

# Perturbed Hankel Determinants: Applications to the Information Theory of MIMO Wireless Communications

Yang Chen  
Department of Mathematics  
Imperial College  
180 Queen's Gate  
London SW7 2BZ, UK  
ychen@ic.ac.uk

Matthew McKay  
Electronic and Computer Engineering Department  
Hong Kong University of Science and Technology  
Clear Water Bay, Kowloon, HONG KONG  
eemckay.ust.hk

01-07-2010

## Abstract

In this paper we compute two important information-theoretic quantities which arise in the application of multiple-input multiple-output (MIMO) antenna wireless communication systems: the distribution of the mutual information of multi-antenna Gaussian channels, and the Gallager random coding upper bound on the error probability achievable by finite-length channel codes. It turns out that the mathematical problem underpinning both quantities is the computation of certain Hankel determinants generated by deformed versions of classical weight functions. For single-user MIMO systems, it is a deformed Laguerre weight, whereas for multi-user MIMO systems it is a deformed Jacobi weight. We apply two different methods to characterize each of these Hankel determinants. First, we employ the ladder operators of the corresponding monic orthogonal polynomials to give an exact characterization of the Hankel determinants in terms of Painlevé differential equations. This turns out to be a Painlevé V for the single-user MIMO scenario and a Painlevé VI for the multi user scenario. We then employ Coulomb fluid methods to derive new closed-form approximations for the Hankel determinants which, although formally valid for large matrix dimensions, are shown to give accurate results for both the MIMO mutual information distribution and the error exponent even when the matrix dimensions are small. Focusing on the single-user mutual information distribution, we then employ both the exact Painlevé representation and the Coulomb fluid approximation to yield deeper insights into the scaling behavior in terms of the number of antennas and signal-to-noise ratio. Among other things, these results allow us to study the asymptotic Gaussianity of the distribution as the number of antennas increase, and to explicitly compute the correction terms to the mean, variance, and higher order cumulants.

# 1 Introduction and Preliminaries

Over the past decade, multiple-input multiple-output (MIMO) systems have been at the forefront of wireless communications research and development. Such systems, which employ multiple antennas at both the transmitter and receiver, have revolutionized the wireless industry and now form the basis of most emerging wireless standards, such as next-generation wireless local area networks (WLAN) and cellular mobile networks. The main reason for this explosion of interest is due to the independent discoveries of Telatar [1] and Foschini [2], which demonstrated that the fundamental information-theoretic capacity of MIMO systems grows *linearly* with the number of antennas. This is important, since traditional methods for increasing capacity, which typically increase valuable system resources such as the transmission power, yield only a logarithmic capacity increase. Indeed, MIMO is now widely recognized as a key technology for meeting the ever-increasing demands for higher-rate data-oriented wireless communications applications and services.

There are various metrics for characterizing the fundamental information-theoretic limits of MIMO systems, each of which have relevance depending on the specific wireless communication configuration. The most commonly studied metric is the so-called *ergodic capacity*, which specifies the maximum achievable average mutual information between the transmitter and receiver. This measure, which assumes that there is sufficient dynamics in the system such that a user's codeword may span a large number of "independent channels", has been studied extensively over the past decade (see e.g., [1–10] and the tutorial discussion [11]). A closely related measure which has received far less attention is the *outage capacity*. This measure is important for characterizing the communication limits of systems which are not highly dynamic (for example, WLANs), or systems which have stringent delay requirements. Compared with the ergodic capacity, the analysis of outage capacity is much more complicated since it requires the *distribution* of the mutual information between the transmitter and receiver, rather than simply the average value. Another important information-theoretic quantity of practical interest is the error probability achievable with block-coding schemes of a given length and rate. Whilst an

exact characterization of this quantity is not tractable in general, methods have been proposed by Gallager for upper bounding this quantity [12]. Such bounds have been well-established for single-antenna systems, however much less is known for MIMO.

In this paper, we aim to present new methods for studying the outage capacity and the error probability of MIMO systems. As we will see, in both cases the fundamental mathematical problem of interest boils down to characterizing the Hankel determinant

$$D_n = \det (\mu_{i+j})_{i,j=0}^{n-1} \tag{1.1}$$

generated from the moments of a certain weight function  $w(x)$ ,

$$\mu_k := \int_J x^k w(x) dx, \quad k = 0, 1, 2, \dots \tag{1.2}$$

with  $J$  denoting the support of the weight. The specific form of the weight depends on the MIMO configuration under investigation. We will consider two important MIMO configurations pertaining to single-user and multi-user MIMO systems respectively. In the single-user case, the weight function is shown to be a deformation of the classical Laguerre weight given by

$$w(x) = x^\alpha e^{-x} (x+t)^\lambda, \quad 0 \leq x < \infty, \alpha > 0, t > 0, \tag{1.3}$$

whereas in the multi-user case the weight function is found to be a deformation of the (shifted) classical Jacobi weight given by

$$w(x) = x^{\alpha_1} (1-x)^{\alpha_2} \left( \frac{x+t}{1-x} \right)^\lambda, \quad x \in [0, 1], \alpha_1 > 0, \alpha_2 > 0, t > 0. \tag{1.4}$$

We employ two different methods from random matrix theory for evaluating the corresponding Hankel determinants, both of which have been used to great effect by the mathematical physics community. First, we derive exact expressions for the Hankel determinants by employing the theory of orthogonal polynomials and their corresponding ladder operators. There exists extensive literature on this subject; for example [13–20]. See also [21–26] for recent applications

of ladder operators to Hermitian matrix ensembles.

Despite these prior contributions, this approach has not been employed to problems in wireless communications and information theory. Through the ladder operator framework, we demonstrate that the Hankel determinant generated from the deformed Laguerre weight (1.3) has a simple representation involving the Painlevé V differential equation. We also derive a discrete difference equation representation, which allows the Hankel determinant to be numerically evaluated iteratively, in terms of the Hankel matrix dimension. For the deformed Jacobi weight (1.4), upon establishing the connection to the multi-user MIMO system model of interest, we employ a known result derived using the ladder operator method in [27] to represent the corresponding Hankel determinant in terms of the Painlevé VI differential equation. This presents a new connection between the Painlevé VI differential equation and the information theory of multi-user MIMO communication systems.

In addition to deriving exact expressions for the Hankel determinants, we also derive approximations for these determinants by employing the general linear statistics results from [28], derived based on Dyson's Coulomb fluid interpretation [29–31]. These asymptotic results are essentially the Hankel analog of Szegő's strong limit theorem on the asymptotic characterization of large Toeplitz determinants, a component of which appeared as early as 1919 [32], where Hankel determinants generated by compactly supported weights were studied. We refer the reader to [33–35] for more related contributions.

As with the ladder operator approach, Coulomb fluid techniques have been used very successfully in the context of mathematical physics, however, they have yet to be adopted by the wireless communications and information theory community. (An exception is the very recent submitted work [36], which employed the Coulomb fluid framework but not the linear statistics method to characterize the large-antenna behavior of the single-user MIMO channel capacity.) A key advantage of the Coulomb fluid methodology is that the expressions which are obtained are in closed-form, and are simpler than those obtained via exact methods. Moreover, although formally valid for large matrix dimensions, we find that the approximations are remarkably accurate for even very small matrix dimensions (e.g.,  $2 \times 2$ ). By adopting this Coulomb fluid

framework, our results also establish the Gaussian behavior of the channel capacity as the number of antennas grow large, for both the single-user and multi-user MIMO systems. This point has been derived previously for the single-user MIMO case, using different methods from [37], however we do not believe that it has been established for the multi-user MIMO scenario.

After deriving the exact Painlevé and approximate Coulomb fluid representations, we then employ both results to gain useful insights into the mutual information distribution. In particular, focusing on the single-user MIMO scenario (i.e., deformed Laguerre weight) with  $n$  transmit and receive antennas, we demonstrate that the Coulomb fluid approximation in fact gives the *exact* distribution of the mutual information, corresponding to a Gaussian, to leading order in  $n$ . We also use the Painlevé V equation to compute the large- $n$  correction terms for the mean, variance, and third cumulant. Among other things, these results demonstrate the interesting phenomenon that the distribution deviates from Gaussian as the signal-to-noise ratio (SNR) increases. The sensitivities of the mean, variance, and third moment, with respect to the SNR are also examined.

The paper is organized as follows. In the rest of this section, we present a detailed discussion of the linear models which underpin the single-user and multi-user MIMO wireless communication scenarios of interest. We also introduce the fundamental information-theoretic measures of outage capacity and error probability, and establish important connections with Hankel determinants generated from deformed Laguerre and Jacobi weights. Then, in Section 2, we introduce the ladder operators of orthogonal polynomials and their associated compatibility conditions, which provide the key ingredients for establishing an exact finite  $n$  characterization of the Hankel determinants in terms of Painlevé differential equations. In Section 3, we introduce, for large  $n$ , the probability density of a class of random variables called *linear statistics*. These results, based on Dyson's Coulomb fluid interpretation [29] and developed further in [28, 30, 31, 38], are very general and embrace a wide class of random matrix models. By employing these general results, we derive closed-form approximations for the mutual information distribution for the Hankel determinants generated by both the deformed Laguerre and Jacobi weights. These results permit very fast evaluation of the error probabilities of single-user and multi-user MIMO

systems, whilst also demonstrating the Gaussian behavior of the mutual information distribution for large  $n$ . In Section 4, focusing on the single-user MIMO scenario, the Coulomb results are compared, for large  $n$ , with the solutions of the continuous  $\sigma$ -form of the Painlevé V. We also compute closed-form asymptotic expressions for the recurrence coefficients corresponding to the deformed Laguerre weight, which are basic variables in our theory. Our analysis will involve a number of complicated integrals, which are derived and tabulated in the Appendix.

## 1.1 Information Theory of MIMO Wireless Systems

Consider a MIMO communication system with  $n_t$  transmit and  $n_r$  receive antennas. The linear model relating the input (transmitted) signal vector  $\mathbf{x}_{n_t \times 1} \in \mathbb{C}^{n_t}$  and output (received) signal vector  $\mathbf{y}_{n_r \times 1} \in \mathbb{C}^{n_r}$  takes the form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} . \tag{1.5}$$

Here,  $\mathbf{n}_{n_r \times 1} \in \mathbb{C}^{n_r}$  is a complex Gaussian vector with zero mean and covariance  $E(\mathbf{n}\mathbf{n}^\dagger) = \mathbf{Q}_n$ . This covariance matrix can account for the effects of both receiver noise as well as multi-user interference, and as such, the selection of  $\mathbf{Q}_n$  will distinguish between the single-user and multi-user MIMO models which we consider subsequently. The matrix  $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ , referred to as the *channel matrix*, represents the wireless fading coefficients between each transmit and receive antenna. This matrix is assumed to be known to the receiver<sup>1</sup>, but not to the transmitter. The channel is modeled stochastically, with distribution depending on the specific wireless environment. Under the realistic assumption that there are sufficient scatterers surrounding the transmit and receive terminals, the channel matrix  $\mathbf{H}$  is accurately modeled according to a complex Gaussian distribution with independent and identically distributed (i.i.d.) elements having zero mean and unit variance. The transmitted signal  $\mathbf{x}$  is designed to meet a power constraint:

$$E(\mathbf{x}^\dagger \mathbf{x}) \leq P . \tag{1.6}$$

---

<sup>1</sup>In practice, this information can be obtained using standard channel estimation techniques.

