i.d with the variance* $1/R$ *with the ratio* $\beta = T/R$ *fixed. Show that,*

$$R_{\mathbf{HH}^\dagger}(\omega) = \frac{\beta}{1-\omega}. \tag{3.87}$$

Solution 9. *Let's start with (3.10),*

$$\phi\left((\mathbf{HH}^\dagger)^n\right) = \sum_{i=1}^{n} N_{n,r}\beta^r \tag{3.88}$$

With the definition of free cumulant we have

$$\sum_{i=1}^{n} N_{n,r}\beta^r = \sum_{\pi \in N(n)} \prod_{B_i \in \pi} \kappa_{|B_i|} \tag{3.89}$$

Remark that (3.89) holds if all free cumulants is equal β *as*

$$\sum_{i=1}^{n} N_{n,r}\beta^r = \sum_{\pi \in N(n)} \prod_{B_i \in \pi} \beta = \sum_{\pi \in N(n)} \beta^r. \tag{3.90}$$

Then the R-transform reads,

$$R(\omega) = \sum_{n=1}^{\infty} \beta\omega^{n-1} = \beta \sum_{n=0}^{\infty} \beta\omega^n \tag{3.91}$$

$$= \frac{\beta}{1-\omega}. \tag{3.92}$$

Theorem 16. *[41] The functional inversion of Stieltjes transform is equal to*

$$G^{-1}(\omega) = R(\omega) + \frac{1}{\omega}. \tag{3.93}$$

Lemma 5. *R-transform of the matrix* $c\mathbf{X}, c \in \mathbb{R}$ *can be expressed as*

$$R_{c\mathbf{X}}(\omega) = cR_{\mathbf{X}}(c\omega) \tag{3.94}$$

Proof. Let us star with (3.36),

$$c \cdot G_{\mathbf{cA}}(s) = G_{\mathbf{A}}\left(\frac{s}{c}\right) \tag{3.95}$$

With $s \to G_{\mathbf{cA}}^{-1}(s)$ we have

$$c = G_{\mathbf{A}}\left(\frac{G_{\mathbf{cA}}^{-1}(s)}{c}\right) \Rightarrow \frac{G_{\mathbf{cA}}^{-1}(s)}{c} = G_{\mathbf{A}}^{-1}(cs) \tag{3.96}$$

Thus,

$$R_{\mathbf{cA}}(\omega) = G_{\mathbf{cA}}^{-1}(\omega) - \frac{1}{\omega} = cG_{\mathbf{A}}^{-1}(c \cdot s)\frac{1}{\omega} - \frac{1}{\omega} \tag{3.97}$$

$$= c \cdot R_{\mathbf{A}}(c\omega) \tag{3.98}$$

□

Example 10. *Consider a projection matrix* $\mathbf{A}$*, and a matrix* $\mathbf{B} = \mathbf{U}\mathbf{A}\mathbf{U}^\dagger$ *where* $\mathbf{U}$ *is a Haar matrix, and*

$$p_A(x) = \frac{\delta(x+1) + \delta(x-1)}{2}.$$

Find the asymptotic eigenvalue distribution of $\mathbf{A} + \mathbf{B}$.

Solution 10. *First we have seen on the examples of free families that,* $\mathbf{A}$ *and* $\mathbf{B}$ *are free. It is immediate that,* $\mathbf{A}$ *and* $\mathbf{B}$ *have same distributions but the eigenvectors are fully uncorrelated. Thus,*

$$R_{\mathbf{A}+\mathbf{B}}(\omega) = 2 * R_{\mathbf{A}}(\omega) \tag{3.99}$$

So let's find Stieltjes transform of A,

$$\begin{aligned} G_{\mathbf{A}}(s) &= \int \frac{1}{s-x} dP_(x) && (3.100) \\ &= \frac{1}{2}\left(\frac{1}{s-1} + \frac{1}{s+1}\right) && (3.101) \end{aligned}$$

The functional inversion of $G_{\mathbf{A}}(s)$ *reads*

$$\begin{aligned} \omega &= G_{\mathbf{A}}\left(G_{\mathbf{A}}^{-1}(\omega)\right) && (3.102) \\ &= \frac{1}{2}\left(\frac{1}{G_{\mathbf{A}}^{-1}(\omega) - 1} + \frac{1}{G_{\mathbf{A}}^{-1}(\omega) + 1}\right) && (3.103) \\ 0 &= G_{\mathbf{A}}^{-1}(\omega)^2 - \frac{1}{\omega} G_{\mathbf{A}}^{-1}(\omega) - 1 && (3.104) \\ G_{\mathbf{A}}^{-1}(\omega) &= \frac{1 \mp \sqrt{1+4\omega^2}}{2\omega} && (3.105) \end{aligned}$$

With theorem 13 we have,

$$R_{\mathbf{A}}(\omega) = \frac{-1 \mp \sqrt{1+4\omega^2}}{2\omega} \tag{3.106}$$

We have two solutions. Now remark that

$$\begin{aligned} \lim_{\omega\to 0} R_{\mathbf{A}}(\omega) &= \lim_{\omega\to 0} \kappa_{\mathbf{A},1} + \sum_{n=2}^{\infty} \kappa_{\mathbf{A},\mathbf{n}} \omega^{n-1} && (3.107) \\ &= \kappa_{\mathbf{A},1} = \phi(\mathbf{A}) && (3.108) \end{aligned}$$

where the mean is 0*. Then we can determine the right solution as*

$$\begin{aligned} 0 &= \lim_{\omega\to 0} \frac{-1 \mp \sqrt{1+4\omega^2}}{2\omega} && (3.109) \\ 0 &= \lim_{\omega\to 0} \frac{-1}{2\omega} \mp \lim_{\omega\to 0} \frac{\sqrt{1+4\omega^2}}{2\omega} && (3.110) \\ 0 &= -\frac{1}{2} \mp \lim_{\omega\to 0} \frac{\sqrt{1+4\omega^2}}{2\omega} && (3.111) \end{aligned}$$

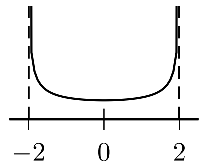


Figure 3.2: Arc Sine Law: The pdf of sum of two free random matrices both have binary eigenvalue distributed.

thus the one with positive sign is the right solution. Then

$$R_{\mathbf{A}}(\omega) = R_{\mathbf{A}}(\omega) = \frac{-1+\sqrt{1+4\omega^2}}{2\omega} \tag{3.112}$$

$$\begin{aligned} R_{\mathbf{A+B}}(\omega) &= \frac{-1+\sqrt{1+4\omega^2}}{\omega} \\ G^{-1}_{\mathbf{A+B}}(s) &= \frac{\sqrt{1+4s^2}}{s} \end{aligned} \tag{3.113}$$

Then the Stieltjes transform reads

$$\begin{aligned} s &= \frac{\sqrt{1+4G^2_{\mathbf{A+B}}(s)}}{G_{\mathbf{A+B}}(z)} \\ \Rightarrow G_{\mathbf{A+B}}(s) &= \frac{1}{\sqrt{s^2-4}} \end{aligned} \tag{3.114}$$

Finally by using inversion formula of the Stieltjes transform we have

$$\begin{aligned} p_{A+B}(x) &= -\frac{1}{\pi}\lim_{y\to 0}\Im G_{\mathbf{A+B}}(x+jy) \\ &= -\frac{1}{\pi}\lim_{y\to 0}\Im\frac{1}{\sqrt{(x+iy)^2-4}} \\ &= -\frac{1}{\pi}\Im\frac{1}{\sqrt{x^2-4}} \\ &= \frac{1}{\pi}\frac{1}{\sqrt{4-\lambda^2}} \end{aligned} \tag{3.115}$$

We have two remarkable observations. The first one is: if we randomly rotate the eigenvectors which is haar-unitary operation, then the matrix $\mathbf{A}$ is free with the randomly rotated one ($\mathbf{UAU}^{\dagger}$). Second, adding two free element which have discrete densities has continuous density as shown in figure 10. Therefore, having an intuition about free probability regards to (classical) probability theory is wrong. On the other hand, thinking about freeness regarding to interdependency of eigenvectors is a good intuition.

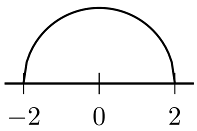


Figure 3.3: Free Central Limit Theorem: Semi Circle Law.

Corollary 2 (Free Central Limit Theorem [43])**.** *Let* $\mathbf{H}_n$ *be a free identical family of random matrices with the eigenvalues zero mean variance* 1 *for all* $1 \leq n \leq N$*. Then the asymptotic eigenvalue distribution of*

$$\mathbf{H} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{H}_n \tag{3.116}$$

converges in distribution to semicircle.

$$p_{\mathbf{H}}(x) = \frac{1}{2\pi}\sqrt{4-x^2} \qquad \lambda \in (-2, 2) \tag{3.117}$$

Proof.

$$R_{\mathbf{H}}(\omega) \quad = \quad \frac{1}{\sqrt{N}} \sum_{n=1}^{N} R_{\mathbf{H}_i}(\frac{\omega}{\sqrt{N}}) \tag{3.118}$$

Since the matrices identical-free, we have

$$\begin{aligned} R_{\mathbf{H}}(\omega) &= \frac{N}{\sqrt{N}} R_{\mathbf{H}_i}(\frac{\omega}{\sqrt{N}}) \\ &= \sqrt{N} R_{\mathbf{H}_i}(\frac{\omega}{\sqrt{N}}) \\ &= \sqrt{N}\left(\kappa_1 + \kappa_2 \frac{z}{\sqrt{N}} + \kappa_3 \frac{\omega^2}{N} + ...\right) \\ &= \sqrt{N}\left(0 + \frac{\omega}{\sqrt{N}} + \kappa_3 \frac{\omega^2}{N} + ...\right) \end{aligned} \tag{3.119}$$

where recall that: first order free cumulant is mean, the second order one is variance. As $N \to \infty$, the cumulants which higher than 2 vanishes, thus

$$\lim_{N\to\infty} \sqrt{N}\left(0 + \frac{\omega}{\sqrt{N}} + \kappa_3 \frac{\omega^2}{N} + ...\right) = \omega \tag{3.120}$$

One can simply follow same steps as previous example then will find the semi-circle distribution. □

Compression Of Random Matrix

Consider a $R \times R$ matrix $\mathbf{H}$ such that,

$$\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_R]. \tag{3.121}$$

Suppose, the $R \times T, (T \leq R)$ matrix $\mathbf{H}_\beta$ is defined as

$$\mathbf{H}_\beta = [\mathbf{h}_1, \cdots, \mathbf{h}_T]. \tag{3.122}$$

Assume that, you know R-transform of $R \times R$ matrix $\mathbf{H}\mathbf{H}^\dagger$. Then we want to address to determine of R-transform of $T \times T$ matrix $\mathbf{H}_\beta^\dagger \mathbf{H}_\beta$. The ideas compressing $R \times R$ matrix $\mathbf{H}\mathbf{H}^\dagger$ to $T \times T$ matrix by using projection matrix. As an example, let the $R \times R$ diagonal matrix $\mathbf{P}$ be a projection matrix such that,

$$p_{\mathbf{P}}(x) = (1 - \beta)\delta(x) + \beta\delta(x - 1). \tag{3.123}$$

For $R = 4, \beta = 1/2$,

$$\mathbf{H}\mathbf{H}^\dagger = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{pmatrix}; \quad \mathbf{P}\mathbf{H}\mathbf{H}^\dagger\mathbf{P} = \begin{pmatrix} h_{11} & h_{12} & 0 & 0 \\ h_{21} & h_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

which is called also corner of matrix. It is immediate see that, eigenvalue distribution of $T \times T$ corner of matrix $\mathbf{H}\mathbf{H}$ is equivalent to $\mathbf{H}_\beta^\dagger \mathbf{H}_\beta$. Then we can characterize any compressed version matrix which has real-valued eigenvalue distribution with uncompressed version in terms of cumulants as

Theorem 17 (Theorem 14.10 in [43]). *Consider the $R \times R$ hermitian random matrix $\mathbf{X}$. Let the $R \times R$ diagonal matrix $\mathbf{P}$ be distributed as*

$$p_{\mathbf{P}}(x) = (1 - \beta)\delta(x) + \beta\delta(x - 1). \tag{3.124}$$

Moreover define,

$$\mathbf{X}_\beta = \mathbf{X}\mathbf{P} \tag{3.125}$$

Then the asymptotic eigenvalue distribution of $\mathbf{X}_\beta$ converges almost surely to limit

$$p_{\mathbf{X}_\beta}(x) = (1 - \beta)\delta(x) + \beta p_{\mathbf{Y}}(x) \tag{3.126}$$

such that the R-transform of $\mathbf{Y}$ satisfies

$$R_{\mathbf{Y}}(\omega) = R_{\mathbf{X}}(\beta\omega) \tag{3.127}$$

as $R \to \infty$ with β fixed.

Example 11. *Let the entries of the $R \times T$ matrix $\mathbf{H}_\beta$ be i.i.d with the variance $1/R$ with the ratio $\beta = T/R \leq 1$ fixed. Then find the R-transform of $\mathbf{H}_\beta^\dagger \mathbf{H}_\beta$ for any $\beta \leq 1$.*

Solution 11. *With (3.87) for* $\beta = 1$*,*

$$R_{\mathbf{H}_1^\dagger \mathbf{H}_1}(\omega) = \frac{1}{1-\omega}. \tag{3.128}$$

Thus we have,

$$R_{\mathbf{H}_\beta^\dagger \mathbf{H}_\beta}(\omega) = R_{\mathbf{H}_1^\dagger \mathbf{H}_1}(\beta\omega) = \frac{1}{1-\beta\omega} \tag{3.129}$$

Theorem 18. *[24] Consider an invertible hermitian matrix* $\mathbf{X}$ *. Then we have,*

$$\frac{1}{R_{\mathbf{X}}(\omega)} = R_{\mathbf{X}^{-1}}\left(-R_{\mathbf{X}}(\omega)(1+\omega R_{\mathbf{X}}(\omega))\right). \tag{3.130}$$

Theorem 18 has an interesting connection with Replica Analysis for vector precoding for wireless MIMO wystems [24]. Moreover it can be useful to find performance measure of MIMO. In the following we address to find the minimum mean square error (MMSE) MIMO:

A-Quick Application I

Consider the wireless MIMO system described as

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{x} + \mathbf{n} \tag{3.131}$$

where $\mathbf{x}$, $\mathbf{y}$ $\mathbf{H}$, $\mathbf{n}$, γ are the channel input, the channel output,, the channel matrix, additive Gaussian noise (AWGN), and the signal -to- noise ratio, respectively. The entries of $\mathbf{x}$ and $\mathbf{n}$ are assumed to be iid zero mean and unit variance. Moreover let the entries of the $R \times T$ matrix $\mathbf{H}$ be iid with zero mean variance $1/R$.

In this subsection, we will address an important performance measure for (3.131): the minimum mean-square-error (MMSE) such that [60]

$$\frac{1}{T}\min_{\mathbf{M}\in\mathbb{C}^{R\times T}} E\left[\|\mathbf{x}-\mathbf{M}\mathbf{y}\|^2\right] = \int_0^\infty \frac{1}{1+\gamma x} dP_{\mathbf{H}^\dagger\mathbf{H}}(x) \tag{3.132}$$

Consider the MIMO system defined in (3.131). Let define a bi-variate function such that,

$$g(x,y) \triangleq 1 + x - xy - \sqrt{(x+1)^2 + xy(xy+2-2x)}. \tag{3.133}$$

Then we have,

$$\frac{1}{T}\min_{\mathbf{M}\in\mathbb{C}^{R\times T}} E\left[\|\mathbf{x}-\mathbf{M}\mathbf{y}\|^2\right] = \frac{2-g(\gamma,\beta)}{2+2\gamma-g(\gamma,\beta)} \tag{3.134}$$

as $T, R \to \infty$ with ratio $\beta = T/R$ fixed.

Solution

For simplicity let us define,

$$\mathbf{X} = \mathbf{I} + \gamma \mathbf{H}^{\dagger}\mathbf{H} \tag{3.135}$$

Thus,

$$\frac{1}{T} \min_{\mathbf{M} \in \mathbb{C}^{R \times T}} E\left(\|\mathbf{x} - \mathbf{M}\mathbf{y}\|^2\right) = \phi\left(\mathbf{X}^{-1}\right) = R_{\mathbf{X}^{-1}}(\omega)\big|_{\omega=0} \tag{3.136}$$

With the additive free convolution and the scaling property of R-transform, we have

$$R_{\mathbf{X}}(\omega) = 1 + \frac{1}{\gamma - \beta\omega} \tag{3.137}$$

Recall theorem 18,

$$\frac{1}{R_{\mathbf{X}}(\omega)} = R_{\mathbf{X}^{-1}}\left(-R_{\mathbf{X}}(\omega)(1 + \omega R_{\mathbf{X}}(\omega))\right). \tag{3.138}$$

To solve the problem we must find the ω such that,

$$-R_{\mathbf{X}}(\omega)(1 + \omega R_{\mathbf{X}}(\omega) = 0 \tag{3.139}$$

or explicitly

$$-\frac{1 + \gamma - \gamma\beta\omega}{1 - \gamma\beta\omega}\left[1 + \frac{\omega(1 + \gamma - \gamma\beta\omega)}{1 - \gamma\beta\omega}\right] \tag{3.140}$$

where outside term of bracket generates one solution, inside the bracket generates two solutions. The right solution can be simply found by following criteria

$$\phi(\mathbf{X}^{-1}) > 0 \tag{3.141}$$

since $\mathbf{X}$ is positive definite matrix which means $\mathbf{X}^{-1}$ is also positive definite. Then the right solution comes from inside the bracket with minus sign of square term of the solution quadratic equation such that,

$$\omega = \frac{1 + \gamma - \gamma\beta - \sqrt{(\gamma + 1)^2 + \beta\gamma(\beta\gamma + 2 - 2\gamma)}}{2\gamma\beta} \tag{3.142}$$

Finally,

$$\phi(\mathbf{X}^{-1}) = \frac{1}{R_{\mathbf{X}}(\omega)} = \frac{1 - \gamma\beta\omega}{1 + \gamma - \gamma\beta\omega}. \tag{3.143}$$

which completes the proof.

3.2.3 Multiplicative Free Convolution

Consider the free random matrices $\mathbf{A}$ and $\mathbf{B}$ and assume that, their asymptotic eigenvalue distributions are known. Now we want to address how to infer the asymptotic eigenvalue distribution of $\mathbf{AB}$.