### 1.1.1 Ergodic and Outage Capacity

The Shannon capacity governs the ultimate limits of communication systems. More specifically, this measure defines the highest data rate that can be achieved with negligible errors by any transmission scheme. As such, the Shannon capacity forms a benchmark for the design of practical transmission technologies. Mathematically, the Shannon capacity is defined in terms of the mutual information between the input and output signals, which for the MIMO linear model (1.5) is given by:

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}|\mathbf{H}) &= \mathcal{H}(\mathbf{y}|\mathbf{H}) - \mathcal{H}(\mathbf{y}|\mathbf{x}, \mathbf{H}) \\ &= \mathcal{H}(\mathbf{y}|\mathbf{H}) - \mathcal{H}(\mathbf{n}) \end{aligned} \tag{1.7}$$

with  $\mathcal{H}(\mathbf{y}|\mathbf{H})$  denoting the conditional entropy of  $\mathbf{y}$ , defined in terms of its density  $p(\mathbf{y}|\mathbf{H})$  as follows:

$$\mathcal{H}(\mathbf{y}|\mathbf{H}) = E(-\log p) := - \int_{\mathbb{C}^{n_r}} p(\mathbf{y}|\mathbf{H}) \log p(\mathbf{y}|\mathbf{H}) d\mathbf{y}. \tag{1.8}$$

In general, there are two important measures of capacity—the *ergodic capacity* and the *outage capacity*. The ergodic capacity is the relevant measure for applications with highly dynamic channels; for example, high-mobility wireless applications, in which case the channel matrix  $\mathbf{H}$  varies quickly over time, and therefore each transmission codeword sees a large number of “independent” channel realizations. The ergodic capacity is defined as:

$$C = \max_{p(\mathbf{x})} E_{\mathbf{H}} (I(\mathbf{x}; \mathbf{y}|\mathbf{H})) \tag{1.9}$$

where the maximum is taken over all densities  $p(\mathbf{x})$  of the input vector  $\mathbf{x}$ , subject to the constraint (1.6). It has been proven in [1] that the optimal input density  $p^*(\mathbf{x})$  is multi-variate complex Gaussian with zero mean. As such, the mutual information  $I(\mathbf{x}; \mathbf{y}|\mathbf{H})$  becomes

$$I(\mathbf{x}; \mathbf{y}|\mathbf{H}) = \log \det \left( \mathbf{I}_{n_r} + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger \mathbf{Q}_n^{-1} \right) \tag{1.10}$$

where  $\mathbf{Q} = E(\mathbf{x}\mathbf{x}^\dagger)$  is the input signal covariance. The capacity (1.9) can therefore be reposed as an optimization over  $\mathbf{Q}$  as

$$C = \max_{\mathbf{Q} \geq 0} E_{\mathbf{H}} (I(\mathbf{x}; \mathbf{y} | \mathbf{H})) \quad (1.11)$$

subject to  $\text{tr}(\mathbf{Q}) \leq P$ . This quantity has been studied extensively over the past decade. Indeed, methods have been proposed to calculate the optimal  $\mathbf{Q}$  for a range of MIMO channel models of interest [39, 40], and the expectation has been characterized through the use of both large- and finite-dimensional random matrix theory [1–6, 9, 11, 41]. For the case which we consider, where  $\mathbf{H}$  is assumed to be complex Gaussian with independent zero-mean unit-variance entries, the capacity-achieving input covariance has been derived in [1] and is given by

$$\mathbf{Q}^* = \frac{P}{n_t} \mathbf{I}_{n_t}. \quad (1.12)$$

In practical terms, this implies that the capacity is achieved by sending independent Gaussian signals from each of the transmit antennas with equal power.

Compared with the ergodic capacity, the outage capacity is a much more difficult problem and has received far less attention. In contrast to the ergodic capacity, this measure is suitable for wireless applications with low mobility (e.g., wireless local area networks), in which case the channel is assumed fixed during the transmission of a codeword. The outage capacity corresponding to an outage probability  $P_{\text{out}}$  is defined as the transmission rate which can be supported by  $(1 - P_{\text{out}}) \times 100\%$  of the channel realizations<sup>2</sup>. Although the optimal distribution of  $\mathbf{x}$  which maximizes this quantity is unknown, a sensible choice is to employ the same distribution as that which achieves the ergodic capacity, i.e.,  $p^*(\mathbf{x})$ . In this case, the outage capacity  $C_{\text{out}}$  satisfies:

$$P_{\text{out}}(C_{\text{out}}) = \Pr (I(\mathbf{x}; \mathbf{y}) < C_{\text{out}})$$

---

<sup>2</sup>By *supported*, we mean that the mutual information for a given channel realization is greater than the transmission rate.



$$= \Pr \left( \log \det \left( \mathbf{I}_{n_r} + \frac{P}{n_t} \mathbf{H} \mathbf{H}^\dagger \mathbf{Q}_n^{-1} \right) < C_{\text{out}} \right) \quad (1.13)$$

with  $\mathbf{Q}^*$  denoting the input covariance which maximizes (1.11). Compared with (1.11), the quantity on the right-hand side (r.h.s.) of (1.13) is much more difficult to characterize since it involves the entire distribution of the  $\log \det(\cdot)$  random variable, rather than simply the expected value. As such, a common approach has been to compute the first few moments of the distribution, and then use these to obtain a Gaussian approximation (see e.g., [4, 6]).

The outage probability (1.13) can be calculated via

$$P_{\text{out}}(C_{\text{out}}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{M}(j\omega) \frac{1 - e^{-j\omega C_{\text{out}}}}{j\omega} d\omega \quad (1.14)$$

where  $\mathcal{M}(\cdot)$  denotes the moment generating function of the mutual information, taking the form

$$\begin{aligned} \mathcal{M}(\lambda) &:= E_{\mathbf{H}} (\exp(\lambda I(\mathbf{x}; \mathbf{y} | \mathbf{H}))) \\ &= E_{\mathbf{H}} \left( \det \left( \mathbf{I}_{n_r} + \frac{P}{n_t} \mathbf{H} \mathbf{H}^\dagger \mathbf{Q}_n^{-1} \right)^\lambda \right), \end{aligned} \quad (1.15)$$

and  $j := \sqrt{-1}$ .

Alternatively, the Chernoff bound may be employed, as in [9], to give

$$P_{\text{out}}(C_{\text{out}}) \leq \mathcal{M}(-s) e^{s C_{\text{out}}} \quad (1.16)$$

for any  $s > 0$ . Let

$$G(s) := \log \mathcal{M}(-s),$$

and rewrite the moment generating function in an obvious alternative form:

$$\mathcal{M}(-s) = \int e^{-sq(X)} \mu(dX).$$

Then a straightforward computation shows that

$$G'(s) = -\frac{\int e^{-sq(X)} q(X) \mu(dX)}{\int e^{-sq(X)} \mu(dX)}$$

$$G''(s) = \frac{\int e^{-sq(X)} [q(X)]^2 \mu(dX)}{\int e^{-sq(X)} \mu(dX)} - \left( \frac{\int e^{-sq(X)} q(X) \mu(dX)}{\int e^{-sq(X)} \mu(dX)} \right)^2.$$

If we interpret

$$e^{-sq(X)} \mu(dX)$$

as a probability measure, then for all real  $s$ ,

$$G''(s) = \overline{(q - \bar{q})^2} > 0,$$

where

$$\overline{F(X)} := \frac{\int F(X) e^{-sq(X)} \mu(dX)}{\int e^{-sq(X)} \mu(dX)}.$$

Since  $G(s)$  is convex we can minimize the Chernoff bound by minimizing

$$G(s) + s C_{\text{out}}$$

and minimum of the above is

$$-\frac{1}{2} \frac{(s^* G''(s^*) - C_{\text{out}})^2}{G''(s^*)} + \frac{1}{2} s^{*2} G''(s^*),$$

where  $G(s)$  is minimized at  $s^*$  and we have assumed that  $s^* > 0$ .

### 1.1.2 Upper Bound on Error Probability

Whilst the capacity is an important fundamental quantity, it is difficult to approach in practice since it requires infinitely long codewords and receivers with unbounded complexity. Thus, an important question is to determine what data rates  $R$  are achievable for practical channel coding strategies with fixed and finite codeword length  $N$ , subject to a requirement on the tolerated

error probability  $P_e$ . One method for addressing this problem was proposed by Gallager [12], where a general upper bound on  $P_e$  was derived, assuming that *randomly-selected* length- $N$  block codes of rate  $R$  were employed, along with maximum-likelihood receivers. For the MIMO transmission model (1.5), assuming that the channel is *memoryless* (i.e., each transmission period sees an independent realization of  $\mathbf{H}$ ), the Gallager random coding upper bound on error probability is expressed as [1]

$$P_e(N, R) \leq e^{-NE_r(R)}, \quad (1.17)$$

where  $E_r(R)$  is referred to as the *error exponent*, which is independent of  $N$ . This function is given by

$$E_r(R) = \max_{0 \leq \rho \leq 1} \{E_0(\rho) - \rho R\}, \quad (1.18)$$

with

$$E_0(\rho) = -\log E_{\mathbf{H}} \left[ \int_{\mathbb{C}^{n_r}} \left[ \int_{\mathbb{C}^{n_t}} p(\mathbf{x}) \{p(\mathbf{y}|\mathbf{x}, \mathbf{H})\}^{1/(1+\rho)} d\mathbf{x} \right]^{(1+\rho)} d\mathbf{y} \right], \quad (1.19)$$

where

$$p(\mathbf{y}|\mathbf{x}, \mathbf{H}) = \frac{\exp[-(\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger \mathbf{Q}_n^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x})]}{\det(\pi \mathbf{Q}_n)} \quad (1.20)$$

and  $p(\mathbf{x})$  denotes the density of the input signal  $\mathbf{x}$ . To maximize the error exponent and thus minimize the error probability,  $p(\mathbf{x})$  should be selected so as to maximize  $E_0(\rho)$ . Evaluating this optimal input distribution is very challenging, and a sensible (and more tractable) choice is to assume that  $\mathbf{x}$  takes the capacity-achieving distribution  $p^*(\mathbf{x})$  presented in the previous section; i.e., it is zero-mean complex Gaussian with covariance  $\mathbf{Q}^*$ . In fact, it was shown in [1] that under the assumption that  $\mathbf{x}$  is Gaussian,  $p^*(\mathbf{x})$  is optimum in terms of maximizing  $E_0(\rho)$ .

With this input distribution, it is easy to see that (1.19) particularizes to:

$$E_0(\rho) = -\log E_{\mathbf{H}} \left[ \det \left( \mathbf{I}_{n_r} + \frac{P}{n_t(1+\rho)} \mathbf{H} \mathbf{H}^\dagger \mathbf{Q}_n^{-1} \right)^{-\rho} \right]. \quad (1.21)$$

We remark that whilst more refined bounds on the error probability compared with (1.17) have also been derived [12], these more elaborate bounds still yield the same underlying mathematical problem as that posed in (1.21).

## 1.2 Single-User MIMO and the Deformed Laguerre Weight

Single-user MIMO systems embrace a class of coordinated wireless networks for which all transmissions are scheduled in an orthogonal manner (e.g., in orthogonal time-slots or orthogonal frequency bands), and as such, transmissions do not interfere with one another. In this scenario, the transmitted and received signals conform to the linear model (1.5), with  $\mathbf{n}$  simply reflecting the receiver noise. This noise is spatially uncorrelated (across antennas), and without loss of generality it has covariance

$$\mathbf{Q}_n = \mathbf{I}_{n_r}. \quad (1.22)$$

Due to the normalization of the trace of  $\mathbf{Q}_n$ , the transmit power  $P$  also represents the SNR.

The key quantity of interest for the outage capacity is the moment generating function, which in this case particularizes to

$$\mathcal{M}(\lambda) = E_{\mathbf{H}} \left[ \det \left( \mathbf{I}_{n_r} + \frac{1}{t} \mathbf{H} \mathbf{H}^\dagger \right)^\lambda \right] \quad (1.23)$$

with

$$t := \frac{n_t}{P}. \quad (1.24)$$

Similarly, the error exponent (1.19) admits the same form, but with substitutions  $\lambda = -\rho$  and  $t = (1+\rho)n_t/P$ . In the following, we will focus our discussion on the moment generating function

(1.23), keeping in mind that the application to the error probability is immediate.

Let

$$m := \max\{n_r, n_t\}, \quad n := \min\{n_r, n_t\}, \quad \alpha := m - n$$

and define

$$\mathbf{W} := \begin{cases} \mathbf{H}\mathbf{H}^\dagger, & n_r < n_t \\ \mathbf{H}^\dagger\mathbf{H}, & n_r \geq n_t \end{cases}.$$

The matrix  $\mathbf{W}$  is a complex Wishart random matrix with positive eigenvalues denoted by  $\{x_i\}_{i=1}^n$ . It is well known that the joint probability density function of the eigenvalues read

$$p(x_1, x_2, \dots, x_n) \propto \prod_{i=1}^n w_{\text{Lag}}(x_i) \prod_{1 \leq j < k \leq n} (x_j - x_k)^2, \quad (1.25)$$

where  $x_i \in [0, \infty)$ , and  $w_{\text{Lag}}(\cdot)$  is the classical Laguerre weight

$$w_{\text{Lag}}(x) = x^\alpha e^{-x}.$$

With these definitions, using the identity

$$\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A}), \quad (1.26)$$

we can evaluate the moment generating function as

$$\begin{aligned} \mathcal{M}(\lambda) &= E \left[ \det \left( \mathbf{I}_n + \frac{1}{t} \mathbf{W} \right)^\lambda \right] \\ &= E \left[ \prod_{k=1}^n \left( 1 + \frac{x_k}{t} \right)^\lambda \right] \\ &= \frac{\int_{\mathbb{R}_+^n} \prod_{i < j} (x_i - x_j)^2 \prod_{k=1}^n \left( 1 + \frac{x_k}{t} \right)^\lambda w_{\text{Lag}}(x_k) dx_k}{\int_{\mathbb{R}_+^n} \prod_{i < j} (x_i - x_j)^2 \prod_{k=1}^n w_{\text{Lag}}(x_k) dx_k}. \end{aligned} \quad (1.27)$$

The remaining integrals are evaluated in determinant form via the Andreief-Heine identity:

$$\det(\mu_{i+j})_{i,j=0}^{n-1} = D_n[w] = \frac{1}{n!} \int_{(a,b)^n} \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \prod_{k=1}^n w(x_k) dx_k, \quad (1.28)$$

where

$$\mu_i := \int_a^b x^i w(x) dx, \quad i = 0, 1, 2, \dots \quad (1.29)$$

are moments of the weight  $w$ . See [42] for a recent exposition on this and other related matters.

Obviously the moments would depend on the parameters which may appear in the weight.

With this identity, we immediately obtain

$$\mathcal{M}(\lambda) = t^{-n\lambda} \frac{D_n(t, \lambda)}{D_n(t, 0)} \quad (1.30)$$

where

$$D_n(t, \lambda) = \det(\mu_{i+j}(t, \lambda))_{i,j=0}^{n-1} \quad (1.31)$$

is the Hankel determinant generated from the deformed Laguerre weight

$$w_{\text{dLag}}(x) = w_{\text{dLag}}(x, t, \lambda) := (x+t)^\lambda w_{\text{Lag}}(x), \quad t > 0 \quad (1.32)$$

with moments

$$\mu_k(t, \lambda) := \int_0^\infty x^k w_{\text{dLag}}(x) dx, \quad k = 0, 1, 2, \dots \quad (1.33)$$

**Remark 1** The factor  $D_n(t, 0)$  is simply the Hankel determinant generated from the non-deformed Laguerre weight,  $w_{\text{Lag}}(x)$ , which can be computed exactly in terms of the Barnes  $G$ -

function as

$$D_n(t, 0) = D_{n,\alpha}[w_{\text{Lag}}] = \frac{G(n+1)G(n+\alpha+1)}{G(\alpha+1)}, \quad G(1) = 1. \quad (1.34)$$

Similarly, for  $t = 0$  and fixed  $\lambda$ , it follows that

$$D_n(0, \lambda) = D_{n,\alpha+\lambda}[w_{\text{Lag}}].$$

**Remark 2** The moments are expressed in terms of the Kummer function of the second kind  $U(a; b; z)$  as follows:

$$\begin{aligned} \mu_k(t, \lambda) &:= \int_0^\infty x^{\alpha+k}(x+t)^\lambda e^{-x} dx \\ &= t^{\alpha+\lambda+k+1} \Gamma(\alpha+k+1) U(\alpha+k+1; \alpha+\lambda+k+2; t), \quad k = 0, 1, 2, \dots \end{aligned} \quad (1.35)$$

Note that this expression has been previously reported in [43], and an alternative representation given in [44]. Whilst this identity, combined with (1.34) and (1.30), gives a “closed-form” determinantal representation for the moment generating function, it does not provide useful insights and it also becomes unwieldy to evaluate if the number of antennas become large. Moreover, in this form, it is not amenable to further analysis, such as determining the asymptotic scaling properties. To overcome these problems, in Section 2 we employ the theory of orthogonal polynomials and their corresponding ladder operators to provide a more useful characterization, where we express the Hankel determinant generated from the deformed Laguerre weight in terms of the classical Painlevé V differential equation.

An alternative characterization for the moment generating function which will also prove useful is derived as follows. Starting with (1.27), and applying the transformations  $x_i \rightarrow nx_i$ ,  $i = 1, \dots, n$ , we obtain

$$\mathcal{M}(\lambda) = T^{-n\lambda} \frac{\int_{\mathbb{R}_+^n} \prod_{i<j} (x_i - x_j)^2 \prod_{k=1}^n (T + x_k)^\lambda x_k^\alpha e^{-nx_k} dx_k}{\int_{\mathbb{R}_+^n} \prod_{i<j} (x_i - x_j)^2 \prod_{k=1}^n x_k^\alpha e^{-nx_k} dx_k} \quad (1.36)$$

where  $T := \beta_1/P$ , with  $\beta_1 = n_t/n$ . Equivalently,

$$\mathcal{M}(\lambda) = T^{-n\lambda} \frac{Z_n(\lambda)}{Z_n(0)} \quad (1.37)$$

where

$$Z_n(\lambda) = \int_{\mathbb{R}_+^n} \exp\left(-\Phi(x_1, \dots, x_n) + \lambda \sum_{i=1}^n \log(T + x_i)\right) \prod_{k=1}^n dx_k \quad (1.38)$$

with

$$\Phi(x_1, \dots, x_n) = -2 \sum_{1 \leq j < k \leq n} \log|x_j - x_k| + n \sum_{\ell=1}^n (x_\ell - \beta \log x_\ell) \quad (1.39)$$

and we have defined  $\beta := m/n - 1$ .

This representation will be important for deriving an approximation for the moment generating function based on the Coulomb fluid approach in Section 3.

### 1.3 Multi-User MIMO and the Deformed Jacobi Weight

In this section we introduce the second communication scenario which we will consider, corresponding to a multi-user MIMO system. Such systems are highly relevant for cellular mobile applications, where the transmit-receive communication channel is impaired by interference from other users operating within the same frequency bandwidth. In fact, the key issue of interference presents one of the most important challenges in the successful deployment of MIMO in practice [45]. We will focus on the practical *interference-limited* scenario, where the receiver noise is negligible compared with the multi-user interference. We assume that there are  $K$  interferers, each equipped with  $n_t$  antennas, and transmitting (capacity achieving) independent Gaussian signals with power  $P_I/n_t$  out of each transmit antenna. We make the common assumption (see e.g., [46, 47]), that the interferers have equal-power, which is valid when the interferers are located at similar distances to the receiver. Moreover, as discussed in [46], if this assumption is not met, then the equal power assumption leads to a lower bound on performance. We also assume



that  $n_r \leq n_t$ , which is appropriate for modeling the downlink communication (base-station to mobile transmission) of a interference-limited cellular system, where the number of transmit antennas deployed by the base-station may be large, but the number of receive antennas is highly restricted due to limited space constraints of the cellular mobile device.

Under the assumptions discussed above, the noise term  $\mathbf{n}$  in (1.5) takes the form

$$\mathbf{n} = \sum_{i=1}^K \mathbf{H}_i \mathbf{x}_i \quad (1.40)$$

where  $\mathbf{H}_i$  and  $\mathbf{x}_i$  denote the channel matrix and input vector for the  $i$ th interferer respectively. These are assumed independent across  $i$ , and independent of  $\mathbf{H}$  and  $\mathbf{x}$ . It is convenient to write this in the stacked form

$$\mathbf{n} = \mathbf{H}_I \mathbf{x}_I \quad (1.41)$$

where

$$\mathbf{H}_I = [\mathbf{H}_1, \dots, \mathbf{H}_K], \quad \mathbf{x}_I = [\mathbf{x}_1, \dots, \mathbf{x}_K]^T. \quad (1.42)$$

Here,  $\mathbf{H}_I \in \mathbb{C}^{n_r \times K n_t}$  is complex Gaussian with independent zero-mean unit-variance entries, whilst  $\mathbf{x}_I \in \mathbb{C}^{K n_t}$  is complex Gaussian with independent zero-mean entries having variance  $P/n_t$ . The noise covariance matrix, conditioned on  $\mathbf{H}_I$ , is therefore given by

$$\mathbf{Q}_n = \frac{P_I}{n_t} \mathbf{H}_I \mathbf{H}_I^\dagger. \quad (1.43)$$

In this case, the moment generating function of the capacity particularizes to

$$\mathcal{M}(\lambda) = E_{\mathbf{H}} \left[ \det \left( \mathbf{I}_{n_r} + \frac{P}{P_I} \mathbf{H} \mathbf{H}^\dagger (\mathbf{H}_I \mathbf{H}_I^\dagger)^{-1} \right)^\lambda \right]. \quad (1.44)$$

Similarly, the error exponent (1.19) admits the same form, but with the substitution  $\lambda = -\rho$  and replacing  $P_I$  with  $(1 + \rho)P_I$ . Here,  $\mathbf{H} \mathbf{H}^\dagger$  and  $\mathbf{H}_I \mathbf{H}_I^\dagger$  are independent complex central Wishart.

As for the deformed Laguerre case, in the following we will focus our discussion on the moment generating function (1.44), keeping in mind that the application to the error probability is immediate.

For consistency with previous notation, let us define:

$$n := n_r, \quad m_1 := n_t, \quad m_2 := K n_t, \quad \alpha_1 := m_1 - n, \quad \alpha_2 := m_2 - n. \quad (1.45)$$

Generalizing [48, pp. 312-314] from real to complex matrices, we find that the joint probability density function of the eigenvalues  $f_1, \dots, f_n$  of the random matrix  $\mathbf{H}\mathbf{H}^\dagger(\mathbf{H}_I\mathbf{H}_I^\dagger)^{-1}$  is given by

$$p(f_1, f_2, \dots, f_n) \propto \prod_{k=1}^n \frac{f_k^{\alpha_1}}{(1 + f_k)^{m_1 + m_2}} \prod_{1 \leq i < j \leq n} (f_i - f_j)^2,$$

where  $f_k \in (0, \infty)$ ,  $k = 1, 2, \dots, n$ . With the change of variables<sup>3</sup>

$$f_k = \frac{x_k}{1 - x_k},$$

the above density becomes

$$f(x_1, x_2, \dots, x_n) \propto \prod_{k=1}^n w_{\text{Jac}}(x_k) \prod_{1 \leq i < j \leq n} (x_i - x_j)^2, \quad (1.46)$$

where  $x_k \in (0, 1)$ ,  $k = 1, 2, \dots, n$ , and  $w_{\text{Jac}}(\cdot)$  denotes the shifted classical Jacobi weight

$$w_{\text{Jac}}(x) = x^{\alpha_1}(1 - x)^{\alpha_2}, \quad x \in [0, 1]. \quad (1.47)$$

Note that the classical Jacobi weight has the form

$$\tilde{w}_{\text{Jac}}(x) = (1 - x)^{\alpha_1}(1 + x)^{\alpha_2}, \quad x \in [-1, 1] \quad (1.48)$$

---

<sup>3</sup>We would like to thank Iain Johnstone for pointing this out.

and therefore

$$w_{\text{Jac}}(x) = \frac{\tilde{w}_{\text{Jac}}(1-2x)}{2^{\alpha_1+\alpha_2}}, \quad x \in [0, 1]. \quad (1.49)$$

The moment generating function can be evaluated as

$$\begin{aligned} \mathcal{M}(\lambda) &= E \left[ \prod_{k=1}^n \left( 1 + \frac{P}{P_I} f_k \right)^\lambda \right] \\ &= E \left[ \prod_{k=1}^n \left( 1 + \frac{P}{P_I} \frac{x_k}{1-x_k} \right)^\lambda \right] \\ &= t^{-n\lambda} E \left[ \prod_{k=1}^n \left( \frac{t+x_k}{1-x_k} \right)^\lambda \right] \\ &= t^{-n\lambda} \frac{\int_{(0,1)^n} \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \prod_{k=1}^n w_{\text{Jac}}(x_k) \left( \frac{x_k+t}{1-x_k} \right)^\lambda dx_k}{\int_{(0,1)^n} \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \prod_{k=1}^n w_{\text{Jac}}(x_k) dx_k}, \end{aligned}$$

where

$$t = \frac{P_I}{P - P_I}.$$

We note here that the  $t$  variable has *no*  $n$  dependence when expressed in terms of  $P_I/(P-P_I)$ , rather *unlike* the single user MIMO case. This is mathematically reasonable as the Jacobi weight is compactly supported.