Definition 16 (M-transform and S-transform)**.** *The moment generating function for a hermitian random matrix* $\mathbf{X}$ *is defined as*

$$M_{\mathbf{X}}(s) = \sum_{n=1}^{\infty} = \phi(\mathbf{X}^n)s^n \tag{3.144}$$

or equivalently,

$$M_{\mathbf{X}}(s) = \left(\frac{1}{s}\right) G_{\mathbf{X}}\left(\frac{1}{s}\right) - 1. \tag{3.145}$$

Moreover, the S-transform of $\mathbf{X}$ *is defined*

$$S_{\mathbf{X}}(z) = \frac{1+z}{z} M_{\mathbf{X}}^{-1}(s). \tag{3.146}$$

Theorem 19. *[42] Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be free random matrices such that, either* $\phi(\mathbf{A}) \neq 0$ *or* $\phi(\mathbf{B}) \neq 0$*. Then we have,*

$$S_{\mathbf{AB}}(z) = S_{\mathbf{A}}(z)S_{\mathbf{B}}(z). \tag{3.147}$$

Moreover, R-transform and S-transform has a straightforward relation such that [43],

$$S_{\mathbf{X}}(zR_{\mathbf{X}}(z)) = \frac{1}{R_{\mathbf{X}}(\omega)} \tag{3.148}$$

Example 12. *Let the entries of the* $R \times T$ *matrix* $\mathbf{H}$ *be i.i.d with the variance* $1/R$ *with the ratio* $\beta = T/R$ *fixed. Show that,*

$$S_{\mathbf{H}^\dagger\mathbf{H}}(\omega) = \frac{1}{1+\beta z}. \tag{3.149}$$

Solution 12. *Let us start with (3.129) in the previous example,*

$$R_{\mathbf{H}^\dagger\mathbf{H}}(\omega) = \frac{1}{1-\beta\omega} \tag{3.150}$$

Then with (3.148), we have

$$S_{\mathbf{X}}\left(\frac{\omega}{1-\beta\omega}\right) = 1 - \beta\omega \tag{3.151}$$

Let $z \to \omega(1-\beta\omega)^{-1}$ *which yields* $\omega = z(1+\beta z)^{-1}$*. Then we have*

$$S_{\mathbf{H}^\dagger\mathbf{H}}(z) = 1 - \frac{\beta\omega}{1+\beta z} = \frac{1}{z+\beta z} \tag{3.152}$$

Lemma 6. *Consider the* $R \times T$ *matrix* $\mathbf{X}$*. Then we have the following relation:*

$$S_{\mathbf{XX}^\dagger}(z) = \frac{z+1}{z+\beta} S_{\mathbf{X}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right) \tag{3.153}$$

with $\beta = T/R$*.*

Proof. Recall the lemma (3),

$$G_{\mathbf{XX}^\dagger}(s) = \beta G_{\mathbf{X}^\dagger\mathbf{X}}(s) + \frac{\beta - 1}{s} \tag{3.154}$$

Thus,

$$M_{\mathbf{XX}^\dagger}(s) = \beta M_{\mathbf{X}^\dagger\mathbf{X}}(s) \tag{3.155}$$

$$\Rightarrow M^{-1}_{\mathbf{XX}^\dagger}(z) = M^{-1}_{\mathbf{X}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right) \tag{3.156}$$

Then we have,

$$S_{\mathbf{XX}^\dagger}(z) = \frac{z+1}{z} M^{-1}_{\mathbf{X}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right) \tag{3.157}$$

$$= \frac{z+1}{z}\frac{z}{z+\beta} S_{\mathbf{X}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right) \tag{3.158}$$

$$= \frac{z+1}{z+\beta} S_{\mathbf{X}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right). \tag{3.159}$$

□

Example 13. *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be i.i.d with the variance $1/R$ with the ratio $\beta = T/R$ fixed. Then find the S-transform of $\mathbf{HH}^\dagger$.*

Solution 13. *Let us start with (3.149),*

$$S_{\mathbf{H}^\dagger\mathbf{H}}(\omega) = \frac{1}{1+\beta z} \tag{3.160}$$

With lemma 6 we have

$$S_{\mathbf{HH}^\dagger}(\omega) = \frac{z+1}{z+\beta} S_{\mathbf{H}^\dagger\mathbf{X}}\left(\frac{z}{\beta}\right) \tag{3.161}$$

$$= \frac{1}{z+\beta} \tag{3.162}$$

Lemma 7. *S-transform of the matrix $c\mathbf{X}, c \in \mathbb{R}$ can be expressed as*

$$S_{c\mathbf{X}}(z) = \frac{1}{c} S_{\mathbf{X}}(z). \tag{3.163}$$

Proof.

$$M_{c\mathbf{X}}(s) = \sum_{n=1}^{\infty} \phi(c^n \mathbf{X}^n) s^n = \sum_{n=1}^{\infty} \phi(\mathbf{X}^n)(cs)^n \tag{3.164}$$

$$= M_{\mathbf{X}}(sc) \tag{3.165}$$

Thus we have

$$M_{c\mathbf{X}}^{-1}(s) = \frac{1}{c}M_{\mathbf{X}}^{-1}(s) \tag{3.166}$$

$$S_{c\mathbf{X}}(z) = \frac{1}{c}S_{\mathbf{X}}(z). \tag{3.167}$$

□

Example 14. *Let the entries of the $R \times S$ and $S \times T$ matrices $\mathbf{A}$ and $\mathbf{B}$ be independent identically distributed with zero mean, variances $1/R$ and $1/S$ and*

$$\mathbf{H} \triangleq \mathbf{AB}. \tag{3.168}$$

Moreover assume that $R, S, T \to \infty$ with ratios $\rho = S/R$ and $\beta = R/T$ fixed. Then find the S-transform of $\mathbf{HH}^\dagger$.

Solution 14. *Let first define,*

$$\begin{aligned} \mathbf{C}_{R\times R} &= \mathbf{ABB}^\dagger\mathbf{A}^\dagger \\ \tilde{\mathbf{C}}_{S\times S} &= \mathbf{A}^\dagger\mathbf{ABB}^\dagger \end{aligned} \tag{3.169}$$

where,

$$S_{\mathbf{A}^\dagger\mathbf{A}}(z) = \frac{1}{1+\rho z} \qquad S_{\mathbf{BB}^\dagger}(z) = \frac{1}{z+\beta/\rho} \tag{3.170}$$

Then S-transform reads,

$$S_{\tilde{\mathbf{C}}}(z) = \frac{1}{(1+\rho z)(z+\beta/\rho)} \tag{3.171}$$

$$\begin{aligned} S_{\mathbf{C}}(z) &= \frac{z+1}{z+\rho}S_{\tilde{\mathbf{C}}}\left(\frac{z}{\rho}\right) \\ &= \frac{z+1}{z+\rho}\frac{\rho}{(1+z)(z+\rho\beta/\rho)} \\ &= \frac{\rho}{(z+\rho)(z+\beta)} \end{aligned} \tag{3.172}$$

Moreover S-transform of product of free matrices can be generalize in the following way:

Lemma 8. *Define a random matrix $\mathbf{X}$ as,*

$$\mathbf{X} = \mathbf{A}_N\mathbf{A}_{N-1}\cdots\mathbf{A}_2\mathbf{A}_1 \tag{3.173}$$

where $\mathbf{A}_n$ is the size of $K_n \times K_{n-1}$. Assume that, the family

$$(\{\mathbf{A}_1^\dagger\mathbf{A}_1\}, \{\mathbf{A}_2^\dagger\mathbf{A}_2\}\cdots\{\mathbf{A}_N^\dagger\mathbf{A}_N\}) \tag{3.174}$$

is asymptotically free all sizes K_n tend to infinity with the ratios

$$\chi_n = \frac{K_{n-1}}{K_n}, \quad 1 \le n \le N \tag{3.175}$$

remaining constant. Moreover define the ratios,

$$\rho_n = \frac{K_n}{K_N} \tag{3.176}$$

Then we have,

$$S_{\mathbf{XX}^\dagger}(z) = \frac{z+\rho_N}{z+\rho_0} \prod_{n=1}^{N} S_{\mathbf{A}_n^\dagger \mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{3.177}$$

In theorem 18 we saw the relation between a matrix with its inverse in term of R-transform. In S-transform this relation is quite similar determinant operator such as

$$\det \mathbf{X^{-1}} = \frac{1}{\det \mathbf{X}}.$$

Theorem 20. *[46] Let the hermitian matrix* $\mathbf{X}$ *be invertible. Then we have*

$$S_{\mathbf{X}^{-1}}(z) = \frac{1}{S_{\mathbf{X}^{-1}}(-1-z)}. \tag{3.178}$$

Indeed, the S-transform and the determinant operator has the following explicit relation:

Theorem 21. *[Corollary.5 in [48]]Let the* $R \times R$ *matrix* $\mathbf{X}$ *be bounded and invertible, then we have*

$$\frac{1}{R} \log\det \mathbf{XX}^\dagger = -\int_{-1}^{0} \log S_{\mathbf{XX}^\dagger}(z)dz. \tag{3.179}$$

as $R \to \infty$*.*

Theorem 21 is very interesting result, since it gives a compact formula for mutual information. In the following we show how to find the mutual information of MIMO channels where the channel matrix consist of iid entries. Note that, this is a trivial problem can be solved in term of Stieltjes transform, but to show the capability of Theorem 21 we used non-trivial way. Moreover we will re-refer the Theorem at the application chapter to derive explicit mutual information formula in high SNR regime.

A-Quick Application II

Consider the wireless MIMO system described in (3.131) as

$$\mathbf{y} = \sqrt{\gamma}\mathbf{Hx} + \mathbf{n} \tag{3.180}$$

In this subsection, we will address another important performance measure for (3.131): the mutual information such that,

$$\frac{I(\gamma)}{R} = \int \log(1+\gamma x)dP_{\mathbf{HH}^\dagger}(x) \tag{3.181}$$

Then with Theorem 21 we have,

$$\frac{I(\gamma)}{R} = \log\left(\frac{g(\gamma,\beta)-2}{2\gamma g(\gamma,\beta)^{\beta}}\right) + \frac{g(\gamma,\beta)}{2\gamma} + \beta^2 \log\beta\gamma \tag{3.182}$$

as $T, R \to \infty$ with ratio $\beta = T/R$ fixed.

Solution

In the same way, let us define

$$\mathbf{X} = \mathbf{I} + \gamma\mathbf{H}\mathbf{H}^{\dagger} \tag{3.183}$$

Then R-transform reads,

$$R_{\mathbf{X}}(\omega) = 1 + \frac{\beta}{\frac{1}{\gamma}+\omega} \tag{3.184}$$

Recall the functional relation between R-transform and S-transform (3.148) such that,

$$S_{\mathbf{X}}(zR_{\mathbf{X}}(z)) = \frac{1}{R_{\mathbf{X}}(\omega)} \tag{3.185}$$

Then we have two solution such that,

$$S_{\mathbf{X}}(z) = \frac{1+\beta\gamma+z\gamma \mp \sqrt{(z\gamma-1)^2+\gamma\beta(\gamma\beta+2z\gamma+2)}}{2z\gamma} \tag{3.186}$$

We can simply find the right solution by the following criteria

$$\lim_{\beta\to 0} S_{\mathbf{X}}(z) = 1 \tag{3.187}$$

since the S-transform of identity matrix is 1. Then the one with positive sign of square term of 3.186 fulfils the criteria. With theorem 21 we have,

$$\begin{aligned} \frac{1}{R}\log\det\left(\mathbf{I}+\gamma\mathbf{H}\mathbf{H}^{\dagger}\right) &= -\int_{-1}^{0} \log S_{\mathbf{X}^{\dagger}}(z)dz. \\ &= \log\left(\frac{g(\gamma,\beta)-2}{2\gamma g(\gamma,\beta)^{\beta}}\right) + \frac{g(\gamma,\beta)}{2\gamma} + \beta^2\log\beta\gamma \end{aligned} \tag{3.188}$$

where (3.188) computed by *Mapple 15.*

Chapter 4

Non-hermitian Free Probability

Starting Girko's circular law is probably the best beginning for this section. In 1994 Girko proved that, the eigenvalues of $N \times N$ matrix with independent entries of mean 0, and the variance $1/N$ fall uniformly on a circular disk of radius 1 as $N \to \infty$. Figure 4.1 illustrates this numerically. The theorem is correct whether the matrix is real or complex. When the matrix is real there is a larger attraction of eigenvalues on the real axis and a small repulsion just off axis. This disappears as $N \to \infty$ [50].
Non-hermitian matrices have complex-valued eigenvalue distribution in general. In the hermitian case, we worked on the complex-valued functions to search real-valued eigenvalues, we now have to work on a q-valued function to search complex-valued eigenvalues (see figure 4.2).
To deal with complex-valued eigenvalue distribution, an extended version of

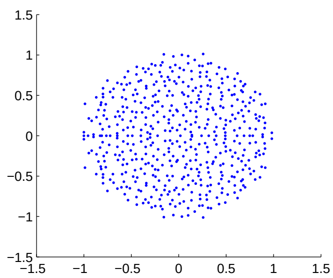


Figure 4.1: The eigenvalues of 500×500 Gaussian random matrices in the complex plane.

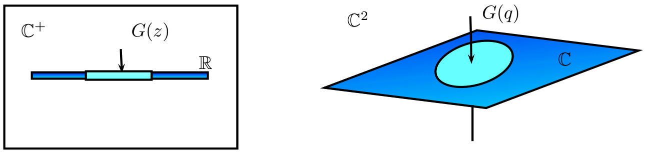


Figure 4.2: Left: Complex-valued operation for a real function in upper complex plane. Right: Quaternion-valued operation for a complex function in hyper complex plane.

free probability which we call “Quatartenionic Free Probability Theory[1]” is introduced.

4.1 Quatartenionic Free Probability Theory

Hermitian matrices have real eigenvalues. The method of choice to deal with real-valued eigenvalue distributions in free probability is to utilize complex analysis, i.e. to represent a real-valued eigenvalue distribution

$$p(x) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \Re\{jG(x + j\epsilon)\} \tag{4.1}$$

as a limit of a complex-valued holomorphic function $G(s)$, which is the Stieltjes transform and as defined before by

$$G(s) = \int \frac{dP(x)}{s - x} \tag{4.2}$$

Complex-valued eigenvalue distributions are often circularly symmetric and, thus, not holomorphic. They can be represented by a pair of holomorphic functions representing real and imaginary part. Instead of real and imaginary part of a complex variable z, one can also consider z, its complex conjugate z^*, and apply the Wirtinger rule [53] for differentiation, i.e.

$$\frac{\partial z}{\partial z^*} = 0 = \frac{\partial z^*}{\partial z}. \tag{4.3}$$

4.1.1 Stieltjes Transform

In order to generalize the Stieltjes transform to two complex variables z and z^*, we first rewrite (4.2) by

$$G(s) = \frac{d}{ds} \int \log(s - x) dP(x) \tag{4.4}$$

[1] A less explicit calculus for non-Hermitian random matrices was already proposed by Jarosz and Nowak in [52]. However we generally follow a explicit calculus govern by Müller in [23]

	Probability Space*	Algebra*
Classical Probability*	Commutative	Commutative
Free Probability*	Non-commutative	Commutative
Quatartenionic Free Probability*	Non-Commutative	Non-commutative

Figure 4.3: Comparison between Classical, Free, and Quatartenionic Free Probability Theories

Further, note that the Dirac function of complex argument can be represented as the limit

$$\begin{aligned} \delta(z-z') &= \frac{1}{\pi}\lim_{\epsilon\to 0}\frac{\epsilon^2}{(|z-z'|^2+\epsilon^2|)^2} \\ &= \frac{1}{\pi}\lim_{\epsilon\to 0}\frac{\partial^2}{\partial z\partial z^*}\log[|z-z'|^2+\epsilon^2] \end{aligned} \tag{4.5}$$

Thus we have

$$p(z) = \frac{1}{\pi}\lim_{\epsilon\to 0}\frac{\partial^2}{\partial z\partial z^*}\int \log[|z-z'|^2+\epsilon^2]dP(z') \tag{4.6}$$

We define the bivariate Stieltjes transform by

$$\begin{aligned} G(s,\epsilon) &= \frac{\partial}{\partial s}\int \log[|s-z|^2+\epsilon^2]dP(z) \\ &= \int \frac{(s-z)^*}{|s-z|^2+\epsilon^2}dP(z) \end{aligned} \tag{4.7}$$

and get the bivariate Stieltjes inversion formula to read

$$p(z) = \frac{1}{\pi}\lim_{\epsilon\to 0}\frac{\partial}{\partial z^*}G(z,\epsilon) \tag{4.8}$$

At first sight the bivariate Stieltjes transform looks quite different from 4.2. However we can write (4.7) as

$$G(s,\epsilon) = \int \left[\begin{pmatrix} s-z & i\epsilon \\ i\epsilon & s^*-z^* \end{pmatrix}^{-1}\right]_{11} dP(z) \tag{4.9}$$

which clearly resembles the form of (4.2). To get an even more striking analogy with (4.2), we can introduce the Stieltjes transform with quaternionic argument $q \equiv v + j\omega, (v,\omega) \in \mathbb{C}^2, i^2 \equiv -1, ij = -ji$

$$G(q) \equiv \int \frac{dP(z)}{q-z} \tag{4.10}$$

and with the respective inversion formula

$$p(z) = \frac{1}{\pi} \lim_{\epsilon \to 0} \frac{\partial}{\partial z^*} \Re G(z + i\epsilon) \tag{4.11}$$

and the definition $\Re(v+i\omega) \equiv v \in \mathbb{C}$.[2] Quaternions are inconvenient to deal with since multiplication of quaternions does not commute, in general. However, any quaternion $q = v + i\omega$ can be conveniently represented by the complex-valued 2×2 matrix

$$\begin{pmatrix} v & \omega \\ -\omega^* & v^* \end{pmatrix}. \tag{4.12}$$

This matrix representation directly connects (4.9) with (4.10) via

$$G(s, \epsilon) = \Re G(s + i\epsilon). \tag{4.13}$$

Finally, the quaternion-valued Stieltjes transform can be expressed as

$$\begin{aligned} G(q) &\equiv \int (1 - q^{-1}z)^{-1} q^{-1} dP(z) && (4.14) \\ &= \sum_{n=0}^{\infty} \int (q^{-1}z)^n q^{-1} dP(z) && (4.15) \\ &= \sum_{n=0}^{\infty} E\left[(zq^{-1})^n\right] q^{-1}. && (4.16) \end{aligned}$$

Note that (4.16) is equivalent to

$$q^{-1} \sum_{n=0}^{\infty} E\left[(q^{-1}z)^n\right]. \tag{4.17}$$

But we'll follow (4.16) for the rest of the work.

4.1.2 Additive Free Convolution

We define the R-transform of quaternion argument p in complete analogy to the complex case in [30] as

$$R(p) = G^{-1}(p) - p^{-1} \tag{4.18}$$

and obtain for free random matrices $\mathbf{A}$ and $\mathbf{B}$, with $R_{\mathbf{A}}(p)$ and $R_{\mathbf{A}}(p)$ denoting the R-transforms of the respective asymptotic eigenvalue distributions,

$$R_{\mathbf{A}+\mathbf{B}}(p) = R_{\mathbf{A}}(p) + R_{\mathbf{A}}(p) \tag{4.19}$$

The scaling law of the R-transform generalizes as follows

$$R_{\mathbf{zA}}(p) = zR_{\mathbf{A}}(pz) \tag{4.20}$$

for $z \in \mathcal{C}$. Note that the order of factors does matter here, since $pz \neq zp$, in general.

[2] Note that real and imaginary part of a quaternion are its first and second complex component, respectively.

Remark 4. *Let* **A** *and* **B** *are free each others. Then we have*

$$G_{\mathbf{A}+\mathbf{B}}(q) = G_{\mathbf{A}}\left(q - R_{\mathbf{B}}\left[G_{\mathbf{A}+\mathbf{B}}(q)\right]\right) \tag{4.21}$$

Proof.

$$\begin{aligned} q &= G_{\mathbf{A}}\left[G_{\mathbf{A}}^{-1}(q)\right] \\ &= G_{\mathbf{A}}\left[G_{\mathbf{A}+\mathbf{B}}^{-1}(q) - G_{\mathbf{B}}^{-1}(q) + \frac{1}{q}\right] \\ &= G_{\mathbf{A}}\left[G_{\mathbf{A}+\mathbf{B}}^{-1}(q) - R_{\mathbf{B}}(q)\right] \end{aligned} \tag{4.22}$$

By substitution $q \to G_{\mathbf{A}+\mathbf{B}}(q)$, we have

$$G_{\mathbf{A}+\mathbf{B}}(q) = G_{\mathbf{A}}\left(q - R_{\mathbf{B}}\left[G_{\mathbf{A}+\mathbf{B}}(q)\right]\right)$$

□

4.1.3 Multiplicative Free Convolution

While additive free convolution generalizes straightforwardly, this is very different for multiplicative free convolution.

Define a modified quaternion-valued Stieltjes transform of a non-Hermitian random matrix **X** as,

$$\mathcal{G}_{\mathbf{X}} = \lim_{\epsilon \to 0} G_{\mathbf{X}}(z + i\epsilon) \tag{4.23}$$

Moreover for any $q \in \mathbb{C}^2$ define the following operation as

$$q^l = \omega q \omega^*, \quad q^r = \omega^* q \omega \tag{4.24}$$

where $\omega \triangleq e^{(j \arg z)/4}$. Let the non-hermitian matrices **A** and **B** are free each others. Then we have [54],

$$R_{\mathbf{AB}}\left(\mathcal{G}_{\mathbf{AB}}\right) = R_{\mathbf{A}}\left(\mathcal{G}_{\mathbf{B}}\right)^l R_{\mathbf{B}}\left(\mathcal{G}_{\mathbf{A}}\right)^r. \tag{4.25}$$

But this is non-trivial formula and very less fruitful compare quaternion valued R transform. On the other hand there is an interesting result in the name of *S-Transfrom over (non-commutative) unital Banach algebra* [51]:

Let b be a unital-(non commutative) Banach algebra over complex number. Define the M-transform in term of the b-valued function,

$$M(b) \triangleq G(b^{-1}) \cdot b^{-1} - 1 \tag{4.26}$$

$$= \sum_{n=1}^{\infty} E\left[(bz)^n\right] \tag{4.27}$$

Then we have b-valued S-transform can be written as,

$$S(b) = b^{-1}(1 + r)M^{-1}(b) \tag{4.28}$$

Let x and y are free operators. Then we have [51]

$$S_{xy}(b) = S_y(b) S_x(S_y(b)^{-1} r S_y(b)). \tag{4.29}$$

Since the b defined as a non-commutative variable, the functional form $b-valued$ Stieltjes transform completely matches with $q-valued$ Stieltjes transform. Thus we have the following conjecture:

Proposition 1. *Let the square random matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *be free each others. Define quaternion-valued M-transform as*

$$\begin{aligned} M(q) &\triangleq G(q^{-1}) \cdot q^{-1} - 1 &(4.30)\\ &= \sum_{n=1}^{\infty} E\left[(qz)^n\right] &(4.31)\end{aligned}$$

Moreover define the quaternion valued S-transform

$$S(r) = r^{-1}(1+r) M^{-1}(r) \tag{4.32}$$

where $r \in \mathbb{C}^2$. *Then we propose*

$$S_{\mathbf{AB}}(r) = S_{\mathbf{B}}(r) S_{\mathbf{A}}(S_{\mathbf{B}}(r)^{-1} r S_{\mathbf{B}}(r)). \tag{4.33}$$

4.1.4 Quaternion-valued functions for hermitian matrices

Recall that, the quaternion-valued Stieltjes can be expanded as,

$$\int \frac{dP(z)}{q-z} = \sum_{n=0}^{\infty} E\left[(qz)^n\right] q^{-1}$$

However the quaternion-valued Stietjes transform for a real distribution can be written

$$\begin{aligned} \int \frac{dP(x)}{q-x} &= \sum_{n=0}^{\infty} \int \frac{x^n dP(x)}{q^{n+1}} \\ &= \sum_{n=0}^{\infty} \frac{m_n}{q^{n+1}} &(4.34)\end{aligned}$$

since $qx = xq, x \in \mathbb{R}$. This yields the same algebra as in the complex case. Therefore quaternion-valued Stieltjes, R and S transforms for a real distribution are simply equivalent to the complex case. Obviously

$$G_{\mathbf{H}}(q) = G_{\mathbf{H}}(s)|_{s=q}\,; \quad R_{\mathbf{H}}(p) = R_{\mathbf{H}}(\omega)|_{\omega=p}\,; \quad S_{\mathbf{H}}(r) = S_{\mathbf{H}}(z)|_{z=r}$$

where $\mathbf{H}$ is a hermitian matrix.

Example 15. *Let* $\mathbf{H}$ *is a semicircle element. Find* $G_{\mathbf{H}}(q)$ *and* $R_{\mathbf{H}}(p)$.

Solution 15.

$$\int \frac{dP_H(x)}{q-x} = \sum_{n=0}^{\infty} \frac{C_n}{q^{2n+1}} \tag{4.35}$$

since odd moments of a even distribution vanishes and C_n is nth Catalan number. By using recursive expression of Catalan number, we have [43],

$$\begin{aligned} G_{\mathbf{H}}(q) &= \frac{1}{q} + \sum_{n=1}^{\infty} \frac{1}{q^{2n+1}} \left(\sum_{m=1}^{n} C_{m-1} C_{n-m} \right) \\ &= q^{-1} + q^{-1} \sum_{n=1}^{\infty} \sum_{m=1}^{n} \frac{C_{m-1}}{q^{2m+1}} \cdot \frac{C_{n-m}}{q^{2(m-n)+1}} \\ &= q^{-1} + q^{-1} \sum_{m=1}^{\infty} \frac{C_{m-1}}{q^{2m+1}} \cdot \left(\sum_{n=m}^{\infty} \frac{C_{n-m}}{q^{2(m-n)+1}} \right) \\ &= q^{-1} + q^{-1} \sum_{m=1}^{\infty} \frac{C_{m-1}}{q^{2m+1}} \cdot G_H(q) \\ &= q^{-1} + q^{-1} G_{\mathbf{H}}^2(q) = q^{-1} \left(1 + G_{\mathbf{H}}^2(q)\right). \end{aligned} \tag{4.36}$$

which yields the following solution,

$$G_{\mathbf{H}}(q) = \frac{1}{2} \left[q - \left(q^2 - 4\right)^{\frac{1}{2}} \right] \tag{4.37}$$

Now with a substitution $q \to G_{\mathbf{H}}^{-1}(q)$ in (4.36) we have,

$$\begin{aligned} 0 &= \left[G_{\mathbf{H}}^{-1}(q)\right]^{-1} (1+q^2) - q \\ &= \left[R_{\mathbf{H}}(q) + q^{-1}\right]^{-1} (1+q^2) - q \\ &= (1+q^2) - (R_{\mathbf{H}}(q) + q^{-1}) q \end{aligned} \tag{4.38}$$

which yields

$$R_{\mathbf{H}}(q) = q \tag{4.39}$$

Remark 5. *Let* $\mathbf{G}$ *is a full circle element, then we have*

$$R_{\mathbf{G}}(q) = \Im q \tag{4.40}$$

Proof. $\mathbf{G}$ can be decomposed as

$$\mathbf{G} = \frac{\mathbf{H}_1 + j\mathbf{H}_2}{\sqrt{2}} \tag{4.41}$$

where $\mathbf{H}_{1,2}$ are semicircle element and free each others. Then we have

$$R_{\mathbf{G}}(q) = \frac{1}{2}(q + jqj) = \Im q \tag{4.42}$$

□

4.2 R-Diagonal Matrices

In the previous section, we worked on quaternionic free probability which allows to deal with complex-valued eigenvalue distributions of non-Hermitian matrices. In this section we will work on methods for R-diagonal matrices:

Definition 17. *[56] A random matrix* $\mathbf{X}$ *is called R-diagonal if it can be decomposed as* $\mathbf{X} = \mathbf{U}\mathbf{Y}$*, such that* $\mathbf{U}$ *is Haar unitary and free of* $\mathbf{Y} = \sqrt{\mathbf{X}\mathbf{X}^\dagger}$*.*

Recall in the previous chapter we defined $R \times R$ bi-unitarily invariant matrix $\mathbf{X}$ such that, it can be decomposed as $\mathbf{X} = \mathbf{U}\mathbf{Y}$ where $\mathbf{U}$ is $R \times R$ Haar unitary matrix and independent of $R \times R$ matrix $\mathbf{Y}$. As the matrix size grows, independence is converted into freeness according to some freeness result [32]. Therefore bi-unitarily invariant matrices are asymptotically R-diagonal[3].

R-diagonal matrices have circularly symmetric eigenvalue distribution. In order to determine the boundary of such distributions, we define the following measures [47]:

$$in(\mathbf{X})^2 \triangleq \int \frac{1}{x} dP_{\mathbf{X}\mathbf{X}^\dagger}(x) \tag{4.43}$$

$$out(\mathbf{X})^2 \triangleq \int x dP_{\mathbf{X}\mathbf{X}^\dagger}(x) \tag{4.44}$$

(where these integrals are computed by using the conventions $1/0 = \infty$ and $1/\infty = 0$). It is obvious that, $out(\mathbf{X})^2$ is the 2nd moment of singular value distribution of $\mathbf{X}$ and when $\mathbf{H}$ is invertible (or has no zero eigenvalues), $in(\mathbf{X})^2$ is the 2nd moment of singular value distribution of $\mathbf{X}^{-1}$.

4.2.1 Which matrices are R-diagonal?

It is immediate to see that a Haar-unitary matrix $\mathbf{T}$ and a (i.i.d) Gaussian random matrix $\mathbf{H}$ are asymptotically R-diagonal matrices such that they can decomposed as

$$\mathbf{T} = \mathbf{T}\mathbf{I}; \qquad \mathbf{H} = \mathbf{U}\mathbf{Q} \tag{4.45}$$

where $\mathbf{U}$ is a Haar-unitary matrix and $\mathbf{Q}$ is a quarter circle distributed random matrix. Moreover, with the following theorems we here present some important class R-diagonal matrices:

Theorem 22. *[47] Let the matrix* $\mathbf{X}_i$ *be free family of R-diagonal matrices for all* $1 \leq n \leq N$*. Then,*

- *Sum of free R-diagonal matrices:* $\sum_i \mathbf{X}_i$
- *Product of free R-diagonal matrices:* $\prod \mathbf{X}_i$
- *Power of a R-diagonal matrices:* $\mathbf{X}_i^p$

[3] Note that independent R-diagonal matrices are free of each others.

are R-diagonal too.

Theorem 23 (Proposition 6.1.1 in [43])**.** *Let the matrix* $\mathbf{X}$ *be R-diagonal and free of the matrix* $\mathbf{Y}$*. Then* $\mathbf{XY}$ *is R-diagonal too.*

Theorem 24. *[43] Let the free hermitian matrices* $\mathbf{X}$ *and* $\mathbf{Y}$ *have a symmetric (even) eigenvalue distribution on the real line. Then the matrix* $\mathbf{XY}$ *is R-diagonal.*

4.2.2 Additive Free Convolution

On the operations such a sum or product of R-diagonal random matrix can be performed without quaternionic free calculus.

Consider a hermitian matrix $\widetilde{\mathbf{X}}$ such that the empirical eigenvalue distribution of $\widetilde{\mathbf{X}}$ is

$$p_{\widetilde{\mathbf{X}}}(x) = \frac{p_{\sqrt{\mathbf{XX}^\dagger}}(x) + p_{\sqrt{\mathbf{XX}^\dagger}}(-x)}{2} \tag{4.46}$$

where $\widetilde{\mathbf{X}}$ is symmetrized singular value version of $\mathbf{X}$.

Theorem 25. *[Proposition 3.5 in [47]] Let the asymptotically free random matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *be R-diagonal. Define*

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{4.47}$$

Then we have,

$$R_{\widetilde{\mathbf{C}}}(\omega) = R_{\widetilde{\mathbf{A}}}(\omega) + R_{\widetilde{\mathbf{B}}}(\omega). \tag{4.48}$$

At first sight, calculus with respect to the symmetrized singular distribution looks non-trivial. However the following two lemmas make the problem as trivial as in hermitian case.

Lemma 9. *[Symmetrization Lemma I] Let* $\mathbf{X}$ *be a rectangular non-Hermitian random matrix in general. Then we have,*

$$G_{\widetilde{\mathbf{X}}}(s) = sG_{\mathbf{XX}^\dagger}\left(s^2\right) \tag{4.49}$$

Lemma 10. *[Symmetrization Lemma II] Let the matrix* $\mathbf{X}$ *be defined as in the previous lemma. Then we have,*

$$S_{\widetilde{\mathbf{X}}}(z) = \left[\frac{z+1}{z} S_{\mathbf{XX}^\dagger}(z)\right]^{\frac{1}{2}} \tag{4.50}$$

Recall (3.148), such that $zS(z)$ and $zR(z)$ are functionally inverse each others [43]. Thus, (4.50) allows us to switch from S-transform from $\mathbf{XX}^\dagger$ to R-transform of $\widetilde{\mathbf{X}}$. and vice-versa. We also refer section 2.3 in [56] for an equivalent argument.

Example 16. *Let* $\mathbf{X}$ *is a deformed quarter circle element. Then find the* $\tilde{R}_{\mathbf{X}}(\omega)$.

Solution 16. *Recall (3.170) such that the S-transform of* $\mathbf{XX}^\dagger$ *reads*

$$S_{\mathbf{XX}^\dagger}(z) = (z+\beta)^{-1}. \tag{4.51}$$

Then with the inversion formula between R-transform and S-transform 3.148

$$\omega\tilde{R}_{\mathbf{X}}(\omega)\left(\omega\tilde{R}_{\mathbf{X}}(\omega)+1\right)\cdot\left(\omega\tilde{R}_{\mathbf{X}}(\omega)+\beta\right)^{-1} = \omega^2 \tag{4.52}$$

which yields the following solution[4]*,*

$$\tilde{R}_{\mathbf{X}}(\omega) = \sqrt{\left(\frac{\omega}{2}-\frac{1}{2\omega}\right)^2+\beta}+\frac{\omega}{2}-\frac{1}{2\omega} \tag{4.53}$$

4.2.3 Multiplicative Free Convolution

Trace operator is cyclic-invariant, this allows us to work in complex valued free probability by means of S-transform to deal with a multiplication of non-hermitian matrices.

Beside, the following theorem gives a straightforward way to switch from singular values to eigenvalues of R-diagonal matrices and vice versa:

Theorem 26. *[47] Let the random matrix* $\mathbf{X}$ *be R-diagonal such that it can decomposed as* $\mathbf{X} = \mathbf{UY}$ *where* $\mathbf{U}$ *is haar unitary matrix and free of* $\mathbf{Y} = \sqrt{\mathbf{XX}^\dagger}$*. Then we have*

(i) The eigenvalue distribution $P_{\mathbf{X}}(z)$*is circularly invariant with its boundary*

$$supp\,(P_{\mathbf{X}}) = [in(\mathbf{X})^{-1}, out(\mathbf{X})] \times_p [0, 2\pi).^5 \tag{4.54}$$

explicitly $P_{\mathbf{X}}$*'supports is the annulus with inner radius* $in(\mathbf{X})^{-1}$ *and outer radius* $out(\mathbf{X})$*.*

(ii) The S-transform $S_{\mathbf{Y}^2}$ *of* $\mathbf{Y}^2$ *has an analytic continuation to neighbourhood of interval* $(P_{\mathbf{Y}^2}(0)-1, 0]$ *and monotonically deceasing on* $(P_{\mathbf{Y}^2}(0)-1, 0]$ *such that the derivative of S-transform* $S'_{\mathbf{Y^2}} < 0$*, and it takes the values in between*

$$S\left((P_{\mathbf{Y^2}}(0)-1, 0]\right) = (in(\mathbf{X})^{-2}, out(\mathbf{X})^2]. \tag{4.55}$$

(iii) $P_{\mathbf{X}}(0) = P_{\mathbf{Y}}(0)$ *and the radial distribution function*

$$P_{\mathbf{X}}\left(S_{\mathbf{Y}^2}(r-1)^{-\frac{1}{2}}\right) = r; \quad r \in (P_{\mathbf{Y}}(0), 1]. \tag{4.56}$$

(iv) $P_{\mathbf{X}}(z)$ *is the only circularly symmetric probability measure satisfying (iii).*

[4] There are two solutions. Remark, in case $\beta = 1$, the right one must give $\tilde{R}(\omega) = \omega$.

[5] By $\times_p$ we denote polar set product: $A \times_p B = \{ae^{j\theta} | a \in A, \theta \in B\}$

Corollary 3. *[47] With the notation as in Theorem 26, the functional inversion of radial probability measure of* $\mathbf{X}$

$$P_{\mathbf{X}}^{-1}(r) = S_{\mathbf{Y^2}}(r-1)^{-\frac{1}{2}} : (P_{\mathbf{Y}}(0), 1] \to (in(\mathbf{X})^{-1}, out(\mathbf{X})] \tag{4.57}$$

has an analytical continuation to a neighbourhood of its domain and monotonically increasing on $(P_{\mathbf{Y}}(0), 1]$ *such that the derivative* $(P_{\mathbf{X}}^{-1})' > 0$*. Moreover the radial density of* $\mathbf{X}$ *such that*

$$2\pi r\, p_{\mathbf{X}}(z)|_{|z|=r} = \frac{dP_{\mathbf{X}}(r)}{dr}, \quad r \in (in(\mathbf{X})^{-1}, out(\mathbf{X})] \tag{4.58}$$

has an analytical continuation to neighbourhood of $(in(\mathbf{X})^{-1}, out(\mathbf{X})]$*.*

Theorem 26 and its Corollary play central to characterize non-hermitian random matrices. To comprehend the manner let us do an example:

Example 17. *Let the entries of the* $T \times T$ *matrix* $\mathbf{G}$ *be independent identically distributed with variance* $1/T$ *and the matrix* $\mathbf{P} \in \{0,1\}^{T\times T}$ *be diagonal with* L *non-zero entries. Then, show that the empirical eigenvalue distribution of* $\mathbf{H} = \mathbf{GP}$ *converges almost surely to*

$$p(z) = (1-\phi)\delta(z) + \begin{cases} \frac{1}{\pi} & |z| < \sqrt{\phi} \\ 0 & elsewhere \end{cases} \tag{4.59}$$

Solution 17. *Remark that* $\mathbf{HH}^\dagger$ *is the square equivalent of deformed quarter circle law (eigenvalues) element. Thus*

$$S_{\mathbf{HH}^\dagger}(z) = \frac{1}{z+\phi} \tag{4.60}$$

With theorem 26, we have,

$$P_{\mathbf{H}}^{-1}(r) = \frac{1}{\sqrt{S_{\mathbf{HH}^\dagger}(r-1)}} = \sqrt{r+\phi-1} \tag{4.61}$$

Then the probability measure (radial) reads

$$P_{\mathbf{H}}(r) = (1-\phi) + r^2 \tag{4.62}$$

Moreover, the zero measure of the distribution is

$$P_{\mathbf{H}}(0) = (1-\phi)\delta(z) \tag{4.63}$$

Where the asymptotic eigenvalue distribution of $\mathbf{H}$ *can easily find as*

$$\begin{aligned} p_{\mathbf{H}}(z) &= (1-\phi)\delta(z) + \left(\frac{1}{2\pi r}\frac{dP_{\mathbf{H}}(r)}{dr}\right)\Big|_{|z|=r} \\ &= (1-\phi)\delta(z) + \frac{1}{\pi} \end{aligned} \tag{4.64}$$

Finally we need determine the boundary of the density. Since the distribution has some zero measure, then inner radius of the density reads,

$$in(\mathbf{H})^{-1} = 0 \tag{4.65}$$

The outer radius of the density reads,

$$out(\mathbf{H}) = \frac{1}{\sqrt{S_{\mathbf{H}\mathbf{H}^\dagger}(0)}} = \sqrt{\phi}. \tag{4.66}$$

A large class of R-diagonal matrices have the property to behave as if they are identical with respect to multiplication:

Theorem 27. *[Proposition 3.10 in [47]] Let the random matrices $\mathbf{H}_n$ be asymptotically free R-diagonal elements, and their asymptotic eigenvalue distributions of $\mathbf{H}_n$ be identical for all n. Then the asymptotic eigenvalue distributions of*

$$\prod_{n=1}^{N} \mathbf{H}_n \tag{4.67}$$

and $\mathbf{H}_1^N$ are identical.

Theorem 27 allows us to derive the asymptotic eigenvalue distribution many types of practically relevant random matrices with stardard methods for the transformation of probability densities. We will re-refer this weird consequence of R-diagonal matrices in the application chapter.