Applying the Andreief-Heine identity (1.28), we obtain

$$\mathcal{M}(\lambda) = t^{-n\lambda} \frac{D_n(t, \lambda, m_1, m_2)}{D_n(t, 0, m_1, m_2)} \quad (1.50)$$

where

$$D_n(t, \lambda, m_1, m_2) = \det (\mu_{i+j}(t, \lambda, m_2, m_2))_{i,j=0}^{n-1} \quad (1.51)$$

is the Hankel determinant generated from the deformed (shifted) Jacobi weight

$$w_{\text{dJac}}(x) = w_{\text{dJac}}(x, t, \lambda) := w_{\text{Jac}}(x) \left( \frac{x+t}{1-x} \right)^\lambda \quad (1.52)$$

with moments

$$\mu_k(t, \lambda, m_1, m_2) := \int_0^1 x^k w_{\text{dJac}}(x) dx, \quad k = 0, 1, 2, \dots \quad (1.53)$$

**Remark 3** The factor  $D_n(t, 0, m_1, m_2)$  is simply the Hankel determinant generated from the (shifted) non-deformed Jacobi weight,  $w_{\text{Jac}}(x)$ , which can be computed exactly in terms of the Barnes  $G$ -function. To this end, we apply the transformation (1.49) to give

$$\begin{aligned} D_n(t, 0, m_1, m_2) &= D_{n, \alpha_1, \alpha_2}[w_{\text{Jac}}] \\ &= \frac{1}{2^{(\alpha_1 + \alpha_2 + n)^n}} \frac{1}{n!} \int_{(-1, 1)^n} \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \prod_{k=1}^n \tilde{w}_{\text{Jac}}(x_k) dx_k \\ &= \frac{(2\pi)^n}{4^{n(\alpha_1 + \alpha_2 + n)}} \frac{\Gamma\left(\frac{\alpha_1 + \alpha_2 + 1}{2}\right) G^2\left(\frac{\alpha_1 + \alpha_2 + 1}{2}\right) G^2\left(\frac{\alpha_1 + \alpha_2}{2} + 1\right)}{G(\alpha_1 + \alpha_2 + 1) G(\alpha_1 + 1) G(\alpha_2 + 1)} \\ &\quad \times \frac{G(n+1)G(n+\alpha_1+1)G(n+\alpha_2+1)G(n+\alpha_1+\alpha_2+1)}{G^2\left(n + \frac{\alpha_1 + \alpha_2 + 1}{2}\right) G^2\left(n + \frac{\alpha_1 + \alpha_2}{2} + 1\right) \Gamma\left(n + \frac{\alpha_1 + \alpha_2 + 1}{2}\right)} \end{aligned} \quad (1.54)$$

where, to obtain the last equality, we have invoked [35, Eq. (1.6)].

Similarly, for  $t = 0$  and fixed  $\lambda$ , it follows that

$$D_n(0, \lambda, m_1, m_2) = D_{n, \alpha_1 + \lambda, \alpha_2 - \lambda}[w_{\text{Jac}}]. \quad (1.55)$$

**Remark 4** The moments can be expressed in terms of the hypergeometric function  ${}_2F_1(\cdot)$  as follows:

$$\mu_k(t, \lambda, m_1, m_2) = t^\lambda \Gamma(1 + k + \alpha_1) \Gamma(1 + \alpha_2 - \lambda) {}_2F_1(1 + k + \alpha_1, -\lambda, 2 + k + \alpha_1 + \alpha_2 - \lambda, -1/t).$$

This relation was also pointed out in [49]. Whilst this identity, combined with (1.55) and (1.50), gives a “closed-form” determinantal representation for the moment generating function, it does not provide useful insights and it also becomes unwieldy to evaluate if the number of antennas become large. To overcome these problems, in Section 2 we employ the theory of orthogonal polynomials and their corresponding ladder operators to provide a more useful characterization,

where we express the Hankel determinant generated from the deformed Jacobi weight in terms of the classical Painlevé VI differential equation.

Similar to (1.37), it will also be useful to note the following equivalent representation for the moment generating function,

$$\mathcal{M}(\lambda) = T^{-n\lambda} \frac{Z_n(\lambda)}{Z_n(0)} \quad (1.56)$$

where  $T = t$ ,

$$Z_n(\lambda) = \int_{(0,1)^n} \exp \left( -\Phi(x_1, \dots, x_n) + \lambda \sum_{\ell=1}^n \log \left( \frac{x_\ell + T}{1 - x_\ell} \right) \right) \prod_{k=1}^n dx_k \quad (1.57)$$

and

$$\Phi(x_1, \dots, x_n) = -2 \sum_{1 \leq i < j \leq n} \log |x_i - x_j| - n \sum_{\ell=1}^n (\varphi_1 \log x_\ell + \varphi_2 \log(1 - x_\ell)) \quad (1.58)$$

with

$$\varphi_1 = \frac{m_1}{n} - 1, \quad \varphi_2 = \frac{m_2}{n} - 1. \quad (1.59)$$

Once again, this representation will be critical for employing the Coulomb fluid methodology in Section 3.

## 2 Exact Characterization via the Ladder Operator Method

In the previous section, we demonstrated strong relationships between two information measures of MIMO channels, namely, the outage capacity and the error probability, and certain Hankel determinants. For single-user MIMO systems, the Hankel determinant of interest was generated via the moments of a deformation of the Laguerre weight, whereas for multi-user MIMO systems, it was generated via the moments of a deformed Jacobi weight. In this section, we present an exact characterization of these Hankel determinants by employing the theory of orthogonal

polynomials and their corresponding “raising and lowering” ladder operators. Before presenting the main results, we first introduce some preliminary material which will prove useful.

## 2.1 Preliminaries of Orthogonal Polynomials and their Ladder Operators

Here we provide a brief discussion to highlight the connections between orthogonal polynomials and the Hankel determinants of interest, as well as provide basic properties of orthogonal polynomials which will be needed. From (1.25) and (1.46), the joint eigenvalue distributions arising in the single-user and multi-user MIMO scenarios admit the generic form

$$p(y_1, \dots, y_n) \propto \prod_{k=1}^n \tilde{w}(y_k) \prod_{1 \leq i < j \leq n} (y_j - y_i)^2, \quad y_i \in (a, b)$$

with  $\tilde{w}(\cdot)$  denoting the weight. Moreover, in both cases, the key quantity of interest takes the generic form

$$D_n = \frac{1}{n!} \int_{(a,b)^n} \prod_{1 \leq i < j \leq n} (y_j - y_i)^2 \prod_{k=1}^n w(y_k) dy_k = \det \left( \int_a^b y^{i+j} w(y) dy \right)_{i,j=0}^n$$

where

$$w(y) = \tilde{w}(y)g(y)$$

denotes the deformed weight. As a direct consequence of the Vandermonde determinant factor, by applying elementary operations, this Hankel determinant can be equivalently expressed as

$$D_n = \det \left( \int_a^b P_i(y) P_j(y) w(y) dy \right)_{i,j=0}^n$$

where  $P_j(\cdot)$  represents any monic polynomial of degree  $j$ , written as

$$P_j(z) = z^j + \mathbf{p}_1(j) z^{j-1} + \dots \tag{2.1}$$

If we orthogonalize the polynomial sequence  $\{P_n(y)\}$  with respect to  $w(y)$  over the interval  $[a, b]$ , i.e.,

$$\int_a^b P_i(y)P_j(y)w(y)dy = h_i\delta_{i,j}, \quad i, j = 0, 1, 2, \dots \quad (2.2)$$

with  $h_i$  denoting the square of the  $L^2$  norm of  $P_i$ , then the Hankel determinant  $D_n$  evaluates to

$$D_n = \prod_{k=0}^{n-1} h_k .$$

Thus, we clearly see that the problem of computing the Hankel determinants of interest becomes one of characterizing the class of polynomials which are orthogonal with respect to the deformed weight function. To attack this problem, we require some definitions and tools, as given below.

We start by noting that if

$$\mu_i := \int_a^b y^i w(y) dy,$$

exists for all  $i = 0, 1, 2, \dots$ , then the theory of orthogonal polynomials states that  $P_n(y)$  for  $n = 0, 1, 2, \dots$  satisfies the three term recurrence relations,

$$zP_n(z) = P_{n+1}(z) + \alpha_n P_n(z) + \beta_n P_{n-1}(z). \quad (2.3)$$

The above sequence of polynomials can be generated from the orthogonality conditions, the recurrence relations and the initial conditions,

$$P_0(z) = 1, \quad \beta_0 P_{-1}(z) = 0.$$

For example

$$P_1(z) = z - \alpha_0 = z - \frac{\mu_1}{\mu_0}.$$

Substituting (2.1) into the recurrence relations, an easy computation shows that

$$\mathfrak{p}_1(n) - \mathfrak{p}_1(n+1) = \alpha_n, \quad (2.4)$$

with  $\mathfrak{p}_1(0) := 0$ . A telescopic sum of (2.4) gives

$$\alpha_n = -\sum_{j=0}^{n-1} \mathfrak{p}_1(j). \quad (2.5)$$

From the recurrence relation (2.3) and the orthogonality relations (2.2), we find

$$\beta_n = \frac{h_n}{h_{n-1}}. \quad (2.6)$$

We shall see that  $\mathfrak{p}_1(n)$  plays an important role in later developments. For more information on orthogonal polynomials, we give reference to Szegő's treatise [50].

Next, we present three lemmas, which represent the ladder operators of orthogonal polynomials, as well as some supplementary conditions. Note that these results have been known for quite sometime; we reproduce them here for the convenience of the reader using the notation of [24], where one can also find a list of references to the literature. We would like to mention here that Magnus [51] was perhaps the first to apply these lemmas—albeit in a slightly different form—to random matrix theory and the derivation of Painlevé equations. It should be mentioned that Tracy and Widom also made use of the compatibility conditions in their systematic study of finite  $n$  matrix models [52]. See also [53, 54]. In presenting these lemmas, we use the following definition:

$$\mathfrak{v} = -\log w.$$

**Lemma 1** *Suppose  $\mathfrak{v} = -\log w$  has a derivative in some Lipschitz class with positive exponent. The lowering and raising operators satisfy the following:*

$$P'_n(z) = -B_n(z)P_n(z) + \beta_n A_n(z)P_{n-1}(z) \quad (2.7)$$



$$P'_{n-1}(z) = [B_n(z) + v'(z)]P_{n-1}(z) - A_{n-1}(z)P_n(z), \quad (2.8)$$

where

$$A_n(z) := \frac{1}{h_n} \int_a^b \frac{v'(z) - v'(y)}{z - y} P_n^2(y) w(y) dy \quad (2.9)$$

$$B_n(z) := \frac{1}{h_{n-1}} \int_a^b \frac{v'(z) - v'(y)}{z - y} P_n(y) P_{n-1}(y) w(y) dy. \quad (2.10)$$

A direct computation produces two fundamental supplementary (compatibility) conditions valid for all  $z \in \mathbb{C} \cup \{\infty\}$  and these are stated in the next lemma.

**Lemma 2** *The functions  $A_n(z)$  and  $B_n(z)$  satisfy the conditions:*

$$B_{n+1}(z) + B_n(z) = (z - \alpha_n)A_n(z) - v'(z) \quad (S_1)$$

$$1 + (z - \alpha_n)[B_{n+1}(z) - B_n(z)] = \beta_{n+1}A_{n+1} - \beta_n A_{n-1}(z) \quad (S_2).$$

It turns out that there is an equation which gives better insight into the coefficients  $\alpha_n$  and  $\beta_n$ , if  $(S_1)$  and  $(S_2)$  are suitably combined to produce a “sum rule” on  $A_n(z)$ . We state this in the next lemma. The sum rule, we shall see later, gives important information about the logarithmic derivative of the Hankel determinant.

**Lemma 3** *The functions  $A_n(z)$ ,  $B_n(z)$ , and the sum*

$$\sum_{j=0}^{n-1} A_j(z),$$

*satisfy the conditions:*

$$B_n^2(z) + v'(z)B_n(z) + \sum_{j=0}^{n-1} A_j(z) = \beta_n A_n(z) A_{n-1}(z). \quad (S'_2)$$

**Remark 5** If the  $v'(z)$  is a rational function in  $z$  then we observe that the divided difference

$$\frac{v'(z) - v'(y)}{z - y}$$

is also a rational function in  $z$  and  $y$ , from which we infer that (2.9) and (2.10) are rational functions of  $z$ . It is now clear that the compatibility conditions in Lemmas 2 and 3 would give further insights into the recurrence coefficients and certain auxiliary quantities that appear in the theory.

## 2.2 Painlevé V Continuous $\sigma$ -Form and the Deformed Laguerre Weight

In this subsection, we focus on the deformed Laguerre scenario, for which the main challenge is to characterize the numerator of (1.30). This is given by the following key result:

**Theorem 1:** The Hankel determinant of the deformed Laguerre weight  $w(x)$  in (1.32) admits the following representation:

$$D_n(t, \lambda) = t^{n\lambda} \exp\left(\int_{\infty}^t \frac{H_n(x) - n\lambda}{x} dx\right) \quad (2.11)$$

where  $H_n(t)$  satisfies the Painlevé V continuous Jimbo-Miwa-Okamoto  $\sigma$ -form:

$$(tH_n'')^2 = [tH_n' - H_n + H_n'(2n + \alpha + \lambda) + n\lambda]^2 - 4(tH_n' - H_n + \delta_n) [(H_n')^2 + \lambda H_n'] \quad (2.12)$$

with  $\delta_n := n(n + \alpha + \lambda)$ .

The remainder of this subsection is devoted to the proof of Theorem 1.