Moreover, the R-diagonal matrices has an interesting consequence of additive free additive convolution regarding to singular values:

Theorem 28. *[Theorem 3.4 in [57]]Let $\mathbf{X}$ be R-diagonal matrix, and decomposed as $\mathbf{X} = \mathbf{U}\sqrt{\mathbf{Y_1}}$ such that $\mathbf{U}$ is Haar-unitary matrix and free with $\sqrt{\mathbf{Y_1}} = \sqrt{\mathbf{X}\mathbf{X}^\dagger}$. Furthermore let the asymptotic eigenvalue distribution of $\mathbf{X}$ be*

$$\phi\delta(z) + \begin{cases} p_{\mathbf{H}}(z) & P_{\mathbf{H}}^{-1}(\phi) < |z| \le b \\ 0 & elsewhere \end{cases} \tag{4.68}$$

Moreover, define a summation of identical free matrices as

$$\mathbf{Y}_\beta = \sum_{n=1}^{\frac{1}{\beta}} \mathbf{Y}_n \tag{4.69}$$

Then the asymptotic eigenvalue distribution of $\mathbf{X}_\beta = \mathbf{U}\sqrt{\mathbf{Y}}_\beta$ satisfies,

$$p_{\mathbf{H}_\beta}(z) = \phi_\beta\delta(z) + \begin{cases} p_{\mathbf{H}}\left(\frac{z}{\sqrt{\beta}}\right) & \sqrt{\beta}P_{\mathbf{H}}^{-1}(\beta^{-1}\phi_\beta + 1 - \beta^{-1}) < |z| \le \sqrt{\beta}P_{\mathbf{H}}^{-1}(1) \\ 0 & elsewhere \end{cases} \tag{4.70}$$

where $\phi_N = \max(0, 1 + \phi N - N)$.

Theorem 28 immediately inspires us to propose the following theorem:

Theorem 29. *Consider a $R \times R$ R-diagonal matrix $\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_R]$ whose eigenvalue distribution*

$$\phi\delta(z) + \begin{cases} p_{\mathbf{H}}(z) & P_{\mathbf{H}}^{-1}(\phi) < |z| \le b \\ 0 & elsewhere \end{cases} \tag{4.71}$$

where $P_{\mathbf{H}}^{-1}(r)$ is the functional inversion of radial probability measure (CDF). Moreover, let the $R \times T; (T \le R)$ matrix $\mathbf{H}_\beta = [\mathbf{h}_1, \cdots, \mathbf{h}_T]$. Define

$$\phi_\beta = \max\left(0, 1 + \phi\beta^{-1} - \beta^{-1}\right) \tag{4.72}$$

Then empirical eigenvalue distribution of $\mathbf{H}_{\beta_u}$ such that

$$\mathbf{H}_{\beta,u} = \mathbf{U}\sqrt{\mathbf{H}_\beta^\dagger \mathbf{H}_\beta} \tag{4.73}$$

converges almost surely to limit distribution satisfies

$$p_{\mathbf{H}_{\beta,u}}(z) = \phi_\beta\delta(z) + \begin{cases} \frac{1}{\beta}p_{\mathbf{H}}(z) & P_{\mathbf{H}}^{-1}(\beta\phi_\beta + 1 - \beta) < |z| \le b \\ 0 & elsewhere \end{cases} \tag{4.74}$$

as $T, R \to \infty$ with $\beta = T/R \le 1$ fixed.

Theorem 29 is has very relevant role to analyse for rectangular random matrices. With Theorem 29 we can very easily find the asymptotic eigenvalue distribution of *singular* equivalent of rectangular random matrices.

In the following chapter will be application chapter for MIMO System and we will characterize practically relevant random matrices with the following definitions:

Definition 18. *Consider the $R \times T, (T \le R)$ matrix $\mathbf{H}$ and the $R \times (R - T)$ null matrix $\mathbf{N}$. Let the $R \times R$ matrix $\mathbf{H}_s = [\mathbf{H}|\mathbf{N}]$. Then we have*

$$\mathbf{H_s}\mathbf{H_s}^\dagger = \mathbf{H}\mathbf{H}^\dagger. \tag{4.75}$$

We call $\mathbf{H}_s$ as the square equivalence *of $\mathbf{H}$.*

Moreover in case the matrix is R-diagonal that also means the matrix biunitarily invariant, then square equivalence can be replaced by *square equivalent*:

Definition 19. *Let the $R \times R$ random matrix be $\mathbf{H}$ be R-diagonal such that,*

$$\mathbf{H}_\beta = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_R] \tag{4.76}$$

Moreover define a $R \times T$ random matrix as

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_T] \tag{4.77}$$

with the ratio $\beta = T/R \leq 1$ *fixed. Define an arbitrary* $R \times R$ *diagonal matrix matrix* $\mathbf{P}$ *such that, the diagonal entries*

$$p_{\mathbf{P}}(x) = (1-\beta)\delta(x) + \beta\delta(x-1) \tag{4.78}$$

Moreover define $\mathbf{HP} = \mathbf{H}_p$*, then we have,*

$$\mathbf{H}_p\mathbf{H}_p^\dagger \equiv \mathbf{H}_\beta\mathbf{H}_\beta^\dagger \tag{4.79}$$

Thus we called $\mathbf{H}_p$ *as a square equivalent of a rectangular random matrix* $\mathbf{H}_\beta$*.*

In case the matrix singular equivalent

Definition 20. *Consider the* $R \times T$ *matrix* $\mathbf{H}$*. Let the* $T \times T$ *matrix* $\mathbf{U}$ *be Haar-unitary matrix and free of* $\mathbf{H}^\dagger\mathbf{H}$*. Moreover define,*

$$\mathbf{H}_u = \mathbf{U}\sqrt{\mathbf{H}^\dagger\mathbf{H}} \tag{4.80}$$

Then, as far as concerning singular value distribution, we have

$$\mathbf{H}_u^\dagger\mathbf{H}_u \equiv \mathbf{H}^\dagger\mathbf{H}. \tag{4.81}$$

We call $\mathbf{H}_u$ *as the* singular equivalent *of* $\mathbf{H}$*.*

Chapter 5

Application to MIMO Systems

In this chapter, we will work on the application of free probability for MIMO System. The chapter divided into two section, the first section is under the Rayleigh fading assumption such that the channel does not have not line-of-sight component. Second section is under the Richian fading assumption where the channel has line-of-sight component.

We will refer practically relevant channel models and the main consideration will be asymptotic analysis of mutual information and asymptotic eigenvalue distribution of *square equivalence* and *singular equivalent* of rectangular channel matrix.

5.1 Rayleigh Fading Channels

In this section, we will address the wireless MIMO system described as follows

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{x} + \mathbf{n} \tag{5.1}$$

where $\mathbf{H} \in \mathbb{C}^{R\times T}$, $\mathbf{x} \in \mathbb{C}^{T}$, $\mathbf{y} \in \mathbb{C}^{R}$, $\mathbf{n} \in \mathbb{C}^{R}$, γ are the channel, the channel input, the channel output, additive white Gaussian noise, and the signal-to-noise ratio, respectively. Moreover entries of the channel $\mathbf{H} \in \mathbb{C}^{R\times T}$ represent the fading coefficients between each transmission path from a transmit antenna to receive antenna.

The channel input $\mathbf{x}$ and the noise vector $\mathbf{n}$ (AWGN) are assumed i.i.d (independent identically distributed) such that,

$$E[\mathbf{x}\mathbf{x}^{\dagger}] = E[\mathbf{n}\mathbf{n}^{\dagger}] = \mathbf{I} \tag{5.2}$$

Moreover, we assumed the channel has no line-of-sight component. Then the mutual information between the channel input and its output reads,

$$\frac{I(\gamma)}{R} = \int \log(1+\gamma x)dP_{\mathbf{H}\mathbf{H}^{\dagger}}(x) \tag{5.3}$$

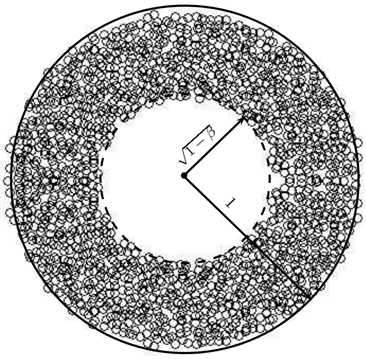


Figure 5.1: The eigenvalues of singular equivalent of 1000×750 iid Gaussian random matrix in the complex plane.

5.1.1 Rich Scattering Channel

In this model we assume the channel $\mathbf{H} \in \mathbb{C}^{R \times T}$ consist of a single scattering matrix such that the entries of $\mathbf{H}$ are assumed be iid with zero mean variance $1/R$.

Let us first try to understand the behaviour of $\mathbf{H}$ by means of both its *square equivalent* and its *singular equivalent*:

Corollary 4. *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be iid with zero mean variance $1/R$. Then, the empirical eigenvalue distribution of the* square equivalent *of $\mathbf{H}$ converges almost surely to*

$$p_{\mathbf{H}_p}(z) = (1-\beta) + \left\{ \begin{array}{cc} \frac{1}{\pi} & |z| \leq \sqrt{\beta} \\ 0 & elsewhere \end{array} \right. \tag{5.4}$$

as $R, T \to \infty$ with ratio $\beta = T/R \leq 1$ fixed.

In other words, the projection of iid square random matrices from R to T dimensions replaces the $R - T$ eigenvalues with greatest modulus by zero eigenvalues.

Corollary 5. *Let the entries of the $R \times T$ matrix $\mathbf{H}$ be iid with zero mean variance $1/R$. Then, the empirical eigenvalue distribution of the* singular equivalent *of $\mathbf{H}$ converges almost surely to*

$$p_{\mathbf{H}_u}(z) = \left\{ \begin{array}{cc} \frac{1}{\beta\pi} & \sqrt{1-\beta} < |z| \leq 1 \\ 0 & elsewhere \end{array} \right. \tag{5.5}$$

as $R, T \to \infty$ with ratio $\beta = T/R \leq 1$ fixed.

In other word, decreasing the rectangular ratio $\beta = T/R$ kills all eigenvalue on the unit disk with the radius $\sqrt{1-\beta}$ as shown in figure 5.

Before having look at the, mutual information formula respective channel model, we will first refer another important performance measure for (5.1) is the minimum mean square-error (MMSE) achieved by a linear receiver, determines signal-to-interference-and-noise ratio (SINR) such that [60],

$$\frac{1}{T}\min_{\mathbf{M}\in\mathbb{C}^{R\times T}} E\left[\|\mathbf{x}-\mathbf{M}\mathbf{y}\|^2\right] = \int_0^\infty \frac{1}{1+\gamma x} dP_{\mathbf{H}^\dagger\mathbf{H}}(x) \tag{5.6}$$

We have already found the MMSE measure for (5.1) in 3.2.2 such that,

$$\frac{1}{T}\min_{\mathbf{M}\in\mathbb{C}^{R\times T}} E\left[\|\mathbf{x}-\mathbf{M}\mathbf{y}\|^2\right] = \frac{2-g(\gamma,\beta)}{2+2\gamma-g(\gamma,\beta)} \tag{5.7}$$

as $T, R \to \infty$ with ratio $\beta = T/R$ fixed. Where the bi-variate function $g(x,y)$ is defined in

$$g(x,y) \triangleq 1 + x - xy - \sqrt{(x+1)^2 + xy(xy+2-2x)}. \tag{5.8}$$

Finally let us refer the mutual information for (5.1) which is already calculated in 3.2.3 as:

$$\frac{I(\gamma)}{R} = \log\left(\frac{g(\gamma,\beta)-2}{2\gamma g(\gamma,\beta)^\beta}\right) + \frac{g(\gamma,\beta)}{2\gamma} + \beta^2 \log\beta\gamma \tag{5.9}$$

as $T, R \to \infty$ with ratio $\beta = T/R$ fixed. For similar result, we also refer the reader to [61].

5.1.2 Channel with certain Scattering Richness

In this subsection we address more the following channel model for (5.1)[17]

$$\mathbf{H} = \mathbf{H}_2\mathbf{H}_1 \tag{5.10}$$

where $\mathbf{H}_1 \in \mathbb{C}^{S\times T}$, $\mathbf{H}_2 \in \mathbb{C}^{R\times S}$ the propagation from the transmit antennas to the scatterers, the propagation from the scatterers to the receive antennas, respectively. The entries of the $R \times S$ matrix $\mathbf{H}_2$ and the $S \times T$ matrix $\mathbf{H}_1$ are assumed iid with zero mean variances $1/R$ and $1/T$ respectively.

The channel model defined in (5.10) generalizes the model defined in (5.1) as it parametrizes various amounts of scattering richness such that,

$$\lim_{\rho\to\infty} \mathbf{H} \equiv \mathbf{X} \tag{5.11}$$

where $\rho = S/R$ and the entries of $\mathbf{X} \in \mathbb{C}^{\mathbb{R}\times\mathbb{T}}$ be iid with zero mean variance $1/R$.

As in the previous subsection, let us first analyse the non-hermitian matrix $\mathbf{H}$ defined in 5.10 by its *square equivalent* and *singular equivalent* as:

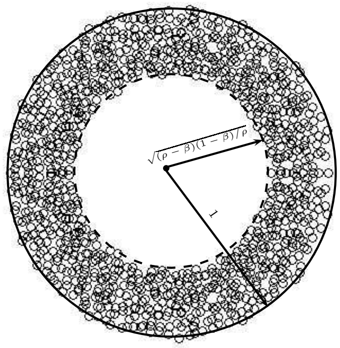


Figure 5.2: The eigenvalues of singular equivalent of 1000×750 $\mathbf{H}$ defined in (5.10) on the complex plane.

Corollary 6. *Let the random matrix* $\mathbf{H}$ *be defined as in (5.10). Then, the empirical eigenvalue distribution of* square equivalent *of* $\mathbf{H}$ *converges almost surely to the limit*

$$p_{\mathbf{H}_p}(z) = \max(1-\beta, 1-\rho)\delta(z) + \begin{cases} \frac{1}{\pi}\frac{\rho}{\sqrt{(\beta-\rho)^2+4\rho|z|^2}} & |z| \leq \sqrt{\beta} \\ 0 & elsewhere \end{cases} \quad (5.12)$$

as $T, R, S \to \infty$ *with ratios* $\rho = S/R$ *and* $\beta = T/R \leq 1$ *fixed.*

On the other the *singular equivalent* of $\mathbf{H}$ is:

Corollary 7. *Let the random matrix* $\mathbf{H}$ *be defined as in (5.10). Then, the empirical eigenvalue distribution of* singular equivalent *of* $\mathbf{H}$ *converges almost surely to the limit,*

$$p_{\mathbf{H}_u}(z) = \begin{cases} \frac{1}{\beta\pi}\frac{\rho}{\sqrt{(1-\rho)^2+4\rho|z|^2}} & \sqrt{(\rho-\beta)(1-\beta)/\rho} < |z| \leq 1 \\ 0 & elsewhere \end{cases} \quad (5.13)$$

as $T, R, S \to \infty$ *with ratios* $\rho = S/R \leq 1$[1] *and* $\beta = T/R \geq 1$ *fixed.*

In other word, decreasing the rectangular ratio $\beta = T/R$ kills all eigenvalues on the unit disk with the radius $\sqrt{(\rho-\beta)(1-\beta)/\rho}$ as shown in figure 7. Note that as general richness parameter ρ increase then the inner radius shows converges to zero.

Finally the mutual information of (5.10) can be written in term of differential form (54 in [17]) as

$$\left(1+\gamma\frac{dI(\gamma)}{d\gamma}\right)^3 + c_2\left(1+\gamma\frac{dI(\gamma)}{d\gamma}\right)^2 + c_1\left(1+\gamma\frac{dI(\gamma)}{d\gamma}\right) + \rho\beta = 0$$

[1] At the proof given in the appendix, we present the result without the constraint $\rho = S/R \leq 1$.

where $c_2 = (\rho + \beta + 1)$ and $c_1 = \rho\beta + \rho + \beta + \rho/\gamma$. Numerical solutions to this differential equation are simply obtained by solving (5.14) via Cardanos formula and integrating numerically. We refer reader to [17] for more general perspective and numerical results.

5.1.3 Completely Correlated MIMO Channel

Consider, the channel matrix $\mathbf{H} \in \mathbb{C}^{R \times T}$ whose entries are Gaussian variables but the entries are correlated . To specify the model, correlation coefficients of all pairs of elements are required [62].

On the other side, we can define a macroscopic variable α such that it is a kind of average measure over all correlation levels between entries. However we must connect the parameter α with the channel matrix. This can be achieved by defining

$$\mathbf{H}_\alpha = \prod_{n=1}^{\alpha} \mathbf{H}_n \tag{5.14}$$

where the entries of $R \times R$ $\mathbf{H}_n, 1 \leq n < \alpha$ and $R \times T$ matrix $\mathbf{H}_\alpha$ are iid with zero mean with variance $1/R$. Note that,$\alpha > 1$ the entries of $\mathbf{H}$ are Gaussian but correlated where the correlation level increases with the number of successive independent matrix product whose entries are iid.

In the following we first analyse the channel matrix defined in 5.14 with assuming $R = T$:

Corollary 8. *Let the entries of the independent $R \times R$ matrices $\mathbf{H}_n$ be iid with zero mean and variance $1/R$ for all n. Then, the empirical eigenvalue distributions of*

$$\prod_{n=1}^{\alpha} \mathbf{H}_n \tag{5.15}$$

and $\mathbf{H}^\alpha$ converge almost surely to the same limit given by

$$p_{\mathbf{H}^\alpha} = \begin{cases} \frac{1}{\pi\alpha}|z|^{\frac{2}{\alpha}-2} & |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{5.16}$$

as $R \to \infty$.

In other words, independent square random matrices with iid entries behave with respect to multiplication asymptotically as if they were identical. This means, that running through the same iid random channel twice or running consecutively through two random channels with the same statistics makes no difference in the large-system limit. By contrast, this does not even hold approximately if the channel is a diagonal matrix.

In the rectangular case, we can analyse the channel matrix $\mathbf{H}_\alpha$ by means of its *singular equivalent*:

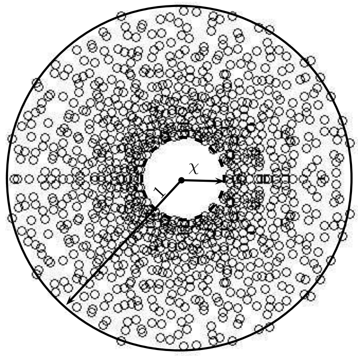


Figure 5.3: The eigenvalues of singular equivalent of 1000×750 $\mathbf{H}$ defined in (5.14) on the complex plane with inner radius $\chi = (1-\beta)^{\frac{\alpha}{2}}$.

Corollary 9. *Let the $R \times T$ matrix $\mathbf{H}_\alpha$ be defined as (5.14). Then, the empirical eigenvalue distribution of* singular equivalent *of $\mathbf{H}_\alpha$ converges almost surely to the limit,*

$$p_{\mathbf{H}_{\alpha,s}} = \begin{cases} \frac{1}{\pi\beta\alpha}|z|^{\frac{2}{\alpha}-2} & (1-\beta)^{\frac{\alpha}{2}} < |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{5.17}$$

as $R, T, \to \infty$ with the ratio $\beta = T/R \leq 1$ fixed.

It is easy to see, with increasing correlation level between entries the eigenvalue distribution of singular equivalent of the channel matrix $\mathbf{H}$ has gradient to zero. The gradient can be easily seen on figure 5.1.3. Therefore we can conclude, the more correlation the entries have, then the eigenvalues will lie more close in the complex plane.