### 2.2.1 Compatibility Conditions, Recurrence Coefficients and Discrete Equations

We start by noting that since

$$v'(x) = -\frac{1}{w(x)} \frac{d}{dx} w(x)$$

for the problem at hand is a rational function, the divided difference

$$\frac{v'(z) - v'(y)}{z - y}$$

will also be a rational function of  $z$  and  $y$ . Consequently, the functions  $A_n(z)$  and  $B_n(z)$  are rational in  $z$ . Therefore  $(S_1)$  and  $(S'_2)$  will give insight into various  $n$  and  $t$  dependent auxiliary quantities, which we shall see later.

For the purpose of applying the ladder operator method, we carry out the preliminary computations,

$$\begin{aligned} v(z) &:= -\log w(z) = -\alpha \log z - \lambda \log(z+t) + z \\ v'(z) &= -\frac{\alpha}{z} - \frac{\lambda}{z+t} + 1 \\ \frac{v'(z) - v'(y)}{z - y} &= \frac{\alpha}{zy} + \frac{\lambda}{(z+t)(y+t)}, \end{aligned}$$

and we see that using the definition of our  $A_n(z)$ ,  $B_n(z)$  and applying integration by parts,

$$A_n(z) = \frac{1 - R_n(t)}{z} + \frac{R_n(t)}{z+t} \quad (2.13)$$

$$B_n(z) = -\frac{n + r_n(t)}{z} + \frac{r_n(t)}{z+t} \quad (2.14)$$

$$R_n(t) := \frac{\lambda}{h_n} \int_0^\infty \frac{[P_n(y)]^2}{y+t} w(y,t) dy \quad (2.15)$$

$$r_n(t) := \frac{\lambda}{h_{n-1}} \int_0^\infty \frac{P_n(y)P_{n-1}(y)}{y+t} w(y,t) dy. \quad (2.16)$$

For the purpose of using  $(S_1)$  and  $(S'_2)$ , in particular  $(S'_2)$ , we first state the following results obtained by substituting  $A_n(z)$  and  $B_n(z)$  given by (2.13)–(2.16):

$$\begin{aligned} B_n^2(z) + v'(z)B_n(z) + \sum_{j=0}^{n-1} A_j(z) &= z^{-2}[(n + r_n)^2 + \alpha(n + r_n)] \\ &+ z^{-1} \left\{ n - \sum_{j=0}^{n-1} R_j + r_n[\lambda - \alpha - t - 2(n + r_n)]/t + (n - \lambda)/t \right\} \end{aligned}$$

$$\begin{aligned}
& + (z+t)^{-1} \left\{ \sum_{j=0}^{n-1} R_j + r_n [t + \alpha - \lambda + 2(n+r_n)]/t - n\lambda/t \right\} \\
& + (z+t)^{-2} [r_n^2 - \lambda r_n].
\end{aligned}$$

Now from  $(S_1)$  we find,

$$-(2n+1+r_{n+1}+r_n) = \alpha - \alpha_n(1-R_n) \quad (2.17)$$

$$r_{n+1}+r_n = \lambda - R_n(t+\alpha_n) \quad (2.18)$$

and from  $(S'_2)$  we find,

$$(n+r_n)^2 + \alpha(n+r_n) = \beta_n(1-R_n)(1-R_{n-1}) \quad (2.19)$$

$$n - \sum_{j=0}^{n-1} R_j + \frac{r_n}{t} [\lambda - \alpha - t - 2(n+r_n)] + \frac{n(\lambda-t)}{t} = \frac{\beta_n}{t} [(1-R_{n-1})R_n + (1-R_{n-1})R_n] \quad (2.20)$$

$$\sum_{j=0}^{n-1} R_j + \frac{r_n}{t} [t + \alpha - \lambda + 2(n+r_n)] - \frac{n\lambda}{t} = -\frac{\beta_n}{t} [(1-R_n)R_{n-1} + (1-R_{n-1})R_n] \quad (2.21)$$

$$r_n^2 - \lambda r_n = \beta_n R_n R_{n-1}. \quad (2.22)$$

**Remark 6** Observe that (2.20) and (2.21) are equivalent. We shall see later that (2.20), when combined with certain relations, performs the sum

$$\sum_{j=0}^{n-1} R_j$$

automatically in closed form.

This sum will provide an important link between the logarithmic derivative of the Hankel determinant with respect to  $t$ ,  $\beta_n$ , and  $r_n$ , which is an essential step in establishing the Painlevé equation.

Note that although the difference relations (2.17)–(2.19) and (2.22) look rather complicated, these can be manipulated to give us insight into the recurrence coefficients  $\alpha_n$  and  $\beta_n$ .

Now the sum of (2.17) and (2.18) gives us a simple expression for the recurrence coefficient

$\alpha_n$  in terms of  $R_n$ :

$$\alpha_n = 2n + 1 + \alpha + \lambda - tR_n. \quad (2.23)$$

From (2.19) and (2.22) have

$$n(n + \alpha) + r_n(\alpha + \lambda + 2n) = \beta_n(1 - R_n - R_{n-1})$$

or equivalently

$$\beta_n(R_n + R_{n-1}) = \beta_n - n(n + \alpha) - r_n(\alpha + \lambda + 2n). \quad (2.24)$$

Now substituting (2.22) and (2.24) into either (2.20) or (2.21) to eliminate  $R_n$  and  $R_{n-1}$  leaves us the following very simple form for  $\sum_{j=0}^{n-1} R_j$ , which will play a crucial role later,

$$t \sum_{j=0}^{n-1} R_j = n(n + \alpha + \lambda) - \beta_n - tr_n. \quad (2.25)$$

But in view of (2.23), we have

$$\begin{aligned} t \sum_{j=0}^{n-1} R_j &= n(n + \alpha + \lambda) - \sum_{j=0}^{n-1} \alpha_j \\ &= n(n + \alpha + \lambda) + \mathbf{p}_1(n). \end{aligned} \quad (2.26)$$

Comparing (2.25) with (2.26) gives

$$\mathbf{p}_1(n) = -\beta_n - tr_n.$$

Note that  $\mathbf{p}_1(n)$  also depends on  $t$ , although this is not always displayed.

We are now in a position to find an expression for  $\beta_n$  in terms of  $r_n$  and  $R_n$ . This is found

by eliminating  $R_{n-1}$  from (2.24) and (2.22) resulting in

$$\beta_n = \frac{1}{1 - R_n} \left[ r_n(2n + \alpha + \lambda) + \frac{r_n^2 - \lambda r_n}{R_n} + n(n + \alpha) \right].$$

We summarize the above in the following lemma:

**Lemma 4** *The recurrence coefficients  $\alpha_n$  and  $\beta_n$  are expressed in terms of the auxiliary quantities  $r_n$  and  $R_n$  as:*

$$\alpha_n = 2n + 1 + \alpha + \lambda - tR_n \quad (2.27)$$

$$\beta_n = \frac{1}{1 - R_n} \left[ r_n(2n + \alpha + \lambda) + \frac{r_n^2 - \lambda r_n}{R_n} + n(n + \alpha) \right]. \quad (2.28)$$

Furthermore,

$$\sum_{j=0}^{n-1} R_j = n(n + \alpha + \lambda) - \beta_n - tr_n, \quad (2.29)$$

$$p_1(n) = -\beta_n - tr_n. \quad (2.30)$$

### 2.2.2 $t$ Evolution and Painlevé V: Continuous $\sigma$ -Form

In this next stage of the proof, we keep  $n$  fixed and vary  $t$ . The differential relations generated here when combined with the difference relations obtained previously will give us the desired Painlevé equation.

A straightforward computation shows that

$$\frac{d}{dt} \log h_n = R_n. \quad (2.31)$$

But, from (2.6), it follows that

$$\frac{d\beta_n}{dt} = \beta_n(R_n - R_{n-1}) \quad (2.32)$$

$$= \beta_n R_n - \frac{r_n^2 - \lambda r_n}{R_n}, \quad (2.33)$$

where the last equality follows from (2.22).

Differentiating

$$0 = \int_0^\infty x^\alpha (x+t)^\lambda e^{-x} P_n(x) P_{n-1}(x) dx$$

with respect to  $t$  produces

$$\begin{aligned} 0 &= \lambda \int_0^\infty (x+t)^{\lambda-1} e^{-x} P_n(x) P_{n-1}(x) dx + \int_0^\infty x^\alpha (x+t)^\lambda e^{-x} \left[ \frac{d}{dt} \mathbf{p}_1(n) x^{n-1} + \dots \right] P_{n-1}(x) dx \\ &= \lambda \int_0^\infty \frac{P_{n-1}(x) P_n(x)}{x+t} w(x) dx + h_{n-1} \frac{d}{dt} \mathbf{p}_1(n), \end{aligned}$$

finally resulting in

$$\frac{d}{dt} \mathbf{p}_1(n) = -r_n. \quad (2.34)$$

Upon noting (2.4), this implies

$$\frac{d\alpha_n}{dt} = r_{n+1} - r_n. \quad (2.35)$$

Now differentiating (2.30) with respect to  $t$  and noting (2.34), we find

$$\begin{aligned} \frac{d\mathbf{p}_1(n)}{dt} &= -\frac{d\beta_n}{dt} - \frac{d}{dt}(tr_n) \\ &= -\frac{d\beta_n}{dt} - r_n - t \frac{dr_n}{dt} \\ &= -r_n. \end{aligned}$$

The above result combined with (2.33) gives

$$\frac{d\beta_n}{dt} = -t \frac{dr_n}{dt} = \beta_n R_n - \frac{r_n^2 - \lambda r_n}{R_n}. \quad (2.36)$$

Because (2.28) expresses  $\beta_n$  as a quadratic in  $r_n$ , we see that  $r_n$  satisfies a Riccati equation;

$$t \frac{dr_n}{dt} = \frac{r_n^2 - \lambda r_n}{R_n} - \frac{R_n}{1 - R_n} \left[ r_n(2n + \alpha + \lambda) + \frac{r_n^2 - \lambda r_n}{R_n} + n(n + \alpha) \right]. \quad (2.37)$$

In fact there is another Riccati equation satisfied by  $R_n$ , which can be found as follows. Eliminating  $r_{n+1}$  from (2.18) and (2.35), and upon referring to (2.28), we find that

$$2r_n = t \frac{dR_n}{dt} + \lambda - R_n(t + 2n + \alpha + \lambda - tR_n). \quad (2.38)$$

Let us now eliminate  $r_n(t)$  from (2.37) and (2.38), which results in a second order ordinary differential equation (o.d.e.) satisfied by  $R_n(t)$ , where  $n$ ,  $\alpha$ , and  $\lambda$  appear as parameters. A further linear fractional change of variable

$$R_n(t) = 1 - \frac{1}{1 - y(t)} \quad \text{or} \quad y = 1 - \frac{1}{1 - R_n(t)},$$

shows that  $y(t)$  satisfies a Painlevé V:

$$y'' = \frac{3y - 1}{2y(y - 1)} (y')^2 - \frac{y'}{t} + \frac{(y - 1)^2}{t^2} \left( \frac{\alpha^2}{2} y - \frac{\lambda^2}{2y} \right) + \frac{(2n + 1 + \alpha + \lambda) y}{t} - \frac{y(y + 1)}{2(y - 1)}. \quad (2.39)$$

We note this is

$$P_V \left( \frac{\alpha^2}{2}, -\frac{\lambda^2}{2}, 2n + 1 + \alpha + \lambda, -\frac{1}{2} \right).$$

For the continuous  $\sigma$ -form of this  $P_V$ , note that

$$\begin{aligned} H_n &:= t \frac{d}{dt} \log D_n \\ &= t \frac{d}{dt} \sum_{j=0}^{n-1} \log h_j \\ &= t \sum_{j=0}^{n-1} R_j \\ &= n(n + \alpha + \lambda) - \beta_n - tr_n \end{aligned} \quad (2.40)$$

$$= n(n + \alpha + \lambda) + \mathfrak{p}_1(n), \quad (2.41)$$

where the last two equations follow from (2.29) and (2.30) of Lemma 4.



From (2.34), (2.40), and (2.41), we obtain expressions for  $\beta_n$  and  $r_n$  in terms of  $H_n$  and  $H'_n$ ,

$$\beta_n = n(n + \alpha + \lambda) + tH'_n - H_n \quad (2.42)$$

$$r_n = -H'_n, \quad (2.43)$$

where  $'$  denotes  $\frac{d}{dt}$ .