Moreover, in physics literature there is an interesting measure for square non-hermitian random matrices that how much the pairs of eigenvalues lying close in the complex plane, called in literature *left-right eigenvector correlation* [63]:

Consider a $R \times R$ non-Hermitian random matrix $\mathbf{X}$ with eigenvalue decomposition

$$\mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{W}^{-1} = \sum_i \lambda_i \mathbf{v}_i \mathbf{w}_i^\dagger \tag{5.18}$$

where $\mathbf{V}$ is called right eigenvector matrix and $\mathbf{W} = \mathbf{V}^{-1}$ is called left eigenvectors. Then the correlation between right-left eigenvectors is defined [63]

$$C_{\mathbf{X}}(z) = \frac{\pi}{R}\sum_{i=1}^{R}(\mathbf{w}_i^\dagger \mathbf{w})(\mathbf{v}_i^\dagger \mathbf{v})\delta(z - z_i) \tag{5.19}$$

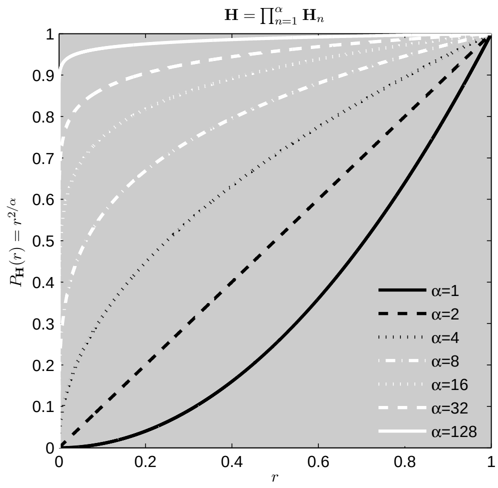


Figure 5.4: The asymptotic radial CDF of eigenvalues of the channel matrix defined (5.14) respective parameter α with $\beta = 1$.

Theorem 30. *Let the random matrix* $\mathbf{H}$ *defined in (5.14). Then the correlation between right-eigenvector defined in (5.19) of singular equivalent of* $\mathbf{H}$ *is*

$$C_{\mathbf{H}_u}(z) = \begin{cases} \frac{1}{\beta}(1 - |z|^{\frac{2}{\alpha}})|z|^{\frac{2}{\alpha}-2} & (1-\beta)^{\frac{\alpha}{2}} < |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{5.20}$$

as $T, R \to \infty$ *with ratio* $\beta = T/R \leq 1$.

Note that, we will calculate mutual information of 5.14 for more general model in Corollary 10.

5.1.4 Mutual Information in High SNR Regime

High SNR analysis of MIMO channels is one of most the interesting area of MIMO channels. Indeed when you google the title"MIMO Channels " you will get $\cong 12 \cdot 10^6$ results. On the other hand when you google the title "MIMO Channels with high SNR", will get $\cong 13 \cdot 10^5$ results.

In this subsection we will present an explicit expression of mutual information in High SNR regime with respect to product of free random matrices.

Recall, the ergodic mutual information at the SNR level γ can be written as

$$\mathcal{I}(\gamma) = E\left[\log\det\left(\mathbf{I} + \gamma\mathbf{H}\mathbf{H}^\dagger\right)\right] \tag{5.21}$$

Furthermore in high SNR Regime $1 \ll \gamma$, it is reasonable to conjecture

$$\begin{aligned} \mathcal{I}(\gamma) &= R \log \gamma + E\left[\log \det \mathbf{H}\mathbf{H}^{\dagger}\right] + o\left(\frac{1}{\gamma}\right) && (5.22) \\ &\cong R \log \gamma + E\left[\log \det \mathbf{H}\mathbf{H}^{\dagger}\right] && (5.23) \end{aligned}$$

We now address the following MIMO Channel $\mathbf{H}$

$$\mathbf{H} = \mathbf{A}_N \mathbf{A}_{N-1} \cdots \mathbf{A}_2 \mathbf{A}_1 \tag{5.24}$$

where $\mathbf{A}_n$ is the size of $K_n \times K_{n-1}$. Assume that, the family

$$(\{\mathbf{A}_1^{\dagger}\mathbf{A}_1\}, \{\mathbf{A}_2^{\dagger}\mathbf{A}_2\} \cdots \{\mathbf{A}_N^{\dagger}\mathbf{A}_N\}) \tag{5.25}$$

is asymptotically free all sizes K_n tend to infinity with the ratios

$$\chi_n = \frac{K_{n-1}}{K_n}, \quad 1 \leq n \leq N \tag{5.26}$$

remaining constant. Moreover define the ratios,

$$\rho_n = \frac{K_n}{K_N}. \tag{5.27}$$

Theorem 31. *Let the random matrix* $\mathbf{H}$ *defined as (5.24) and for simplicity let* $K_N = R$*. Suppose that,*

$$rank(\mathbf{H}\mathbf{H}^{\dagger}) = \frac{R}{\alpha} \tag{5.28}$$

Moreover define a function,

$$g(a) \triangleq (1-a)\log(a-1) + a \log a - 1 \tag{5.29}$$

such that $g(1) = -1$ *and define a vector* $\mathbf{x} \triangleq [1, \alpha, \alpha\rho_N, \alpha\rho_0]$*. Then the mutual information reads,*

$$\begin{aligned} \frac{I(\gamma)}{R} &= \frac{1}{\alpha}\left\{\log \gamma + \sum_{n=1}^{4}(-1)^n g(x_n) - \sum_{n=1}^{N}\int_{-1}^{0} \log S_{\mathbf{A}_n^{\dagger}\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right) dz \right. \\ &\qquad \left. - \int_0^{\frac{1}{\gamma}} G_{\mathbf{H}\mathbf{H}^{\dagger}}(-s) + \frac{\alpha - 1}{s} ds \right\} && (5.30) \end{aligned}$$

where in high SNR regime (5.30) reads,

$$\frac{I(\gamma)}{R} \cong \frac{1}{\alpha}\left\{\log \gamma + \sum_{n=1}^{4}(-1)^n g(x_n) - \sum_{n=1}^{N}\int_{-1}^{0} \log S_{\mathbf{A}_n^{\dagger}\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right) dz\right\} \tag{5.31}$$

As an application of Theorem 31, can be on the Asymptotic Eigenvalue Distribution of Concatenated Vector-Valued Fading Channels [17] as follows:

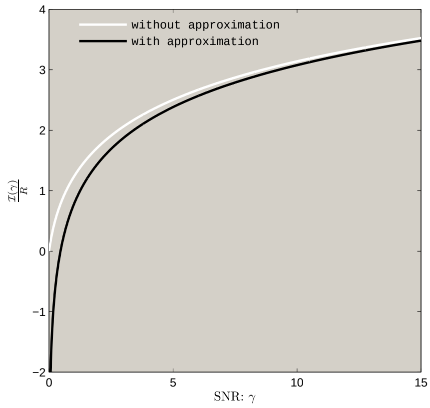


Figure 5.5: Mutual information for a channel composed of two scatterer matrices with the ratios $\rho_0 = 3$ and $\rho_1 = 4$ and the comparison between high-snr approximated one and non-approximated one
.

Corollary 10. *Let the random matrix* $\mathbf{H}$ *defined as (5.24). Furthermore let the entries of the* $K_n \times K_{n-1}$ *matrix* $\mathbf{A}_n$ *be independent and identically distributed with zero mean and respective variances* $1/K_n$*. Let*

$$\alpha = \frac{1}{\min(\rho_n)}, \qquad 0 \leq n \leq N \tag{5.32}$$

Let the function $g(\cdot)$ *be as in (5.29) such that* $g(1) = -1$*. Then the mutual information can be expressed as*

$$\frac{I(\gamma)}{R} = \frac{1}{\alpha}\left\{\log\gamma + g(\alpha) - g(1) + \sum_{n=1}^{N} g(\alpha\rho_{n-1}) - \log\alpha\rho_n - \int_0^{\frac{1}{\gamma}} G_{\mathbf{HH}^\dagger}(-s) + \frac{\alpha-1}{s}ds\right\}. \tag{5.33}$$

where in high SNR regime (5.33) reads as

$$\frac{I(\gamma)}{R} \cong \frac{\log\gamma + g(\alpha) - g(1)}{\alpha} + \frac{1}{\alpha}\sum_{n=1}^{N} g(\alpha\rho_{n-1}) - \log\alpha\rho_n \tag{5.34}$$

Moreover it was shown by Müller in [17], the Stieltjes transform of $\mathbf{HH}^\dagger$ satisfies:

$$G_{\mathbf{HH}^\dagger}(s)\prod_{n=1}^{N}\frac{sG_{\mathbf{HH}^\dagger}(s) + 1 - \rho_{n-1}}{\rho_n}. - sG_{\mathbf{HH}^\dagger}(s) = 1. \tag{5.35}$$

(5.35) allows us to find how much the approximated measure deviates from the actual measure.

Example 18. *Let the entries of the $R \times S$ matrix $\mathbf{A}_2$ and $S \times T$ $\mathbf{A}_2$ be independent and identically distributed with zero mean and variance $1/R$ and $1/S$. Consider a MIMO channel $\mathbf{H}$ described as*

$$\mathbf{H} = \mathbf{A}_2\mathbf{A}_1 \tag{5.36}$$

Then, in high SNR regime the mutual information can be approximated as

$$\frac{I(\gamma)}{R} \cong \frac{\log\gamma + g(\alpha) - g(1)}{\alpha} + \frac{1}{\alpha}\sum_{n=1}^{2} g(\alpha\rho_{n-1}) - \log\alpha\rho_n \tag{5.37}$$

as $R, S, T \to \infty$ with ratios $\rho_1 = S/R$ and $\rho_0 = T/R$ fixed.

As we see figure **??**, high SNR approximation gives good result even for quite moderate low SNR level. But this statement is not true in general. With the help Theorem 29, it is reasonable to conjecture that, when the $\beta >> 1$ or $\beta << 0$, then the high SNR approximation draw reliable portrait even for moderate SNR levels such as 8dB and its neighbourhood.

5.2 Richian Fading Channels

In this section we will work on MIMO Channels by taking account the line-of-sight. We will consider the following MIMO system described by

$$\mathbf{H} = \sigma\mathbf{H}_2\mathbf{H}_1 + \mathbf{H}_0 \tag{5.38}$$

with $\mathbf{H}_0$, $\mathbf{H}_1$, $\mathbf{H}_2$, σ denoting the line-of-sight path, the propagation from the transmit antennas to the scatterers, the propagation from the scatterers to the receive antennas, and the attenuation of the scattered paths relative to the line-of sight paths, respectively, is one of them. Note that the terms to be summed in

$$\mathbf{H}\mathbf{H}^\dagger = \sigma^2(\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger) + (\mathbf{H}_0\mathbf{H}_0^\dagger) + \sigma(\mathbf{H}_2\mathbf{H}_1\mathbf{H}_0^\dagger + \mathbf{H}_0\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger) \tag{5.39}$$

where terms in each parenthesis in (5.39) are Hermitian matrices but they are not free. On the other hand, the term in (5.38) are free but they are non-Hermitian matrices.

In this section, we will make use of an extension of free probability to non-Hermitian random matrices that introduced in the previous chapter to analyse the asymptotic eigenvalue distribution and the mutual information of MIMO channels with line-of-sight.

As in the previous section we assumed the entries of the matrices correspond the scattering component of the channel matrix (5.38), $\mathbf{H}_1 \in \mathbb{C}^{S\times T}$ and $\mathbf{H}_2 \in \mathbb{C}^{T\times R}$ are assumed to have iid with zero mean and variances $1/S$ and $1/R$, respectively, where T, S and R denote the number of transmit antennas, of scatterers, and of receive antennas, respectively. On the other hand, the entries of the matrix $\mathbf{H}_0$ corresponds the line-of-sight component of the channel matrix (5.38), is not iid.

Theorem 32. *Let the entries of $R \times S$ matrix $\mathbf{A}_1$ and $S \times T$ matrix $\mathbf{A}_2$ be independent and identically distributed with zero mean variances $1/R$ and $1/S$ and m^{th} moments upper bounded by $\alpha_m R^{-m/2}$ for some α_m and all $m \leq 1$. Let the $R \times T$ matrix $\mathbf{B}$ be arbitrary matrix free of $\mathbf{A_1 A}_2$ such that the empirical distribution of eigenvalues of $\mathbf{BB}^\dagger$ converges, as $R, T \to \infty$ to a limit distribution with Stieltjes transform $G_{\mathbf{BB}^\dagger}(s)$ defined in (3.31). Furthermore, let*

$$\mathbf{C} = \sigma \mathbf{A}_1 \mathbf{A}_2 + \mathbf{B} \tag{5.40}$$

with $\sigma \in \mathcal{C}$ and define

$$g(x) = \frac{\sigma(\beta x^2 - \rho x^2 - \rho) + \sigma\sqrt{(\beta x^2 - \rho x^2 - \rho)^2 - 4(\rho\beta x^2)(x^2 - \rho)}}{2x(\rho - x^2)} \tag{5.41}$$

Then, the empirical distribution of eigenvalues of $\mathbf{CC}^\dagger$ converges almost surely to a limit distribution whose Stieltjes transform satisfies

$$G_{\mathbf{CC}^\dagger}(s) = \frac{1}{\sqrt{s}} G_{\tilde{\mathbf{B}}}\left(\sqrt{s} - g\left[\sqrt{s} G_{\mathbf{CC}^\dagger}(s)\right]\right). \tag{5.42}$$

as $R, S, T \to \infty$ with $\rho = S/R$ and $\beta = T/R \leq 1$ fixed

5.2.1 Rician Mechanism

Rician fading occurs if the received signal has line-of-sight component: the propagation from transmit antennas to receiver antennas. We will explain the mechanism by means of basic wave propagation.

Before having look the propagation of line-of-sight, let us focus on the propagation from scatterers to transmit antennas. Imagine we receive signals comes from scatterer objects as shown in figure 5.2.1. Then, the phase profile of received signals by the receive antenna index r and a scatterer object index s is

$$\Theta_{s,r} = exp(j\omega_{s,r}) = exp\left(j\left[\theta_s - (t-1)\frac{2\pi d}{\lambda}\sin(\alpha_s)\right]\right) \tag{5.43}$$

Moreover it was well-explained by Müller in [17] why the entries of a scattering matrix be assumed iid by making an analogy with a method to generate random number called *Linear congruential random generator* [64] which used up to 4th version of Matlab. The algorithm is as follows,

$$X_{n+1} = (aX_n + c) mode \quad m \quad n \leq 0. \tag{5.44}$$

with seed X_0 and a=1.

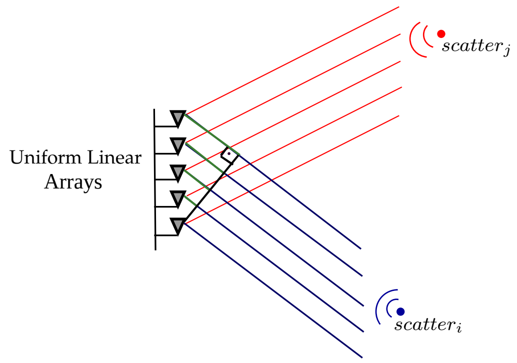


Figure 5.6: Propagation from scatterers to uniform linear array receiver, scatterers are assumed in the far-field.

When we look at figure 5.2.1, it is obvious, that each scatter acts as random number generator with its distance as seed and the sine of its angle times the element spacing as increment. For more detail discussion, we refer the reader to [17].

Now, let us focus on propagation from transmit antenna to receiver antennas as shown figure 5.2.1. First remark, in case SU-MIMO (single user), the user behaves as a single scatterer object. Therefore we conclude that the matrix $\mathbf{H}_0$ corresponds to line-of-sight component of (5.38) will be unit rank in case of SU-MIMO (single user).

On the other hand in MU-MIMO, each user behaves as a single scatterer object, therefore at first it seems rank of the matrix $\mathbf{H}_0$ will be the number of users. But some users may not experience line-of-sight in general. Then we define

$$\phi \equiv \frac{L}{T} \tag{5.45}$$

where L represents the number of user experience line-of-sight. The parameter ϕ specifies the relative rank of line-of-sight component of the MU-MIMO system, which we will call *line-of-sight fraction*.

5.2.2 MU-MIMO systems with LOS

Multiuser (MU) multiple-input multiple-output (MIMO) system have received a great deal of attention recently as they also serve as a models to describe the propagation of virtual MIMO systems were the multiple antennas are not co-located but belong to different cooperating users. The capacity region of a MU-MIMO system depends on the singular values of the channel matrix that governs the propagation from all (virtual) transmitting antennas to all (virtual) receiving antennas.

As we discussed previous subsection, in multi-user MIMO systems, the line-of-sight component of the channel matrix is not limited to rank one, as the

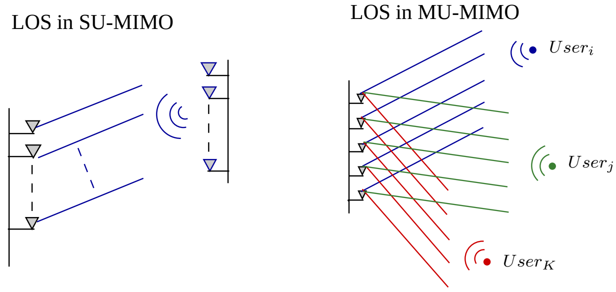


Figure 5.7: The Propagation of line-of-sight for single user MIMO and multiple users MIMO, where users are assumed in the far-field.

antennas need not be co-located. Still, its rank is typically lower than the rank of the scattered component as the existence of a direct path is less probable than the existence of an indirect path. With the scattered component having higher rank, but lower power, the question which of the two components is more important is non-trivial. Furthermore, it is expected that the interplay of both components is important to understand the properties of MU-MIMO systems.

In the following, we will address the wireless MIMO system described by

$$\mathbf{y} = \sqrt{\frac{\gamma}{\phi + \sigma^2}}\mathbf{H}\mathbf{x} + \mathbf{n} \tag{5.46}$$

where $\mathbf{H}$, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{n}$, γ are the channel defined in (5.38) the channel input, the channel output, additive white Gaussian noise (AWGN), and the signal-to-noise ratio, respectively. The entries of $\mathbf{x}$ and $\mathbf{n}$ are assumed to be iid with zero mean and unit variance.

The asymptotic eigenvalue distribution of line-of-sight component $\mathbf{H}_0$ i.e.

$$\mathbf{H}_0\mathbf{H}_0^\dagger = \mathbf{GPG} \tag{5.47}$$

where the entries of $R \times T$ matrix $\mathbf{G}$ are iid with zero mean variance $1/R$ (with ratio $\beta = T/R$ fixed), and the $T \times T$ matrix $\mathbf{P}$ is a diagonal matrix with the distribution of diagonal entries

$$p_{\mathbf{P}}(x) = (1-\phi)\delta(x) + \phi\delta(x-1) \tag{5.48}$$

Recall the formula (3.41)

$$G_{\mathbf{H}_0\mathbf{H}_0^\dagger}(s) = G_{\mathbf{X}}\left(s + \beta \int \frac{x dP_{\mathbf{P}}(x)}{x G_{\mathbf{H}_0\mathbf{H}_0^\dagger}(s) - 1}\right) \tag{5.49}$$

with $\mathbf{X} = \mathbf{0}$ we have,

$$\frac{1}{G_{\mathbf{H}_0\mathbf{H}_0^\dagger}(s)} = s + \frac{\beta\phi}{G_{\mathbf{H}_0\mathbf{H}_0^\dagger}(s) - 1} \tag{5.50}$$

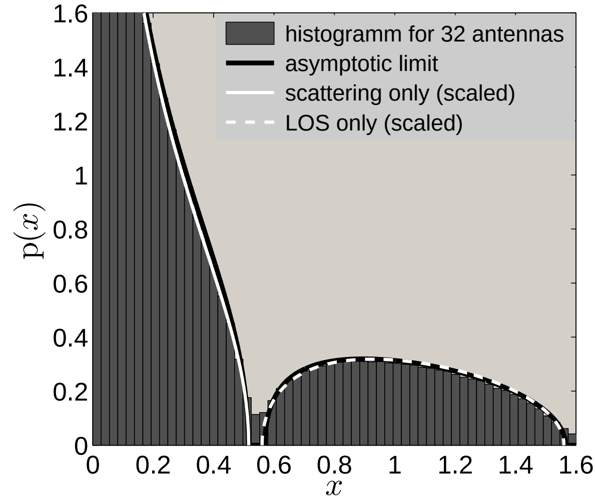


Figure 5.8: Probability density function of the singular values of the matrix $\mathbf{H}$ in (5.38) for $4\sigma = 4\phi = 1$. The dashed lines show scaled and shifted versions of pure scattering ($\phi = 0$) and pure line-of-sight ($\sigma = 0$), respectively.

(5.50) which is the Marchenko-Pastur distribution defined in (3.9) with parameter $\beta\phi$,

$$G_{\mathbf{H}_0\mathbf{H}_0^\dagger}(s) = \frac{1}{2} + \frac{1-\phi\beta}{2s} \mp \sqrt{\frac{(1-\phi\beta)^2}{4s^2} - \frac{1+\phi\beta}{2s} + \frac{1}{4}} \tag{5.51}$$

With the help of (5.51), Theorem 32 allows to calculate the asymptotic singular value distribution of the channel (5.38) This examples was chosen, since the relative scattering attenuation σ and line-of-sight fraction ϕ are small, in practice. In that case, the asymptotic singular value distribution of $\mathbf{H}$ decomposes into two bulks with each bulk being shaped very similar to the cases of pure scattering and pure line-of-sight when scaled or shifted appropriately. This deviates strongly from the quarter circle law that would be obtained, if $\mathbf{H}$ were composed of iid entries. The mutual information of the channel defined in (5.38) and measured in nats is given by

$$\lim_{R\to\infty} \frac{\mathcal{I}(\gamma,\sigma,\phi)}{R} = \int \log\left(1+\frac{x}{s}\right) dP_{\mathbf{H}\mathbf{H}^\dagger}(x)\bigg|_{s=\frac{\phi+\sigma^2}{\gamma}} \tag{5.52}$$

$$= \int_{\frac{\phi+\sigma^2}{\gamma}}^{\infty} G_{\mathbf{H}\mathbf{H}^\dagger}(-s) + \frac{1}{s} ds \tag{5.53}$$

$$= 2\int_{\sqrt{\frac{\phi+\sigma^2}{\gamma}}}^{\infty} \frac{1}{s} - G_{\tilde{\mathbf{H}}}(-s) ds \tag{5.54}$$

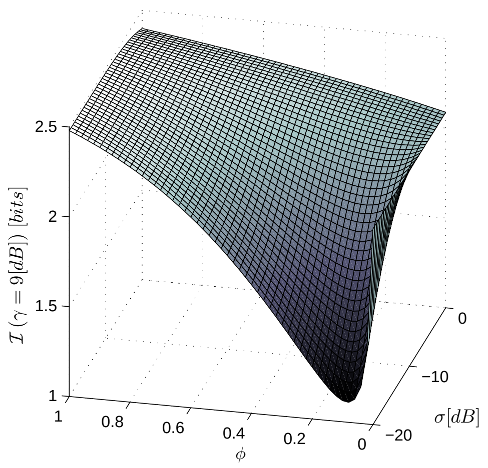


Figure 5.9: Mutual information for $\gamma = 9$[dB], $\beta = \rho = 1$ versus σ and ϕ.

where (5.54) and (5.54) follow from the mutual-information lemma (3.52) [2] It is shown in Figure 5.2.2 for a fixed signal-to-noise ratio of 9 dB.

One can observe that small values of the line-of-sight fraction ϕ and the relative scattering attenuation σ that are typical in many practical scenarios are quite deleterious for the mutual information of the channel.

Furthermore, the figure seems to suggest that blocking the line of sight is better than a small, but non-zero value of the line-of-sight fraction ϕ. However, Figure. 5.2.2 is plotted for constant SNR and blocking the line of sight will surely decrease the SNR.

The hit in mutual information for small line-of-sight fraction and relative scattering attenuation is exacerbated in practice by the fact that analog-to-digital conversion and precise estimation of the scattered paths is challenging in the presence of much stronger direct paths.

5.2.3 Conclusion

Line of sight strongly influences the eigenvalue distribution of multi-user MIMO channels. If the line-of-sight component is significantly stronger than the scattered paths and/or the fraction of users who experience line of sight is small, the eigenvalue distribution is composed of two separate bulks, one corresponding to the scattered paths and one corresponding to the direct paths. In that case, the

[2] Symmetrized version of Stieltjes transform defined in (4.49) as:

$$G_{\widetilde{\mathbf{H}}}(s) = sG_{\mathbf{HH}^\dagger}\left(s^2\right).$$

asymptotic eigenvalue distribution can be accurately approximated by a scaled version of pure scattering and a shifted version of pure line of sight which, in contrast to the exact solution, can be given in closed explicit form.

Appendix A

Proofs

A.1 Theorem 29

Let us start with the $R \times R$ R-diagonal matrix $\mathbf{H}$ defined in the Theorem as

$$\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_R]. \tag{A.1}$$

Moreover we defined $R \times T; (T \leq R)$ matrix $\mathbf{H}_\beta$ as

$$\mathbf{H}_\beta = [\mathbf{h}_1, \cdots, \mathbf{h}_T]. \tag{A.2}$$

Define a $R \times R$ diagonal matrix $\mathbf{P}$ whose diagonal terms distributed as

$$p_{\mathbf{P}}(x) = (1-\beta)\delta(x) + \beta\delta(x-1) \tag{A.3}$$

Then remark that,

$$\mathbf{P}\mathbf{H}\mathbf{H}^\dagger\mathbf{P} = \mathbf{H}_\beta\mathbf{H}_\beta^\dagger \tag{A.4}$$

Note that,

$$p_{\mathbf{H}_\beta\mathbf{H}_\beta^\dagger}(x) = (1-\beta)\delta(x) + \beta p_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(x). \tag{A.5}$$

With Theorem 18 [Theorem 14.10 in [43]], we have,

$$R_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(\omega) = R_{\mathbf{H}\mathbf{H}^\dagger}(\beta\omega). \tag{A.6}$$

Recall the functional relation between R-transform and S-transform [43],

$$zR(z)S(zR(z)) = z; \quad zS(z)R(zS(z)) = z \tag{A.7}$$

Let start with,

$$\begin{aligned} S_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(z) &= \frac{1}{R_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(zS_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(z))} = \frac{1}{R_{\mathbf{H}\mathbf{H}^\dagger}(\beta zS_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(z))} && \text{(A.8)} \\ &= S_{\mathbf{H}\mathbf{H}^\dagger}\left(\beta \cdot zS_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(z)R_{\mathbf{H}\mathbf{H}^\dagger}\left(zS_{\mathbf{H}_\beta^\dagger\mathbf{H}_\beta}(z)\right)\right) && \text{(A.9)} \\ &= S_{\mathbf{H}\mathbf{H}^\dagger}(\beta \cdot z). && \text{(A.10)} \end{aligned}$$

Moreover we defined in the Theorem

$$\mathbf{H}_{\beta,u} = \mathbf{U}\sqrt{\mathbf{H}_\beta^\dagger \mathbf{H}_\beta} \tag{A.11}$$

With (4.57) in Corollary 3, we have,

$$\begin{aligned} P^{-1}_{\mathbf{H}_{\beta,u}}(r) &= \frac{1}{\sqrt{S_{\mathbf{H}_\beta^\dagger \mathbf{H}_\beta}(r-1)}} = \frac{1}{\sqrt{S_{\mathbf{H}\mathbf{H}^\dagger}(\beta r - \beta)}} \\ &= \frac{1}{\sqrt{S_{\mathbf{H}\mathbf{H}^\dagger}((\beta r + 1 - \beta) - 1)}} \\ &= P^{-1}_{\mathbf{H}}(\beta r + 1 - \beta). \end{aligned} \tag{A.12}$$

Let $r \to P_{\mathbf{H}_{\beta,u}}(r)$, we have

$$P_{\mathbf{H}}(r) = \beta P_{\mathbf{H}_{\beta,u}}(r) + 1 - \beta \tag{A.13}$$

This reads,

$$P_{\mathbf{H}_{\beta,u}}(r) = \frac{1}{\beta} P_{\mathbf{H}}(r) + 1 - \frac{1}{\beta} \tag{A.14}$$

Moreover, zero measure can be easily find as

$$\phi_\beta = P_{\mathbf{H}_{\beta,u}}(0) = \max(0, \beta^{-1} P_{\mathbf{H}}(0) + 1 - \beta^{-1}) \tag{A.15}$$

$$= \max(0, \beta^{-1}\phi + 1 - \beta^{-1}) \tag{A.16}$$

since we defined zero measure of $\mathbf{H}$ as ϕ. Thus we have distribution of $\mathbf{H}_{\beta,\mathbf{u}}$ satisfies with

$$\frac{dP_{\mathbf{H}_{\beta,u}}(r)}{dr} = p_{\mathbf{H}_{\beta,u}}(r) \tag{A.17}$$

$$= \phi_\beta \delta(r) + \frac{1}{\beta} p_{\mathbf{H}}(r) \tag{A.18}$$

$$p_{\mathbf{H}_{\beta,u}}(z) = \left.\frac{p_{\mathbf{H}_{\beta,u}}(r)}{2\pi r}\right|_{r=|z|} \tag{A.19}$$

$$= \phi_\beta \delta(z) + \frac{1}{\beta} p_{\mathbf{H}}(z) \tag{A.20}$$

Final step is to determine the boundary of distribution. It is obvious, that the outer boundary doesn't change since (with (A.12))

$$P^{-1}_{\mathbf{H}_{\beta,u}}(1) = P^{-1}_{\mathbf{H}}(\beta 1 + 1 - \beta) = P^{-1}_{\mathbf{H}}(1). \tag{A.21}$$

In the same way, the inner boundary reads

$$P^{-1}_{\mathbf{H}_{\beta,u}}(\phi_\beta) = P^{-1}_{\mathbf{H}}(\beta\phi_\beta + 1 - \beta). \tag{A.22}$$

Thus the asymptotic eigenvalue distribution of $\mathbf{H}_{\beta,\mathbf{u}}$ converges to limit distribution

$$p_{\mathbf{H}_{\beta,u}}(z) = \phi_\beta \delta(z) + \begin{cases} \frac{1}{\beta} p_{\mathbf{H}}(z) & P^{-1}_{\mathbf{H}}(\beta\phi_\beta + 1 - \beta) < |z| \le b \\ 0 & elsewhere \end{cases} \tag{A.23}$$

as $R, T \to \infty$ with the ratio $\beta = T/R \le 1$ fixed.

A.2 Proof of Theorem 30

Consider a non-Hermitian random matrix $\mathbf{X}$ and its eigenvalue decomposition as,

$$\mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{W} \tag{A.24}$$

such that,

$$\begin{aligned} \mathbf{V} &= [\mathbf{v}_1, ...\mathbf{v}_R] & \text{(A.25)} \\ \mathbf{V}^{-1} = \mathbf{W} &= [\mathbf{w}_1, ...\mathbf{w}_R] & \text{(A.26)} \end{aligned}$$

Then the definition of left-right eigenvector correlation of $\mathbf{H}$[1]

$$C_{\mathbf{H}}(z) = \frac{\pi}{R}\sum_{i=1}^{R}(\mathbf{w}_i^{\dagger}\mathbf{w})(\mathbf{v}_i^{\dagger}\mathbf{v})\delta(z - z_i) \tag{A.27}$$

Remark that,

$$C_{\mathbf{H}^\alpha}(z) = \frac{\pi}{R}\sum_{i=1}^{R}(\mathbf{w}_i^{\dagger}\mathbf{w})(\mathbf{v}_i^{\dagger}\mathbf{v})\delta(z - z_i^\alpha) \tag{A.28}$$

since

$$\mathbf{H}^\alpha = \mathbf{V}\mathbf{\Lambda}^\alpha\mathbf{W} \tag{A.29}$$

Let the entries of the $R \times R$ matrix $\mathbf{H}$ be iid with zero mean variance $1/R$. Then we have, [63] ,

$$C_{\mathbf{H}}(z) = 1 - |z|^2 \quad |z| \leq 1. \tag{A.30}$$

Note that $C(z)$ is a density function. See that, $C_{\mathbf{H}}(z)$ is circularly symmetric such that the radial density read,

$$C_{\mathbf{H}}(r) = 2\pi r(1 - r^2) \quad 0 < \quad r \leq 1. \tag{A.31}$$

Thus we can easily find the $C_{\mathbf{H}^\alpha}(r)$ as

$$\begin{aligned} C_{\mathbf{H}^\alpha}(r) &= \frac{1}{\alpha}C_{\mathbf{H}}\left(r^{\frac{1}{\alpha}}\right)r^{\frac{1}{\alpha}-1} & \text{(A.32)} \\ &= 2\pi\left(1 - r^{\frac{2}{\alpha}}\right)r^{\frac{2}{\alpha}-1} & \text{(A.33)} \end{aligned}$$

Thus we have

$$\begin{aligned} C_{\mathbf{H}^\alpha}(z) &= \frac{1}{2\pi r}\left. C_{\mathbf{H}^\alpha}(r)\right|_{r=|z|} & \text{(A.34)} \\ &= \left(1 - r^{\frac{2}{\alpha}}\right)r^{\frac{2}{\alpha}-2}. & \text{(A.35)} \end{aligned}$$

With theorem 27 we have,

$$\mathbf{H}_\alpha = \prod_{n=1}^{\alpha}\mathbf{H}_i = \mathbf{U}\mathbf{H}^\alpha \tag{A.36}$$

[1]Note in case the finite size analysis the definition is used as $C'(z) = NC(z)$ in asymptotic case in literature it is used as $C(z)$. i.e. we refer to reader see (41) in [54] or [55]

where the matrices $\mathbf{H}_i$ are identical and free of each others and $\mathbf{U}$ be a unitary free of $\mathbf{H}$. Since we look for a singular equivalent, it is equivalent to concern either the power of $\mathbf{H}$ or product of identical-free matrices $\mathbf{H}_i$. Moreover $C(z)$ is a measure of how much the eigenvalues closely lie each others in the complex plane, therefore it is expected that randomly rotation of a R-diagonal matrix doen't change its the measure $C(z)$.

Recall Theorem 29 tells us the eigenvalue distribution singular equivalent of a rectangular random matrix tells us, is the same form of square-form except the inner radius and the normalization constant $\beta = T/R \le 1$. Thus we have,

$$C_{\mathbf{H}_u}(z) = \begin{cases} \frac{1}{\beta}(1-|z|^{\frac{2}{\alpha}})|z|^{\frac{2}{\alpha}-2} & (1-\beta)^{\frac{\alpha}{2}} < |z| \le 1 \\ 0 & elsewhere \end{cases} \tag{A.37}$$

as $T, R \to \infty$ with ratio $\beta = T/R \le 1$.

A.3 Proof of Theorem 31

Let $R_\alpha = rank(\mathbf{XX}^\dagger) = R/\alpha$. Moreover let define the $R_\alpha \times R_\alpha$ Hermitian matrix $\mathbf{Y}$ whose eigenvalue distribution identical to non-zero eigenvalue distribution of $\mathbf{XX}^\dagger$ such that,

$$p_{\mathbf{XX}^\dagger}(x) = (1-\frac{1}{\alpha})\delta(x) + \frac{1}{\alpha}p_{\mathbf{Y}}(x) \tag{A.38}$$

with (3.39),(3.153), we have

$$G_{\mathbf{Y}}(s) = \alpha G_{\mathbf{XX}^\dagger}(s) + \frac{\alpha-1}{s} \tag{A.39}$$

$$S_{\mathbf{Y}}(z) = \frac{z+1}{z+\alpha} S_{\mathbf{XX}^\dagger}\left(\frac{z}{\alpha}\right) \tag{A.40}$$

Since $\mathbf{XX}^\dagger$ and $\mathbf{Y}$ have same non-zero eigenvalues,

$$\lim_{R\to\infty} \frac{I(\gamma)}{R} = \lim_{R\to\infty} \frac{1}{R} \log\det(\mathbf{I} + \gamma\mathbf{XX}^\dagger) \tag{A.41}$$

$$= \lim_{R\to\infty} \frac{1}{R} \log\det(\mathbf{I} + \gamma\mathbf{Y}) \tag{A.42}$$

$$= \frac{1}{\alpha} \lim_{R_\alpha\to\infty} \frac{1}{R_\alpha} \log\det(\mathbf{I} + \gamma\mathbf{Y}) \tag{A.43}$$

$$= \frac{1}{\alpha} \int_0^\infty \log(1+\gamma x) dP_{\mathbf{Y}}(x)$$

$$= \frac{1}{\alpha} \left\{ \log\gamma + \int_0^\infty \log\left(\frac{1}{\gamma} + x\right) dP_{\mathbf{Y}}(x) \right\} \tag{A.44}$$

We first focus on the integral term in (A.44), and starting with formula [48],

$$\log\left(\frac{1}{\gamma} + x\right) = \log x + \int_0^{\frac{1}{\gamma}} \frac{ds}{x+s} \tag{A.45}$$

Thus,

$$\int_0^\infty \log\left(\frac{1}{\gamma}+x\right)dP_{\mathbf{Y}}(x) = \int_0^\infty \log x dP_{\mathbf{Y}}(x) - \int_0^{\frac{1}{\gamma}}\int_0^\infty \frac{dP_{\mathbf{Y}}(x)}{-x-s}ds \tag{A.46}$$

$$= \int_0^\infty \log x dP_{\mathbf{Y}}(x) - \int_0^{\frac{1}{\gamma}} G_{\mathbf{Y}}(-s)ds \tag{A.47}$$

$$= \int_0^\infty \log x dP_{\mathbf{Y}}(x) - \int_0^{\frac{1}{\gamma}} \left(\alpha G_{\mathbf{XX}^\dagger}(-s)+\frac{1-\alpha}{s}\right)ds \tag{A.48}$$

With Theorem 21 (Corollary 5 in [48])such that, for a bounded and invertible matrix $\mathbf{Y}$, we have

$$\int_0^\infty \log x dP_{\mathbf{Y}}(x) = -\int_{-1}^0 \log S_{\mathbf{Y}}(z)dz \tag{A.49}$$

where with lemma 1 and (A.40),

$$S_{\mathbf{Y}}(z) = \frac{z+1}{z+\alpha}S_{\mathbf{XX}^\dagger}\left(\frac{z}{\alpha}\right) \tag{A.50}$$

$$= \frac{z+1}{z+\alpha}\frac{z+\alpha\rho_N}{z+\alpha\rho_0}\prod_{n=1}^{N} S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right) \tag{A.51}$$

Thus, we have,

$$\int_0^\infty \log x dP_{\mathbf{Y}}(x) = -\int_{-1}^0 \log\frac{z+1}{z+\alpha}+\log\frac{z+\alpha\rho_N}{z+\alpha\rho_0}dz - \sum_{n=1}^{N}\int_{-1}^0 \log S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right)dz \tag{A.52}$$

The first integral term in (A.52) can be simplified by defining a function as

$$g(a) \triangleq \int_{-1}^0 \log(z+a)dz = (1-a)\log(a-1)+a\log a-1; \tag{A.53}$$

such that,

$$g(1) = \lim_{u\to 0^+} u\log u - 1 = \lim_{u\to 0^+}\frac{\log u}{1/u}-1 \tag{A.54}$$

$$= \lim_{u\to 0^+}\frac{1/u}{-1/(u^2)}-1 = -\lim_{u\to 0^+} u-1 = -1 \tag{A.55}$$