All we need to do now is to eliminate  $R_n$  to obtain a functional equation satisfied by  $H_n$ ,  $H'_n$  and  $H''_n$ ,

$$f(H_n, H'_n, H''_n, n, t) = 0.$$

For this purpose, we examine two quadratic equations satisfied by  $R_n$ , one of which is simply a rearrangement of (2.28) and reads

$$\frac{r_n^2 - \lambda r_n}{R_n} + \beta_n R_n = \beta_n - r_n(2n + \alpha + \lambda) - n(n + \alpha). \quad (2.44)$$

The other follows from a derivative of (2.42) with respect to  $t$  and (2.33),

$$\beta_n R_n - \frac{r_n^2 - \lambda r_n}{R_n} = tH''_n. \quad (2.45)$$

Solving for  $R_n$  and  $1/R_n$  from the linear system (2.44) and (2.45), we find

$$\begin{aligned} 2\beta_n R_n &= \beta_n - r_n(2n + \alpha + \lambda) - n(n + \alpha) + tH''_n \\ 2\left(\frac{r_n^2 - \lambda r_n}{R_n}\right) &= \beta_n - r_n(2n + \alpha + \lambda) - n(n + \alpha) - tH''_n, \end{aligned}$$

which we rewrite as follows

$$2R_n = 1 + \frac{tH''_n + (t + 2n + \alpha + \lambda)H'_n - H_n + n\lambda}{tH'_n - H_n + n(n + \alpha + \lambda)} \quad (2.46)$$

$$\frac{2}{R_n} = \frac{-tH''_n + (t + 2n + \alpha + \lambda)H'_n - H_n + n\lambda}{(H'_n)^2 + \lambda H'_n}. \quad (2.47)$$

The product (2.46) and (2.47) gives us the desired continuous  $\sigma$ -form (2.12).

It is finally worth noting that with

$$D_n(t, \lambda) =: t^{\delta_n} \tilde{D}_n, \quad (2.48)$$

then after a little computation we find that  $\tilde{D}_n$  satisfies the Toda molecule equation [55]

$$\frac{d^2}{dt^2} \log \tilde{D}_n = \frac{\tilde{D}_{n+1} \tilde{D}_{n-1}}{\tilde{D}_n^2}. \quad (2.49)$$

### 2.3 Discrete $\sigma$ -Form and the Deformed Laguerre Weight

As an alternative to the continuous PV  $\sigma$ -form, the following theorem establishes a *discrete*  $\sigma$ -form satisfied by the logarithmic derivative of the Hankel determinant taken with respect  $t$ ; which is a non-linear difference equation in  $n$ .

**Theorem 2:** The Hankel determinant of the deformed Laguerre weight  $w(x)$  in (1.32) admits the representation (2.11), with  $H_n$  satisfying the discrete  $\sigma$ -form

$$\begin{aligned} & \left[ \frac{n(n+\alpha)t + (\delta^2 H_n + t)[H_n - \delta_n]}{\delta^2 H_n + 2n + \alpha + \lambda + t} \right]^2 - \lambda \frac{n(n+\alpha)t + (\delta^2 H_n + t)[H_n - \delta_n]}{\delta^2 H_n + 2n + \alpha + \lambda + t} \\ & = \left[ \delta_n - H_n + \frac{n(n+\alpha)t + (\delta^2 H_n - t)[H_n - \delta_n]}{\delta^2 H_n + 2n + \alpha + \lambda + t} \right] (H_{n+1} - H_n)(H_n - H_{n-1}) \end{aligned} \quad (2.50)$$

where

$$\delta^2 H_n := H_{n+1} - H_{n-1}$$

denotes a second order difference in  $H_n$ .

The initial conditions are

$$H_1(t) = \frac{d}{dt} \log D_1(t, \lambda), \quad H_2(t) = \frac{d}{dt} \log D_2(t, \lambda)$$

with

$$D_1(t, \lambda) = \mu_0(t), \quad D_2(t, \lambda) = \mu_0(t)\mu_2(t) - \mu_1^2(t),$$

and the moments are defined in (1.35).

The proof follows similar methods as in the continuous case in the previous subsection; namely, we express the auxiliary quantities  $R_n$ ,  $r_n$  and the recurrence coefficient  $\beta_n$  in terms of  $H_n$  and  $H_{n\pm 1}$ , and substitute these into (2.22). To begin, note that since

$$H_n = t \sum_{j=0}^{n-1} R_j,$$

we find that

$$\begin{aligned} H_{n+1} - H_n &= tR_n \\ H_{n-1} - H_{n+1} &= t(R_n + R_{n-1}). \end{aligned} \tag{2.51}$$

Multiplying by  $\beta_n$  and using (2.24), we obtain the following *linear* equation in  $\beta_n$  and  $r_n$ ,

$$(t + \delta^2 H_n)\beta_n - (2n + \alpha + \lambda)t r_n = n(n + \alpha)t. \tag{2.52}$$

There is a further *linear* equation in  $\beta_n$  and  $r_n$ , obtained by rearranging (2.25), which is

$$\beta_n + t r_n = n(n + \alpha + \lambda) - H_n. \tag{2.53}$$

Solving for  $\beta_n$  and  $r_n$  from (2.52) and (2.53) leaves,

$$\beta_n = n(n + \alpha + \lambda) - H_n + \frac{n(n + \alpha)t + (\delta^2 H_n - t)[H_n - n(n + \alpha + \lambda)]}{\delta^2 H_n + 2n + \alpha + \lambda + t} \tag{2.54}$$

$$t r_n = \frac{n(n + \alpha)t + (\delta^2 H_n + t)[H_n - n(n + \alpha + \lambda)]}{\delta^2 H_n + 2n + \alpha + \lambda + t}. \tag{2.55}$$

The discrete  $\sigma$ -form is found by substituting (2.54), (2.55) and

$$t R_n = H_{n+1} - H_n,$$

into (2.22), i.e.,

$$r_n^2 - \lambda r_n = \beta_n R_n R_{n-1} .$$

## 2.4 Painlevé IV Continuous $\sigma$ -Form and the Deformed Jacobi Weight

We now consider the Hankel determinant generated by the deformed Jacobi weight,

$$x^{\alpha_1}(1-x)^{\alpha_2-\lambda}(x+t)^\lambda, \quad x \in (0, 1), \quad (2.56)$$

discussed in Section 1.3. In contrast to the deformed Laguerre weight, existing characterizations for this case are available. Specifically, such deformation was investigated by Magnus [51], where an auxiliary variable similar to our  $R_n$  in the last section was found to satisfy a particular Painlevé VI. The continuous  $\sigma$ -form associated with this  $P_{VI}$  was derived recently in [27]. Other related work dealing with this weight can be found in [56].

To state the results of [51] and [27] in our context, we must first introduce some additional notation. Let  $\{P_m(x)\}$  satisfy the orthogonality relations:

$$\int_0^1 P_m(x)P_n(x) x^{\alpha_1}(1-x)^{\alpha_2-\lambda}(x+t)^\lambda dy = h_m(t)\delta_{m,n} \quad (2.57)$$

and

$$R_m(t) = \frac{\alpha_2 - \lambda}{h_m(t)} \int_0^1 [P_m(y)]^2 y^{\alpha_1} (1-y)^{\alpha_2-\lambda-1}(y+t)^\lambda dy . \quad (2.58)$$

Then

$$y(t) = 1 - \frac{(1+t)R_m(t)}{2m + \alpha_1 + \alpha_2 + 1}$$

satisfies the following  $P_{VI}$ :

$$\begin{aligned} y'' &= \frac{1}{2} \left( \frac{1}{y} + \frac{1}{y-1} + \frac{1}{y+t} \right) (y')^2 \\ &- \left( \frac{1}{t} + \frac{1}{1+t} - \frac{1}{y+t} \right) y' \end{aligned}$$

$$+ \frac{y(y-1)(y+t)}{t^2(1+t)^2} \left( \nu_1 - \frac{\nu_2 t}{y^2} - \frac{\nu_3(1+t)}{(y-1)^2} + \frac{\nu_4 t(1+t)}{(y+t)^2} \right) \quad (2.59)$$

where

$$\nu_1 = \frac{1}{2}(2m + \alpha_1 + \alpha_2 + 1)^2, \quad \nu_2 = -\frac{\alpha_1^2}{2}, \quad \nu_3 = \frac{(\alpha_2 - \lambda)^2}{2}, \quad \nu_4 = \frac{1 - \lambda^2}{2}.$$

Furthermore, let

$$H_m(t) = -t(1+t) \frac{d \log D_m(t)}{dt} + c_1 t + c_2, \quad (2.60)$$

where

$$c_1 := -m(m + \alpha_1 + \alpha_2) - \frac{(\alpha_1 + \alpha_2 - \lambda)^2}{4} \quad (2.61)$$

$$c_2 := \frac{1}{4}[2m(m + \alpha_1 + \alpha_2) + (\alpha_2 - \lambda)(\alpha_1 + \alpha_2 - \lambda) - \lambda(\alpha_1 - \alpha_2 + \lambda)]. \quad (2.62)$$

Then  $H_m$  satisfies the following  $\sigma$ -form of  $P_{VI}$ :

$$\begin{aligned} -H'_m[t(1+t)H''_m]^2 &- [2H'_m(tH'_m - H_m) + (H'_m)^2 + r_1 r_2 r_3 r_4]^2 \\ &= (H'_m - r_1^2)(H'_m - r_2^2)(H'_m - r_3^2)(H'_m - r_4^2), \end{aligned} \quad (2.63)$$

where

$$r_1 = \frac{\alpha_1 + \alpha_2 - \lambda}{2}, \quad r_2 = \frac{\alpha_2 - \lambda + \alpha_1}{2}, \quad r_3 = \frac{2m + \alpha_1 + \alpha_2 - \lambda}{2}, \quad r_4 = \frac{2m + \alpha_1 + \alpha_2 + \lambda}{2},$$

and  $\alpha_1$  and  $\alpha_2$  are defined by (1.45).

As for the single-user case, the Hankel determinant for the multi-user situation can also be expressed as an integral of  $H_m$ :

$$D_m(t) = D_{n, \alpha_1, \alpha_2 - \lambda}[w_{\text{Jac}}] t^{n\lambda} \exp \left( \int_{\infty}^s \frac{c_1 s + c_2 - H_m(s)}{s(1+s)} ds \right).$$

Note here that we have tacitly assumed that  $t = P_I/(P - P_I)$  is strictly positive. However, we expect the equations (2.59) and (2.63) to be formally valid for all  $t \in \mathbb{R}$ .

### 3 Characterization via the Coulomb Fluid Method

In this section, we present an alternative characterization based on the Coulomb fluid method. As we will see, the key benefit of this approach is that it leads to simpler expressions than the exact results obtained via the ladder operator approach. Moreover, whilst this method is based on large- $n$  considerations, we will show numerically that the approximations are very accurate for very small dimensions also. In fact, in the following section, we will demonstrate that the Coulomb fluid approach actually captures the *exact* distribution of the mutual information to leading order in  $n$ .

Whilst the Coulomb fluid has been applied extensively in the context of statistical mechanics, it is relatively unfamiliar amongst the wireless communications and information theory communities. As such, in the following we will first present some basic background material, based mainly on [28], before deriving new results for both the single-user and multi-user MIMO systems of interest.

#### 3.1 Preliminaries of the Coulomb Fluid Method

Consider a function of the form

$$\frac{Z_n(\lambda)}{Z_n(0)} = \exp [-(F_n(\lambda) - F_n(0))] \quad (3.64)$$

where

$$Z_n(\lambda) := \int_{(L,U)^n} \exp \left[ -\Phi(x_1, \dots, x_n) - \lambda \sum_{i=1}^n f(x_i) \right] \quad (3.65)$$

with

$$\Phi(x_1, \dots, x_n) := -2 \sum_{1 \leq j < k \leq n} \log |x_j - x_k| + n \sum_{j=1}^n v(x_j). \quad (3.66)$$

This expression embraces the moment generating function representations for both the single-user MIMO capacity (1.38)–(1.39) and multi-user MIMO capacity (1.57)–(1.59), with appropriate selection of the functions  $f(\cdot)$  and  $v(\cdot)$ , and integration limits  $L$  and  $U$ .

The key motivation for this representation is that it admits a simple intuitive interpretation in terms of statistical physics, as observed in the seminal papers by Dyson [29]. In particular, interpreting the eigenvalues  $x_1, \dots, x_n$  as the positions of  $n$  identically charged particles, the function  $\Phi(x_1, \dots, x_n)$  is recognized as the total energy of the repelling charged particles, confined by a common external potential  $nv(x)$ . The function  $f(x)$  acts as a perturbation to the system, effectively modifying the external potential. For sufficiently large  $n$ , we can approximate the particles as a continuous fluid with a certain (limiting) density,  $\sigma(x)$ , and assume that it is supported on a single interval  $(a, b)$ . This density will correspond to the equilibrium density of the fluid, obtained via the constrained minimization

$$\min_{\sigma} F[\sigma] \quad \text{subject to} \quad \int_a^b \sigma(x) dx = 1 \quad (3.67)$$

with

$$F(\lambda) := \int_a^b \sigma(x) (n^2 v(x) + \lambda n f(x)) dx - n^2 \int_a^b \int_a^b \sigma(x) \log |x - y| \sigma(y) dx dy. \quad (3.68)$$

As a consequence of the Frostman Lemma [57, pg. 65], the equilibrium density satisfies the integral equation

$$v(x) + \frac{\lambda}{n} f(x) - 2 \int_a^b \ln |x - y| \sigma(y) dy = A, \quad x \in [a, b]$$

where  $A$  is the Lagrange multiplier which fixes the constraint that the equilibrium density has

total charge 1. See [57] for a detailed discussion. The above integral equation with logarithmic kernel is converted into a singular integral equation by taking a derivative with respect to  $x$  for  $x \in (a, b)$ ,

$$2\mathcal{P} \int_a^b \frac{\sigma(y)}{x-y} dy = v'(x) + \frac{\lambda}{n} f'(x),$$

where  $\mathcal{P}$  denotes Cauchy principal value.