Then we have,

$$\int_0^\infty \log x dP_{\mathbf{Y}}(x) = g(\alpha)-g(1)+g(\alpha\rho_0)-g(\alpha\rho_N) - \sum_{n=1}^{N}\int_{-1}^0 \log S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right)dz \tag{A.56}$$

Moreover define a vector $\mathbf{x} \triangleq [1, \alpha, \alpha\rho_N, \alpha\rho_0]$. With plugging (A.56) in (A.48), then plugging (A.48) in (A.44) we have,

$$\begin{aligned}\frac{I(\gamma)}{R} &= \frac{1}{\alpha}\left\{\log\gamma + \sum_{n=1}^{4}(-1)^n g(x_n) - \sum_{n=1}^{N}\int_{-1}^{0}\log S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right)dz\right.\\ &\quad \left. - \int_0^{\frac{1}{\gamma}} G_{\mathbf{XX}^\dagger}(-s) + \frac{\alpha-1}{s}ds\right\} \qquad \text{(A.57)}\end{aligned}$$

where it is immediate to see that, in high SNR (A.57) becomes

$$\frac{I(\gamma)}{R} \cong \frac{1}{\alpha}\left\{\log\gamma + \sum_{n=1}^{4}(-1)^n g(x_n) - \sum_{n=1}^{N}\int_{-1}^{0}\log S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\alpha\rho_{n-1}}\right)dz\right\} \qquad \text{(A.58)}$$

A.4 Proof of Theorem 32

For notational convince, let us define

$$\mathbf{A} = \mathbf{A}_1\mathbf{A}_2 \qquad \text{(A.59)}$$

Furthermore to work with rectangular matrices we define the $R \times R$ diagonal matrix $\mathbf{P}$ whose diagonal elements distributed as

$$p_{\mathbf{P}}(x) = (1-\beta)\delta(x) + \beta\delta(x-1) \qquad \text{(A.60)}$$

where ratio $\beta = T/R \leq 1$ fixed. Moreover define

$$\mathbf{XP} = [\mathbf{A}|\mathbf{N}]; \quad \mathbf{YP} = [\mathbf{B}|\mathbf{N}] \qquad \text{(A.61)}$$

where $\mathbf{N}$ is $R \times (R-T)$ null (zero) matrix such that,

$$\begin{aligned}\mathbf{CC}^\dagger &= \mathbf{P}(\mathbf{X}+\mathbf{Y})(\mathbf{X}+\mathbf{Y})^\dagger\mathbf{P} \qquad \text{(A.62)}\\ &\equiv \mathbf{P}(\mathbf{X}+\mathbf{UY})(\mathbf{X}+\mathbf{UY})^\dagger\mathbf{P} \qquad \text{(A.63)}\end{aligned}$$

where that $\mathbf{U}$ is Haar unitary matrix and free of $\mathbf{X}$ and $\mathbf{Y}$. Note that $\mathbf{X}$ and $\mathbf{UY}$ are R-diagonal. With 25[Proposition 3.5 in [47]] (A.63) reads,

$$\begin{aligned}\mathbf{CC}^\dagger &\equiv \mathbf{P}(\tilde{\mathbf{X}}+\tilde{\mathbf{Y}})(\tilde{\mathbf{X}}+\tilde{\mathbf{Y}})\mathbf{P} \qquad \text{(A.64)}\\ &\equiv [\mathbf{P}(\tilde{\mathbf{X}}+\tilde{\mathbf{Y}})\mathbf{P}][\mathbf{P}(\tilde{\mathbf{X}}+\tilde{\mathbf{Y}})\mathbf{P}] \qquad \text{(A.65)}\end{aligned}$$

Thus we have,

$$\begin{aligned}\tilde{\mathbf{C}} &= \mathbf{P}(\tilde{\mathbf{X}}+\tilde{\mathbf{Y}})\mathbf{P}\\ &= \mathbf{P}\tilde{\mathbf{X}}\mathbf{P} + \mathbf{P}\tilde{\mathbf{Y}}\mathbf{P} \qquad \text{(A.66)}\end{aligned}$$

such that, $\mathbf{P}\tilde{\mathbf{X}}\mathbf{P}$ and $\mathbf{P}\tilde{\mathbf{Y}}\mathbf{P}$ are free (1.12 Corollary in [44]). Then R-transform of reads the following,

$$R_{\tilde{\mathbf{C}}}(\omega) = R_{\mathbf{P}\tilde{\mathbf{X}}\mathbf{P}}(\omega) + R_{\mathbf{P}\tilde{\mathbf{Y}}\mathbf{P}}(\omega) \qquad \text{(A.67)}$$

Note that $\tilde{\mathbf{A}} = \mathbf{P}\tilde{\mathbf{X}}\mathbf{P}$ and $\tilde{\mathbf{B}} = \mathbf{P}\tilde{\mathbf{Y}}\mathbf{P}$. Thus,

$$R_{\tilde{\mathbf{C}}}(\omega) = R_{\tilde{\mathbf{A}}}(\omega) + R_{\tilde{\mathbf{B}}}(\omega) \tag{A.68}$$

Moreover, the Stieltjes transform of $\tilde{\mathbf{C}}$ can be present as,

$$\begin{aligned} \omega &= G_{\tilde{\mathbf{B}}}\left[G^{-1}_{\tilde{\mathbf{B}}}(\omega)\right] \\ &= G_{\tilde{\mathbf{B}}}\left[G^{-1}_{\tilde{\mathbf{C}}}(\omega) - G^{-1}_{\tilde{\mathbf{A}}}(\omega) + \frac{1}{\omega}\right] \\ &= G_{\tilde{\mathbf{B}}}\left[G^{-1}_{\tilde{\mathbf{C}}}(\omega) - R_{\tilde{\mathbf{A}}}(\omega)\right] \end{aligned} \tag{A.69}$$

With substitution $\omega \to G_{\tilde{\mathbf{C}}}(s)$, we have

$$G_{\tilde{\mathbf{C}}}(s) = G_{\tilde{\mathbf{B}}}\left(s - R_{\tilde{\mathbf{A}}}\left[G_{\tilde{\mathbf{C}}}(s)\right]\right). \tag{A.70}$$

With Lemma 9 we have,

$$G_{\mathbf{CC}^\dagger}(s) = \frac{1}{\sqrt{s}} G_{\tilde{\mathbf{B}}}\left(\sqrt{s} - R_{\tilde{\mathbf{A}}}\left[\sqrt{s}G_{\mathbf{CC}^\dagger}(s)\right]\right). \tag{A.71}$$

Thus we left with R-transform of $\tilde{\mathbf{A}}$.

Let start with (3.172) in Example 14

$$S_{\mathbf{AA}^\dagger}(z) = \frac{\rho}{(z+\rho)(z+\beta)} \tag{A.72}$$

With Lemma 10 we have,

$$S_{\tilde{\mathbf{A}}}(z) = \left[\frac{z}{z+1}\frac{\rho}{(z+\rho)(z+\beta)}\right]^{\frac{1}{2}} \tag{A.73}$$

With the inversion formula between R-transform and S-transform 3.148 we have,

$$\frac{\rho\omega R_{\tilde{\mathbf{A}}}(\omega)(\omega R_{\tilde{\mathbf{A}}}(\omega)+1)}{(\omega R_{\tilde{\mathbf{A}}}(\omega)+\rho)(\omega R_{\tilde{\mathbf{A}}}(\omega)+\beta)} - \omega^2 = 0. \tag{A.74}$$

Or explicitly

$$\omega(\rho - \omega^2)R^2_{\tilde{\mathbf{A}}} + (\rho - \rho\omega^2 - \beta\omega^2)R^2_{\tilde{\mathbf{A}}} - \rho\beta\omega = 0. \tag{A.75}$$

(A.75) is second order equation and the right solution fulfils [2]

$$R_{\tilde{\mathbf{A}}}(\omega) = \frac{(\beta\omega^2 - \rho\omega^2 - \rho) + \sqrt{(\beta\omega^2 - \rho\omega^2 - \rho)^2 - 4(\rho\beta\omega^2)(\omega^2 - \rho)}}{2\omega(\rho - \omega^2)} \tag{A.76}$$

Finally with scaling property of R-transform (3.94) such that,

$$R_{\sigma\tilde{\mathbf{A}}}(\omega) = \sigma R_{\tilde{\mathbf{A}}}(\sigma\omega)(\sigma) \tag{A.77}$$

One can plug (A.76) with scaling property (A.77) into (A.71) which proves Theorem 32.

[2]We choose the solution where in the case $\beta = 1$ Then the R-transform of satisfies $\tilde{\mathbf{A}}$ reads,

$$R_{\tilde{\mathbf{A}}}(\omega) = \frac{\rho\omega}{\rho - \omega^2}$$

since when $\beta = 1$.

A.5 Corollary 6

With example 14 we have,

$$S_{\mathbf{HH}^\dagger}(r-1) = \frac{\rho}{(r+\rho-1)(r+\beta-1)} \tag{A.78}$$

With the theorem (26), we have

$$\frac{\rho}{\left(P_{\mathbf{H}_p}\left(\frac{1}{\sqrt{r}}\right)+\rho-1\right)\left(P_{\mathbf{H}_p}\left(\frac{1}{\sqrt{r}}\right)+\beta-1\right)} = r \tag{A.79}$$

or explicitly

$$\frac{1}{\rho}\cdot\left(P_{\mathbf{H}_p}(r)+\rho-1\right)\left(P_{\mathbf{H}_p}(r)+\beta-1\right) = \frac{1}{r^2} \tag{A.80}$$

Which is standard quadratic equation with respect to radial probability measure $P_{\mathbf{H}_p}(r)$. The right solution must fulfil the condition, $P_{\mathbf{H}_p}(r) \geq 0$ such that

$$P_{\mathbf{H}_p}(r) = \frac{1}{2}\left[(1-\beta)+(1-\rho)+\sqrt{(\beta-\rho)^2+4\rho r^2}\right]. \tag{A.81}$$

To find zero measure of the distribution,

$$\begin{aligned} P_{\mathbf{H}_p}(0) &= \frac{1}{2}\left[(1-\beta)+(1-\rho)+\sqrt{(\beta-\rho)^2}\right] && \text{(A.82)}\\ &= \frac{1}{2}\left[(1-\beta)+(1-\rho)+|\rho-\beta|\right] && \text{(A.83)}\\ &= \max(1-\beta, 1-\rho) && \text{(A.84)} \end{aligned}$$

We can easily find the density of asymptotic eigenvalue distribution as,

$$\begin{aligned} p_{\mathbf{H}_p}(z) &= \max(1-\beta,1-\rho)\delta(z) + \left.\left(\frac{1}{2\pi r}\frac{dP_{\mathbf{H}_s}(r)}{dr}\right)\right|_{|z|=r} \\ &= \max(1-\beta,1-\rho)\delta(z) + \frac{1}{\pi}\frac{\xi}{\sqrt{(\beta-\xi)^2+4\xi|z|^2}} && \text{(A.85)} \end{aligned}$$

Finally we need determine the boundary of the distribution. Since the distribution has some zero measure, then inner radius of the density reads,

$$in(\mathbf{H_p})^{-1} = 0 \tag{A.86}$$

The outer radius of the density reads,

$$out(\mathbf{H_p}) = \frac{1}{\sqrt{S_{\mathbf{HH}^\dagger}(0)}} = \sqrt{\beta}. \tag{A.87}$$

Thus the empirical eigenvalue distribution of $R\times R$ matrix $\mathbf{H}_p$ converges almost surely to a limit

$$p_{\mathbf{H}_p}(z) = \max(1-\beta,1-\rho)\delta(z) + \begin{cases} \frac{1}{\pi}\frac{\rho}{\sqrt{(\beta-\rho)^2+4\rho|z|^2}} & |z|\leq\sqrt{\beta} \\ 0 & elsewhere \end{cases} \tag{A.88}$$

as $T,R,S\to\infty$ with ratios $\rho = S/R$ and $\beta = T/R$ fixed.

A.6 Proof of Corollary 7

Let first present Corollary 5.2.1 with $\beta = 1$ as

$$p_{\mathbf{H}}(z) = \max(0, 1-\rho)\delta(z) + \begin{cases} \frac{1}{\pi}\frac{\rho}{\sqrt{(1-\rho)^2+4\rho|z|^2}} & |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{A.89}$$

Define

$$\rho = \max(0, 1-\rho) \tag{A.90}$$
$$\rho_\beta = \max(0, 1-\rho\beta^{-1} - \beta^{-1}) \tag{A.91}$$

With Theorem 29 we have,

$$p_{\mathbf{H}_u}(z) = \rho_\beta\delta(z) + \begin{cases} \frac{1}{\pi}\frac{\rho}{\sqrt{(1-\rho)^2+4\rho|z|^2}} & P_{\mathbf{H}}^{-1}(\beta\rho_\beta + 1 - \beta) < |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{A.92}$$

Thus we left to find the inner boundary, which can easily find by With (4.57) in Corollary 3, we have,

$$P_{\mathbf{H}}^{-1}(r) = \frac{1}{\sqrt{S_{\mathbf{H}^\dagger\mathbf{H}_\beta}(r-1)}} \tag{A.93}$$
$$= \sqrt{(r-1+\rho)(r-1+1)\rho} \tag{A.94}$$

Let, $r \to \beta\rho_\beta + 1 - \beta$, thus the inner boundary of the distribution reads,

$$in(\mathbf{H}_u)^{-1} \triangleq P_{\mathbf{H}}^{-1}(\beta\phi_\beta + 1 - \beta) \tag{A.95}$$
$$= \sqrt{(\beta\rho_\beta\rho - \beta)(\beta\rho_\beta + 1 - \beta)\rho} \tag{A.96}$$

Thus the empirical eigenvalue distribution of $\mathbf{H}_u$ converges to limit distribution as,

$$p_{\mathbf{H}_u}(z) = \rho_\beta\delta(z) + \begin{cases} \frac{1}{\pi}\frac{\rho}{\sqrt{(1-\rho)^2+4\rho|z|^2}} & in(\mathbf{H})^{-1} < |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{A.97}$$

Note the in that in case $\rho = S/R \leq 1$, then $\rho_\beta = 0$. Thus, the inner boundary reads

$$in(\mathbf{H}_u)^{-1} = \sqrt{(\rho-\beta)(1-\beta)\rho} \tag{A.98}$$

as we presented at the corollary for simplicity.

A.7 Proof of Corollary 8

Note that, a matrix whose elements are iid is not R-diagonal matrix. But thanks to CLT (Central Limit Theorem) when we successively product two or more than two independent matrix whose element are iid then the products become iid. Obviously successive product of independent matrix whose element are iid is equivalent to successive product of independent matrix whose element are iid Gaussian.

Recall the statement Theorem 27: the empirical eigenvalue distribution of Product of n-times successive product identical-free R-diagonal matrices and nth power of one of that matrix converge to to same limit.

Since R-diagonal matrices are circularly symmetric distributed, we will first focus on the power of radial probability measure at the final step we switch the complex-valued density as:

$$p_R(r) = 2\pi r \, p_{\mathbf{H}_n}(z)\big|_{|z|=r} = 2r \tag{A.99}$$

Now recall the formula for a transformation of random variable for $y = g(r)$ [43]

$$p_Y(y) = p_R(g^{-1}(y)) \left| \frac{g^{-1}(y)}{dy} \right| \tag{A.100}$$

In our case $g(r) = r^{\alpha}$, then we have,

$$\begin{aligned} p_Y(y) &= \frac{1}{\alpha} p_X \left(y^{\frac{1}{\alpha}} \right) y^{\frac{1}{\alpha}-1} \\ &= \frac{2}{\alpha} y^{\frac{2}{\alpha}-1} \qquad \text{(A.101)} \\ p_{\mathbf{H}_n^{\alpha}}(z) &= \frac{1}{2\pi|z|} \, p_Y(y)\big|_{y=|z|} \\ &= \frac{1}{\pi\alpha} |z|^{\frac{2}{\alpha}-2} \qquad \text{(A.102)} \end{aligned}$$

A.8 Proof of Corollary 4 & 9

Start with Corollary 8, such that,

$$\mathbf{H} = \prod_{n=1}^{\alpha} \mathbf{H}_n \tag{A.103}$$

where the entries of independent matrices $\mathbf{H}_n$ are iid. Then with Theorem 29, we have,

$$p_{\mathbf{H}_u} = \begin{cases} \frac{1}{\pi\beta\alpha} |z|^{\frac{2}{\alpha}-2} & P_{\mathbf{H}}^{-1}(1-\beta) < |z| \le 1 \\ 0 & elsewhere \end{cases} \tag{A.104}$$

Thus we left with the innner boundary which can be easily found with (4.57) in Corollary 3,

$$\begin{aligned} P_{\mathbf{H}}^{-1}(1-\beta) &= \frac{1}{S_{\mathbf{H}\mathbf{H}^{\dagger}}(-\beta)} \qquad \text{(A.105)} \\ &= (1-\beta)^{\frac{\alpha}{2}}. \qquad \text{(A.106)} \end{aligned}$$

Moreover Corollary 4 is same form with $\alpha = 1$ such that

$$p_{\mathbf{H}_u} = \begin{cases} \frac{1}{\pi\beta} & (1-\beta) < |z| \leq 1 \\ 0 & elsewhere \end{cases} \tag{A.107}$$

where $\mathbf{H}_s$ as define in Corollary 4.

A.9 Proof of Corollary 10

Just as in prof of Theorem 1 define, $R_\alpha = rank(\mathbf{X}\mathbf{X}^\dagger) = R/\alpha$. Moreover let define the $R_\alpha \times R_\alpha$ Hermitian matrix $\mathbf{Y}$ whose eigenvalue distribution identical to non-zero eigenvalue distribution of $\mathbf{X}\mathbf{X}^\dagger$ such that,

$$p_{\mathbf{X}\mathbf{X}^\dagger}(x) = (1 - \frac{1}{\alpha})\delta(x) + \frac{1}{\alpha} p_{\mathbf{Y}}(x) \tag{A.108}$$

Now, let start with lemma 1 as,

$$S_{\mathbf{X}\mathbf{X}^\dagger}(z) = \prod_{n=1}^{N} \frac{z + \rho_n}{z + \rho_{n-1}} S_{\mathbf{A}_n^\dagger \mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.109}$$

such that,

$$S_{\mathbf{A}_n^\dagger \mathbf{A}_n}(z) = \frac{1}{z\chi_n + 1} \tag{A.110}$$

Thus we have,

$$\begin{aligned} S_{\mathbf{X}\mathbf{X}^\dagger}(z) &= \prod_{n=1}^{N} \frac{z + \rho_n}{z + \rho_{n-1}} \frac{1}{z\frac{\chi_n}{\rho_{n-1}} + 1} && \text{(A.111)} \\ &= \prod_{n=1}^{N} \frac{z + \rho_n}{z + \rho_{n-1}} \frac{\rho_n}{z + \rho_n} = \prod_{n=1}^{N} \frac{\rho_n}{z + \rho_{n-1}} && \text{(A.112)} \end{aligned}$$

see also (19) in [22]. It is easy to see the rank of $\mathbf{X}\mathbf{X}^\dagger$ is

$$rank(\mathbf{X}\mathbf{X}^\dagger) = \min(K_n) = \min(\rho_n) R \tag{A.113}$$

Thus our main rank parameter reads as

$$\alpha = \frac{1}{\min(\rho_n)}, \qquad 0 \leq n \leq N \tag{A.114}$$

Hence,

$$\begin{aligned} S_{\mathbf{Y}}(z) &= \frac{z+1}{z+\alpha} S_{\mathbf{X}\mathbf{X}^\dagger}\left(\frac{z}{\alpha}\right) && \text{(A.115)} \\ &= \frac{z+1}{z+\alpha} \prod_{n=1}^{N} \frac{\alpha\rho_n}{z + \alpha\rho_{n-1}} && \text{(A.116)} \end{aligned}$$

With Theorem 21 (Corollary 5 in [48]) we have

$$\int_0^\infty \log x dP_{\mathbf{Y}}(x) = -\int_{-1}^0 \log S_{\mathbf{Y}}(z)dz \tag{A.117}$$

$$= -\int_{-1}^0 \log \frac{z+1}{z+\alpha} \sum_{n=1}^N \log \frac{\alpha\rho_n}{z+\alpha\rho_{n-1}} dz \tag{A.118}$$

The integral can be easily simplified by using the function $g(\cdot)$ defined in (A.53), thus we have,

$$\int_0^\infty \log x dP_{\mathbf{Y}}(x) = g(\alpha) - g(1) + \sum_{n=1}^N g(\alpha\rho_{n-1}) - \log \alpha\rho_n \tag{A.119}$$

With the same steps as in the proof of Theorem 2, one can simply get,

$$\frac{I(\gamma)}{R} = \frac{1}{\alpha}\left\{\log\gamma + g(\alpha) - g(1) + \sum_{n=1}^N g(\alpha\rho_{n-1}) - \log\alpha\rho_n - \int_0^{\frac{1}{\gamma}} G_{\mathbf{X}\mathbf{X}^\dagger}(-s) + \frac{\alpha-1}{s} ds\right\} \tag{A.120}$$

such that, in high SNR regime (A.120) reads as

$$\frac{I(\gamma)}{R} \cong \frac{\log\gamma + g(\alpha) - g(1)}{\alpha} + \frac{1}{\alpha}\sum_{n=1}^N g(\alpha\rho_{n-1}) - \log\alpha\rho_n \tag{A.121}$$

A.10 Proof of Lemma 8

By Applying (3.153) recursively, we have

$$\begin{aligned}
S_{\mathbf{X}\mathbf{X}^\dagger}(z) &= S_{\mathbf{A}_N\mathbf{A}_{N-1}\mathbf{A}_{N-2}\cdots\mathbf{A}_2\mathbf{A}_1\mathbf{A}_1^\dagger\mathbf{A}_2^\dagger\cdots\mathbf{A}_{N-2}^\dagger\mathbf{A}_{N-1}^\dagger\mathbf{A}_N^\dagger}(z) \\
&= \frac{z+1}{z+\chi_N} S_{\mathbf{A}_N^\dagger\mathbf{A}_N}\left(\frac{z}{\chi_N}\right) S_{\mathbf{A}_{N-1}\mathbf{A}_{N-2}\cdots\mathbf{A}_2\mathbf{A}_1\mathbf{A}_1^\dagger\mathbf{A}_2^\dagger\cdots\mathbf{A}_{N-2}^\dagger\mathbf{A}_{N-1}^\dagger}\left(\frac{z}{\chi_N}\right) \\
&= \frac{z+1}{z+\chi_N} S_{\mathbf{A}_N^\dagger\mathbf{A}_N}\left(\frac{z}{\chi_N}\right) \frac{z/\chi_N+1}{z/\chi_N+\chi_{N-1}} S_{\mathbf{A}_{N-1}^\dagger\mathbf{A}_{N-1}}\left(\frac{z}{\chi_N\chi_{N-1}}\right) \\
&\quad S_{\mathbf{A}_{N-2}\cdots\mathbf{A}_2\mathbf{A}_1\mathbf{A}_1^\dagger\mathbf{A}_2^\dagger\cdots\mathbf{A}_{N-2}^\dagger}\left(\frac{z}{\chi_N\chi_{N-1}}\right) \\
&= \frac{z+\rho_N}{z+\rho_{N-1}} S_{\mathbf{A}_N^\dagger\mathbf{A}_N}\left(\frac{z}{\rho_{N-1}}\right) \frac{z+\rho_{N-1}}{z+\rho_{N-2}} S_{\mathbf{A}_{N-1}^\dagger\mathbf{A}_{N-1}}\left(\frac{z}{\rho_{N-2}}\right) \\
&\quad S_{\mathbf{A}_{N-2}\cdots\mathbf{A}_2\mathbf{A}_1\mathbf{A}_1^\dagger\mathbf{A}_2^\dagger\cdots\mathbf{A}_{N-2}^\dagger}\left(\frac{z}{\rho_{N-2}}\right) \\
&= \frac{z+\rho_N}{z+\rho_{N-1}} S_{\mathbf{A}_N^\dagger\mathbf{A}_N}\left(\frac{z}{\rho_{N-1}}\right) \frac{z+\rho_{N-1}}{z+\rho_{N-2}} S_{\mathbf{A}_{N-1}^\dagger\mathbf{A}_{N-1}}\left(\frac{z}{\rho_{N-2}}\right) \\
&\quad \frac{z+\rho_{N-2}}{z+\rho_{N-3}} S_{\mathbf{A}_{N-2}^\dagger\mathbf{A}_{N-2}}\left(\frac{z}{\rho_{N-3}}\right) S_{\cdots\mathbf{A}_2\mathbf{A}_1\mathbf{A}_1^\dagger\mathbf{A}_2^\dagger\cdots}\left(\frac{z}{\rho_{N-3}}\right)
\end{aligned} \tag{A.122}$$

This recursive relation shows the proof of Lemma 8. Moreover for completeness let's do final step as

$$S_{\mathbf{XX}^\dagger}(z) = S_{\mathbf{A}_1\mathbf{A}_1^\dagger}\left(\frac{z}{\rho_1}\right)\prod_{n=2}^{N}\frac{z+\rho_n}{z+\rho_{n-1}}S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.123}$$

$$= \frac{z/\rho_1+1}{z/\rho_1+\chi_1}S_{\mathbf{A}_1^\dagger\mathbf{A}_1}\left(\frac{z}{\chi_1\rho_1}\right)\prod_{n=2}^{N}\frac{z+\rho_n}{z+\rho_{n-1}}S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.124}$$

$$= \frac{z+\rho_1}{z+\rho_0}S_{\mathbf{A}_1^\dagger\mathbf{A}_1}\left(\frac{z}{\rho_0}\right)\prod_{n=2}^{N}\frac{z+\rho_n}{z+\rho_{n-1}}S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.125}$$

$$= \prod_{n=1}^{N}\frac{z+\rho_n}{z+\rho_{n-1}}S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.126}$$

$$= \frac{z+\rho_N}{z+\rho_0}\prod_{n=1}^{N}S_{\mathbf{A}_n^\dagger\mathbf{A}_n}\left(\frac{z}{\rho_{n-1}}\right) \tag{A.127}$$

A.11 Proof of Lemma 9

$$G_{\widetilde{\mathbf{X}}}(s) = sG_{\mathbf{XX}^\dagger}\left(s^2\right) \tag{A.128}$$

Proof. Then the Stieltjes transform of $\tilde{p}_\lambda(x)$ reads,

$$G_{\widetilde{\mathbf{X}}}(s) = \int_{-\infty}^{\infty}\frac{p_{\widetilde{\mathbf{X}}}dx}{s-x} \tag{A.129}$$

$$= \frac{1}{2}\int_0^{\infty}\frac{p_{\sqrt{\mathbf{XX}^\dagger}}(x)+p_{\sqrt{\mathbf{XX}^\dagger}}(x)}{s-x}dx$$

$$= \frac{1}{2}\int\left(\frac{1}{s-x}+\frac{1}{s+x}\right)dP_{\sqrt{\mathbf{XX}^\dagger}}(x)$$

$$= s\int\frac{1}{s^2-x^2}dP_{\sqrt{\mathbf{XX}^\dagger}}(x) = s\int\frac{1}{s^2-x}dP_{\mathbf{XX}^\dagger}(x)$$

$$= sG_{\mathbf{XX}^\dagger}\left(s^2\right) \tag{A.130}$$

□

A.12 Proof of Lemma 10

$$S_{\widetilde{\mathbf{X}}}(z) = \left[\frac{z+1}{z}S_{\mathbf{XX}^\dagger}(z)\right]^{\frac{1}{2}} \tag{A.131}$$

Proof. With lemma 9, the M-transform reads the following,

$$M_{\widetilde{\mathbf{X}}}(z) + 1 = \frac{1}{z} G_{\widetilde{\mathbf{X}}} \left(\frac{1}{z} \right) = \frac{1}{z^2} G_{\mathbf{X}\mathbf{X}^\dagger} \left(\frac{1}{z^2} \right) \tag{A.132}$$

Thus,

$$M_{\widetilde{\mathbf{X}}}(z) = M_{\mathbf{X}\mathbf{X}^\dagger}(\omega)\big|_{\omega=z^2} \Rightarrow M^{-1}_{\widetilde{\mathbf{X}}}(z) = \left[M^{-1}_{\mathbf{X}\mathbf{X}^\dagger}(z) \right]^{\frac{1}{2}}. \tag{A.133}$$

Finally the S-transform reads,

$$S_{\widetilde{\mathbf{X}}}(z) = \frac{z+1}{z} M^{-1}_{\widetilde{\mathbf{X}}}(z) = \frac{z+1}{z} \left[M^{-1}_{\mathbf{X}\mathbf{X}^\dagger}(z) \right]^{\frac{1}{2}} \tag{A.134}$$

$$= \left[\frac{z+1}{z} S_{\mathbf{X}\mathbf{X}^\dagger}(z) \right]^{\frac{1}{2}} \tag{A.135}$$

□

Bibliography

[1] Claude Shannon, “A mathematical Theory of Communication”. Univ Illiois Press 1949.

[2] Burt V. Bronk: Exponential Ensemble for Random Matrices, Journal of Mathematical Physics Volume 6, NUMBER 2, FEBRUARY 1965.

[3] Ezio Biglieri, Robert Calderbank, Anthony Constantinides, Andrea Goldsmith, Arogyaswami Paulraj, H. Vincent Poor, “MIMO Wireless Communications” Cambridge University Press 2007.

[4] Emre Telatar, “Capacity of Multi-Antenna Channels, European Transactions on Telecommunications”, 1999.

[5] G. J. Foschini and M. J. Gans, On limits of wireless communication in a fading environment when using multiple antennas, Wireless Personal Commun., vol. 6, no. 3, pp. 311335, Mar. 1998.

[6] Giacinto Gelli, “ Lectures Notes on Wireless Networks” University of Federico II, Naples Italy, 2008.

[7] Wolfram Mathworld, “Narayana Numbers”, http://mathworld.wolfram.com/NarayanaNumber.html.

[8] Carl D. Meyer, ”Matrix Analysis and Applied Linear Algebra”, by SIAM in April 2000.

[9] Romain Couillet and Merouane Debbah, “Random Matrix methods for wireless communications”, Cambridge University Press 2011.

[10] Keith Knight, “Mathematical Statistic”, Chapman & Hall/CRC 2000.

[11] D.Voiculesco, “Multiplication of certain non-commutative random variables”, Journal of Operator Theory 18 (1987), 223-235.

[12] V. L. Girko, Circular Law, Theory. Prob. Appl., vol. 29, pp. 694-706, 1984.

[13] Wigner, E. ”Characteristic Vectors of Bordered Matrices with Infinite Dimensions.” Ann. of Math. 62, 548-564, 1955.

[14] V. A. Marchenko and L.A. Pastur, " Distribution of eigenvalues for some sets of random matrices",Mat. Sb. (N.S.), 72(114):4, 507536, 1967.

[15] Jack W. Silverstein, " Eigenvalues and Eigenvectors of large dimensional sample covariance matrices" in Random Matrices and Their Applications, Joel E. Cohen, Harry Kesten, and Charles M. Newman, Eds., pp. 153159. American Mathematica Society, Providence, RI, 1986.

[16] Borade S. Zheng L, and Gallager R.," Amplify-and-Forward in Wireless Relay Networks: Rate Diversity and Network Size", Transactions on Information Theory, vol. 53, pp. 3302-3318, 2007.

[17] E. Biglieri, G. Taricco, and A. Tulino, Performance of spacetime codes for a large number of antennas, IEEE Transactions on Information Theory, vol. 48, no. 7, pp. 17941803, Jul. 2002.

[18] Ralf R. Mūller,"On the Asymptotic Eigenvalue Distribution of Concatenated Vector-Valued Fading Channels", IEEE Transactions on Information Theory, vol. 48, no. 7, pp. 2086-2091, Jul 2002.

[19] Ralf R. Mūller, "A Random Matrix Model of Communication via Antenna Arrays", IEEE Transactions on Information Theory, vol. 48, no. 9, pp.2495-2506, Sep 2002.

[20] Ralf R. Mūller, "Applications of Large Random Matrices in Communications Engineering" Invited for International Conference on Advances in the Internet, Processing, Systems, and Interdisciplinary Research (IPSI), Sveti Stefan, Montenegro, Oct 2003.

[21] Ralf R. Müller, A random matrix model for communication via antenna arrays, IEEE Transactions on Information Theory, vol. 48, no. 9, pp. 24952506, Sep. 2002.

[22] Ralf R. Müller, On the asymptotic eigenvalue distribution of concatenated vectorvalued fading channels, IEEE Transactions on Information Theory, vol. 48, no. 7, pp. 20862091, Jul. 2002.

[23] Ralf R. Müller, Burak Cakmak, "Channel Modelling of MU-MIMO Systems by Quaternionic Free Probability" IEEE International Symposium on Information Theory (ISIT), Massachusetts- Cambridge, USA, July 2012.

[24] Ralf R. Müller, D. Guo, A. Moustakas, "Vector precoding for wireless MIMO systems and its replica analysis",IEEE Journal on Selected Areas in Communications, vol. 26, no. 3, pp. 530-540, Apr 2008.

[25] R. R. Mūller, Random matrices, free probability and the replica method, in European Signal Processing Conference, Vienna, Austria,Sep. 2004.

[26] Antonia M. Tulino, Sergio Verdu, Random Matrix Theory and Wireless Communications, Foundations and Trends in Communications and Information Theory, vol. 1, no. 1, Jun. 2004.

[27] R. Couillet and M. Debbah, Random Matrix Methods for Wireless Communications. Cambridge University Press, 2011.

[28] S. Verdu and S. Shamai, Spectral efficiency of CDMA with random spreading, IEEE Trans. on Information Theory, vol. 45, pp. 622640, Mar. 1999.

[29] S. Verdu,"Multiuser Detection", Cambridge, UK: Cambridge University Press 1998.

[30] A. M. Tulino, A. Lozano, and S. Verdu, Impact of antenna correlation on the capacity of multiantenna channels, IEEE Transactions on Information Theory, vol. 51, no. 7, pp. 24912509, Jul. 2005.

[31] F. Hiai and D. Petz, "The Semicircle Law, Free Random Variables and Entropy", American Mathematical Society, 2000.

[32] F. Hiai and D. Petz, "The Semicircle Law, Free Random Variables and Entropy", American Mathematical Society, 2006.

[33] F. Hiai and D. Petz, " Asymptotic freeness almost everywhere for random matrices" Acta Sci. Math. Szeged, vol. 66, pp. 801-826,2000.

[34] T. J. Stieltjes, Recherches sur les fractions continues, Annales de la Faculte des Sciences de Toulouse, vol. 8, pp. 1122, 1894.

[35] Jack W. Silverstein," The Stieltjes Transform and its Role in Eigenvalues Behaviour of Large Dimensional Random Matrices" RANDOM MATRIX THEORY AND ITS APPLICATIONS - Multivariate Statistics and Wireless Communications, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore - Vol. 18.

[36] Jack W. Silverstein, Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices, Journal of Multivariate Analysis, vol. 55, pp. 331339, 1995.

[37] Jack W. Silverstein and Z.D. Bai, On the empirical distribution of eigenvalues of a class of large dimensional random matrices, Journal of Multivariate Analysis, vol. 54, pp. 175192, 1995.

[38] S. Thorbjϕrnsen, Mixed moments of Voiculescus Gaussian random matrices, Journal of Functional Analysis, vol. 176, pp.213246, 2000.

[39] D.Voiculesco," Multiplication of certain non-commutative random variables", Journal of Operator Theory **18** (1987), 223-235.

[40] D. Voiculescu, Limit laws for random matrices and free products, Inventiones Mathematicae, vol. 104, pp. 201220, 1991.

[41] Roland. Speicher, "Free probability theory and non-crossing partitions",39e Séminaire Lotharingien de Combinatoire, 1997.

[42] N. Raj Rao, Roland Speicher "Multiplication of free random variables and the S-transform:The case of vanishing mean', Electronic Communications in Probability, July 2007.

[43] Alexandru Nica and Roland Speicher, "Lectures on the Combinatorics of Free Probability", London Mathematical Society Lecture Note Series 365, 2006.

[44] Alexandru Nica and Roland Speicher, "On the multiplication of free n-tuples of non-commutative random variables", Functional Analysis (math.FA); Operator Algebras- Quantum Algebra, arXiv:funct-an/9604011v1 30 Apr 1996.

[45] Franz Lehner," Cumulants in non-commutative probability theory I. Non-commutative Exchangeability systems ", arXiv:math/0210442v3[math.CO] 2 September 2004.

[46] Uffe Haagerup and Hanne Schultz, "Brown Measures of Unbounded Operators Affilated with a Finite von Neumann Algebra", arXiv:math/06052251v1 [math.OA], May 2006.

[47] Uffe Haagerup, Flemming Larsen, "Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras", Journal of Functional Analysis **176**.331-367 (2000) November 30,1999.

[48] Vladislav Kargin," Lyapunov Exponent of Free Operators", arxiv:0712.1278v2 [math.PR] 28 Nov 2008.

[49] Fumio Hiai and Dènes Petz, "The Semicircle Law, Free Random Variables and Entropy", American Mathematical Society Providence, RI 2000.

[50] emphAlan Edelman, N. Raj Rao Random Matrix Theory, Cambridge University Press 2005

[51] Kenneth J. Dykema, On the S-transform over Banach Algebra, arXiv:math/0501083v2 [math.OA] 11 Jan 2005.

[52] A. Jarosz and M. A. Nowak, Random Hermitian versus random non- Hermitian operatorsunexpected links, Journal of Physics A: Mathematical and General, vol. 39, pp. 10 10710 122, 2006.

[53] R. Remmert, "Theory of Complex Functions", New York: SpringerVerlag, 1991.

[54] Z. Burda, R. A. Janik, and M. A. Nowak, Multiplication law and S-transform for non-hermitian random matrices, arXiv:1104.2452v1, Apr.2011.

[55] R. A. Janik, W. Norenberg, M. A. Nowak, G. Papp, I. Zahed, "Correlations of Eigenvectors for Non-Hermitian Random-Matrix Models" arXiv:cond-mat/9902314v1 23 Feb 1999.

[56] Philippe Bianne, Franz Lehner, " Computation of some Examples of Brown's spectral Measure in Free Probability", ESI- Institute for Mathematical Physics , January 7, 2000.

[57] Gabriel H. Tucci," Limits laws for geometric means of free random variables", arXiv:0802.4226v2, October 2010.

[58] Christoly Biely and Stefan Thurner,"Random matrix ensembles of time-lagged correlation matrices": Derivation of eigenvalue spectra and analysis of financial time-series. 19 September 2007.

[59] Thomas M. Cover and Joy A. Thomas, "Elements of Information Theory", John Wiley & Sons, New York, 1991.

[60] S. Verdu,"Multiuser Detection", Cambridge, UK: Cambridge University Press 1998.

[61] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," IEEE Trans. on information Theory, vol. 45, pp. 622-640, March 1999.

[62] Ezio Bigleri, "Coding for Wireless Channels", Springer 2005.

[63] J. T. Chalker and B. Mehlig, "Eigenvector statistics in non-Hermitian random matrix ensembles", arXiv:cond-mat/9809090v1 [cond-mat.dis-nn] 4 September 1998.

[64] D. E. Knuth, The Art of Computer Programming, 2nd ed. Reading, MA: Addison-Wesley, 1981, vol. 2.