If  $v(x)$  is convex in a set of positive measure, the solution to this problem can be found [28], with the optimal  $\sigma(\cdot)$  taking the form

$$\sigma(x) = \sigma(x, \lambda) = \sigma_0(x) + \frac{\varrho(x, \lambda)}{n}, \quad (3.69)$$

where

$$\sigma_0(x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi^2} \int_a^b \frac{v'(x) - v'(y)}{(x-y)\sqrt{(b-y)(y-a)}} dy \quad (3.70)$$

denotes the limiting density of the original system (i.e., in the absence of any perturbation), and

$$\varrho(x, \lambda) = \lambda \tilde{\varrho}(x) = \frac{\lambda}{2\pi^2 \sqrt{(b-x)(x-a)}} \mathcal{P} \int_a^b \frac{\sqrt{(b-y)(y-a)}}{y-x} f'(y) dy \quad (3.71)$$

represents the deformation of this density caused by the external perturbation. The solution theory of singular integral equations can be found in the monographs [58], [59], and [60]. See also [61] for numerous examples on the application singular integral equations to problems in elasticity, and [62] for the  $L_p$  version of the theory.

The boundary parameters  $a$  and  $b$  are chosen to satisfy the supplementary conditions

$$\int_a^b \frac{v'(x)}{\sqrt{(b-x)(x-a)}} dx = 0 \quad (3.72)$$



and

$$\frac{1}{2\pi} \int_a^b \frac{xv'(x)}{\sqrt{(b-x)(x-a)}} dx = 1. \quad (3.73)$$

With these results, for sufficiently large  $n$ , the ratio (3.64) is then approximated by

$$\frac{Z(\lambda)}{Z(0)} = \exp \left[ -\lambda^2 \frac{\mathcal{S}_1^{\text{Comm.}}(T)}{2} - \lambda \mathcal{S}_2^{\text{Comm.}}(T) \right] \quad (3.74)$$

where

$$\mathcal{S}_1^{\text{Comm.}}(T) = \int_a^b f(x) \tilde{\varrho}(x) dx, \quad \mathcal{S}_2^{\text{Comm.}}(T) = n \int_a^b f(x) \sigma_0(x) dx. \quad (3.75)$$

**Remark 7:** With the above results, the moment generating functions (1.37) and (1.56) become

$$\mathcal{M}(\lambda) \approx \exp \left[ -\lambda^2 \frac{\mathcal{S}_1^{\text{Comm.}}(T)}{2} - \lambda (\mathcal{S}_2^{\text{Comm.}}(T) + n \log T) \right] \quad (3.76)$$

which corresponds to a Gaussian distribution with mean and variance given by

$$\mu_{\text{Coulomb}} = -\mathcal{S}_2^{\text{Comm.}}(T) - n \log T, \quad \sigma_{\text{Coulomb}}^2 = -\mathcal{S}_1^{\text{Comm.}}(T). \quad (3.77)$$

Therefore, the outage probability can be obtained via

$$P_{\text{out}}(C_{\text{out}}) \approx \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{C_{\text{out}} - \mu}{\sqrt{2\sigma^2}} \right) \right]. \quad (3.78)$$

**Remark 8:** Based on the Coulomb fluid method, the error exponent (1.18) is approximated as follows:

$$E_r(R) = \max_{0 \leq \rho \leq 1} \left\{ \rho^2 \frac{\mathcal{S}_1^{\text{Comm.}}(T)}{2} - \rho (\mathcal{S}_2^{\text{Comm.}}(T) + n \log(T)) - \rho R \right\}. \quad (3.79)$$

The key challenge is to evaluate the quantities  $\mathcal{S}_1^{\text{Comm.}(T)}$  and  $\mathcal{S}_2^{\text{Comm.}(T)}$  for the single-user and multi-user MIMO scenarios. These problems are addressed in the following subsections. As we will see, in both cases we will need to solve numerous integrals which are quite complicated and are not readily available. Thus, to aid the reader, we have succinctly compiled the solutions to these integrals in the Appendix, along with some detailed derivations.

### 3.2 Coulomb Fluid and the Deformed Laguerre Weight

In this case, we have the particularizations

$$f(x) = -\log(T+x), \quad v(x) = x - \beta \log x, \quad L = 0, \quad U = \infty. \quad (3.80)$$

First consider the constants,  $a$  and  $b$ . These are determined by the equations,

$$\int_a^b \frac{x - \beta}{\sqrt{(b-x)(x-a)}} \frac{dx}{2\pi} = 1, \quad \int_a^b \frac{1 - \beta/x}{\sqrt{(b-x)(x-a)}} dx = 0.$$

With the integral identities (6.16)–(6.18), we obtain

$$\frac{a+b}{4} = 1 + \frac{\beta}{2}, \quad \sqrt{ab} = \beta, \quad (3.81)$$

which leads to

$$a = 2 + \beta - 2\sqrt{1 + \beta}, \quad b = 2 + \beta + 2\sqrt{1 + \beta}. \quad (3.82)$$

Now consider the limiting density,  $\sigma_0(x)$ . In this case, with (3.70), (3.81), and the integral identity (6.16), it can be easily verified that

$$\sigma_0(x) = \frac{1}{2\pi} \frac{\sqrt{(b-x)(x-a)}}{x}, \quad a < x < b \quad (3.83)$$

which is the celebrated Marčenko-Pastur law [63, 64]. Substituting this distribution along with  $f(x) = -\log(T+x)$  into (3.75), and integrating using the identities (6.1), (6.3), and (6.5) gives

$$\begin{aligned} \mathcal{S}_2^{\text{Comm.}}(T) = & -\frac{n}{2} \left[ (a+b) \log \left( \frac{\sqrt{T+a} + \sqrt{T+b}}{2} \right) - \frac{(\sqrt{T+a} - \sqrt{T+b})^2}{2} \right. \\ & \left. - \sqrt{ab} \log \left( \frac{(\sqrt{ab} + \sqrt{(T+a)(T+b)})^2 - T^2}{(\sqrt{a} + \sqrt{b})^2} \right) \right]. \end{aligned} \quad (3.84)$$

We note that an equivalent expression can also be obtained by changing variables  $x \rightarrow (\beta+1)x$  and invoking an integral result from [65]. The derivation of our result here, based on applying the Schwinger parametrization of the log function (6.15), has the advantage of being much more direct, and moreover it can be used to derive other integral expressions encountered with the Coulomb fluid approach, as shown in the Appendix. Such parametrization is ubiquitous in the analytical computation of integrals arising in quantum field theory; see for example [66].

For  $\tilde{\varrho}(x)$ , substituting  $f'(x) = -1/(T+x)$  into (3.71) and using the integral identity (6.10), we calculate

$$\tilde{\varrho}(x) = \frac{1}{2\pi\sqrt{(b-x)(x-a)}} \left( 1 - \frac{\sqrt{(T+a)(T+b)}}{x+T} \right). \quad (3.85)$$

Substituting this into (3.75), and applying the integral identities (6.1) and (6.2) we find,

$$\mathcal{S}_1^{\text{Comm.}}(T) = -2 \log \left[ \frac{1}{2} \left( \frac{T+a}{T+b} \right)^{1/4} + \frac{1}{2} \left( \frac{T+b}{T+a} \right)^{1/4} \right]. \quad (3.86)$$

Figure 1 plots the Gaussian approximation to the distribution of the normalized mutual information (per antenna) of single-user MIMO systems, based on combining (3.86), (3.84), and (3.77), and compares with the true distribution generated via numerical simulations. The results are shown for a relatively low SNR,  $P = 5$  dB, and for various antenna configurations. In all cases the Gaussian approximation is very accurate, even for  $n$  as low as 2. The situation changes, however, when  $P$  is increased, as shown clearly in Figure 2. In particular, as  $P$  increases, it is evident that the distribution starts to deviate from Gaussian, and that this deviation is most significant for small  $n$ . This behavior is in line with the results of [67], which shows that the tail

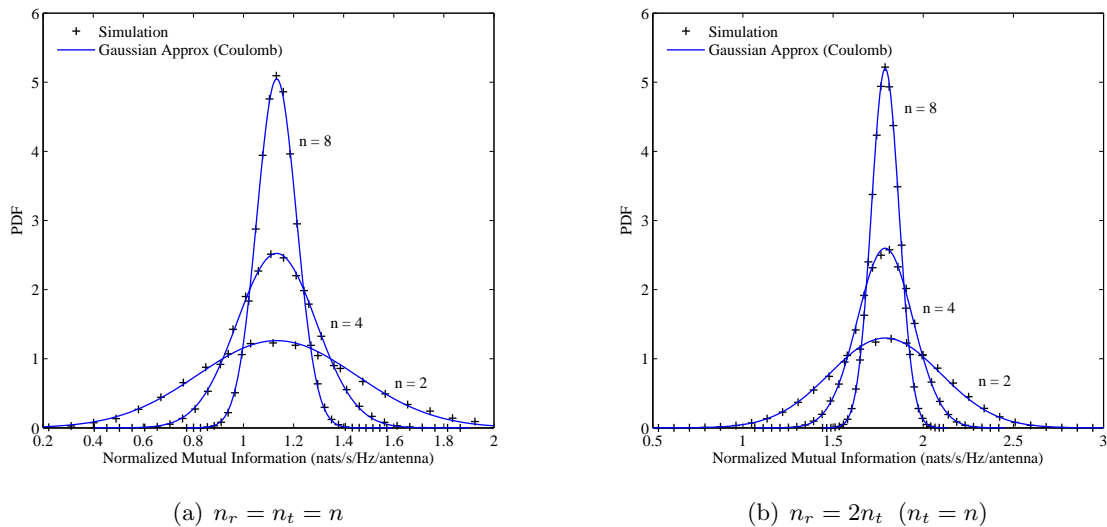


Figure 1: PDF of normalized mutual information  $I(\mathbf{x}, \mathbf{y})/n$  for the single-user MIMO scenario (deformed Laguerre case). Results shown for SNR,  $P = 5$  dB, and different antenna configurations. In all cases shown, the Coulomb fluid approximation is very accurate.

of the distribution behaves like an exponential random variable, rather than a Gaussian, when  $P \rightarrow \infty$  (and  $n$  fixed). We investigate this phenomenon in more detail in Section 4 (see also [36]).

Figure 3 compares the Coulomb fluid approximation for the error exponent, based on combining (3.86), (3.84) and (3.79), with the true error exponent computed via numerical simulation of (1.21) and (1.18). Interestingly, we see that in all cases, including both low *and high*  $P$ , the Coulomb fluid approximation is extremely accurate.

### 3.3 Coulomb Fluid and the Deformed Jacobi Weight

In this case, we have the particularizations:

$$f(x) = -\log\left(\frac{x+T}{1-x}\right), \quad v(x) = -\varphi_1 \log x - \varphi_2 \log(1-x), \quad L=0, \quad U=1. \quad (3.87)$$

First consider the limiting density  $\sigma_0(x)$ . This can be obtained by direct evaluation of (3.70). Alternatively, we may apply a simple transformation of the limiting density  $\rho(\cdot)$  corresponding

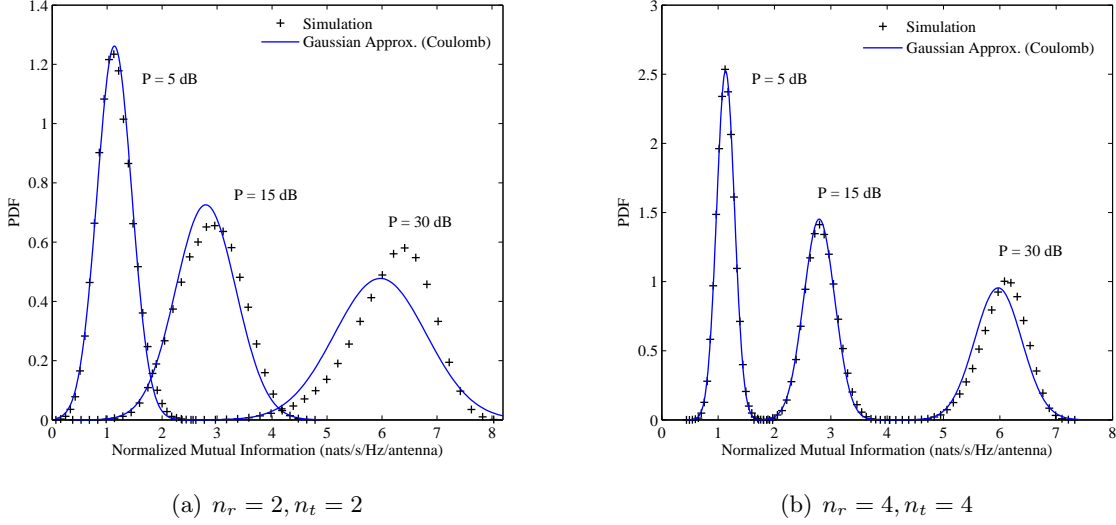


Figure 2: PDF of normalized mutual information  $I(\mathbf{x}, \mathbf{y})/n$  for the single-user MIMO scenario (deformed Laguerre case). Results shown for different SNR values. The Coulomb fluid approximation is very accurate when the SNR  $P$  is low, however it becomes less accurate (the mutual information distribution deviates from Gaussian) as  $P$  increases.

to the classical Jacobi weight (1.48), given by [35]

$$\rho(y) = \frac{n}{\pi} \frac{1 + (\varphi_1 + \varphi_2)/2}{1 - y^2} \sqrt{(B_n - y)(y - A_n)}, \quad y \in (A_n, B_n) \quad (3.88)$$

$$A_n := \frac{1}{(2n + \alpha_1 + \alpha_2 + 2)^2} \left[ \alpha_2^2 - \alpha_1^2 - 4\sqrt{n(n + \alpha_1)(n + \alpha_2)(n + \alpha_1 + \alpha_2)} \right] \quad (3.89)$$

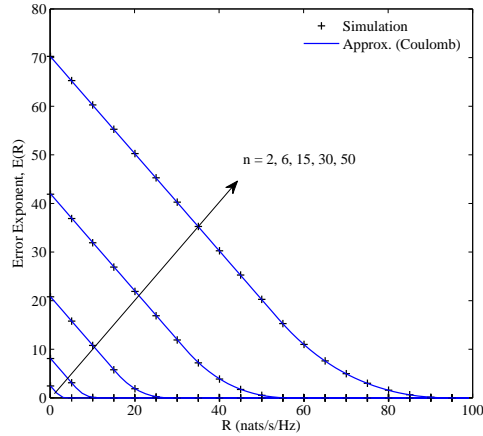
$$B_n := \frac{1}{(2n + \alpha_1 + \alpha_2 + 2)^2} \left[ \alpha_2^2 - \alpha_1^2 + 4\sqrt{n(n + \alpha_1)(n + \alpha_2)(n + \alpha_1 + \alpha_2)} \right] \quad (3.90)$$

$$A_n \rightarrow A := \frac{1}{(\varphi_1 + \varphi_2 + 2)^2} \left[ \varphi_2^2 - \varphi_1^2 - 4\sqrt{(1 + \varphi_1)(1 + \varphi_2)(1 + \varphi_1 + \varphi_2)} \right] \quad (3.91)$$

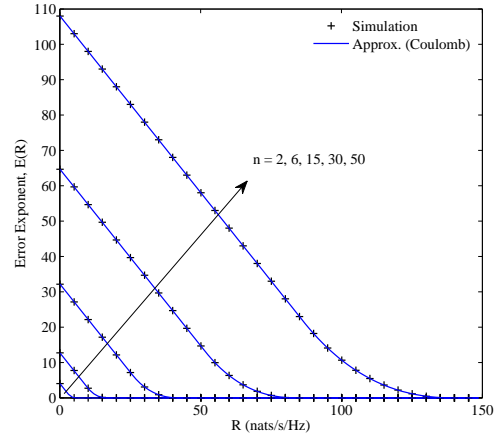
$$B_n \rightarrow B := \frac{1}{(\varphi_1 + \varphi_2 + 2)^2} \left[ \varphi_2^2 - \varphi_1^2 + 4\sqrt{(1 + \varphi_1)(1 + \varphi_2)(1 + \varphi_1 + \varphi_2)} \right]. \quad (3.92)$$

Here, the quantities  $A$  and  $B$  were obtained by recalling that  $\alpha_1 = n\varphi_1$  and  $\alpha_2 = n\varphi_2$ , and taking the limit  $n \rightarrow \infty$ . To relate this density to  $\sigma_0(x)$ , we first note that

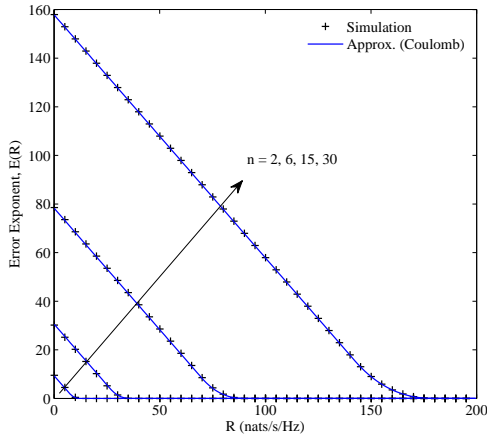
$$\mathbf{v}(x) = -\frac{\log w_{\text{Jac}}(x)}{n} \quad (3.93)$$



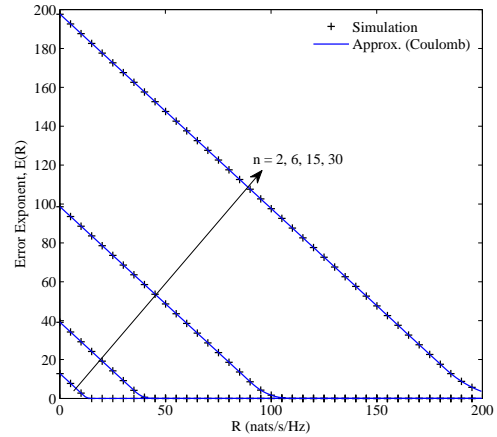
(a)  $n_r = n_t = n$ ,  $P = 10$  dB



(b)  $n_r = 2n_t$  ( $n_t = n$ ),  $P = 10$  dB



(c)  $n_r = n_t = n$ ,  $P = 30$  dB



(d)  $n_r = 2n_t$  ( $n_t = n$ ),  $P = 30$  dB

Figure 3: Error exponent for the single-user MIMO scenario (deformed Laguerre case). Results shown for SNR,  $P = 10$  dB. The Coulomb fluid approximation to the error exponent is very accurate for both low *and* high SNR ratios.

where  $w_{\text{Jac}}(x)$  is the classical Jacobi weight (1.47). Now, from the relation (1.49) we have

$$\begin{aligned} \mathbf{v}'(x) &= -\frac{w'_{\text{Jac}}(x)}{nw_{\text{Jac}}(x)} \\ &= -\frac{\tilde{w}'_{\text{Jac}}(1-2x)}{n\tilde{w}_{\text{Jac}}(1-2x)} \end{aligned} \quad (3.94)$$

which after substituting into (3.70) gives the desired result

$$\begin{aligned} \sigma_0(x) &= \frac{2}{n}\rho(1-2x) \\ &= [1 + (\varphi_1 + \varphi_2)/2] \frac{\sqrt{(b-x)(x-a)}}{\pi x(1-x)}, \quad a < x < b \end{aligned} \quad (3.95)$$

with

$$a := \frac{1-B}{2}, \quad b := \frac{1-A}{2}. \quad (3.96)$$

Substituting this distribution along with  $f(x) = -\log[(T+x)/(1-x)]$  into (3.75), applying the partial fraction decomposition

$$\frac{\sqrt{(b-x)(x-a)}}{x(1-x)} = \frac{1}{\sqrt{(b-x)(x-a)}} \left( 1 - \frac{ab}{x} + \frac{(1-a)(1-b)}{x-1} \right) \quad (3.97)$$

and integrating using the identities (6.1), (6.3)–(6.8), we find

$$\begin{aligned} \frac{\mathcal{S}_2^{\text{Comm.}}(T)}{n[1 + (\varphi_1 + \varphi_2)/2]} &= 2 \log \left( \frac{\sqrt{1-a} + \sqrt{1-b}}{\sqrt{T+a} + \sqrt{T+b}} \right) - \sqrt{ab} \log \left( \frac{1 - (\sqrt{ab} - \sqrt{(1-a)(1-b)})^2}{(\sqrt{ab} + \sqrt{(T+a)(T+b)})^2 - T^2} \right) \\ &+ \sqrt{(1-a)(1-b)} \log \left( \frac{(T+1)^2 - (\sqrt{(T+a)(T+b)} - \sqrt{(1-a)(1-b)})^2}{4(1-a)(1-b)} \right). \end{aligned}$$

We note that an alternative solution was also computed in [10], requiring the numerical evaluation of a certain fixed-point equation.

For  $\tilde{\rho}(x)$ , we substitute  $f'(x) = -1/(T+x) - 1/(1-x)$  into (3.71) and apply the integral

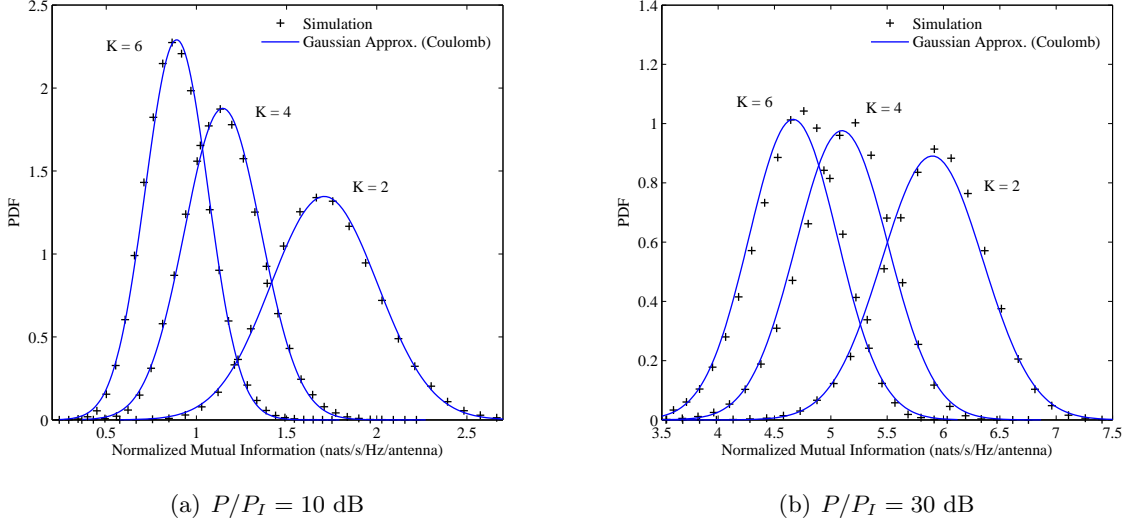


Figure 4: PDF of normalized mutual information  $I(\mathbf{x}, \mathbf{y})/n$  for the multi-user MIMO scenario (deformed Jacobi case). Results shown for  $n_t = 4$  and  $n_r = 3$ . The Coulomb fluid approximation is very accurate when the signal-to-interference ratio  $P/P_I$  is low, however it becomes less accurate (the mutual information distribution deviates from Gaussian) as  $P/P_I$  increases.

identity (6.10) along with (6.11) to yield

$$\tilde{q}(T) = \frac{1}{2\pi\sqrt{(b-x)(x-a)}} \left[ \frac{\sqrt{(1-a)(1-b)}}{1-x} - \frac{\sqrt{(T+a)(T+b)}}{x+T} \right]. \quad (3.98)$$

Substituting this into (3.75), and applying the integral identities (6.2), (6.5), (6.8), (6.9), we obtain

$$\mathcal{S}_1^{\text{Comm.}}(T) = \log \left[ \frac{4\sqrt{(T+a)(T+b)(1-a)(1-b)}}{(T+1)^2 - (\sqrt{(1-a)(1-b)} - \sqrt{(T+a)(T+b)})^2} \right]. \quad (3.99)$$

Figure 4 plots the Gaussian approximation to the distribution of the normalized mutual information (per antenna) of multi-user MIMO systems, based on combining (3.99), (3.98), and (3.77), and compares with the true distribution generated via numerical simulations. As for the single-user MIMO scenario, the Gaussian approximation is accurate regardless of the number of interferers  $K$ , especially when the signal-to-interference ratio  $P/P_I$  is not high. As  $P/P_I$  is increased, once again the distribution starts to deviate from Gaussian; however, this deviation



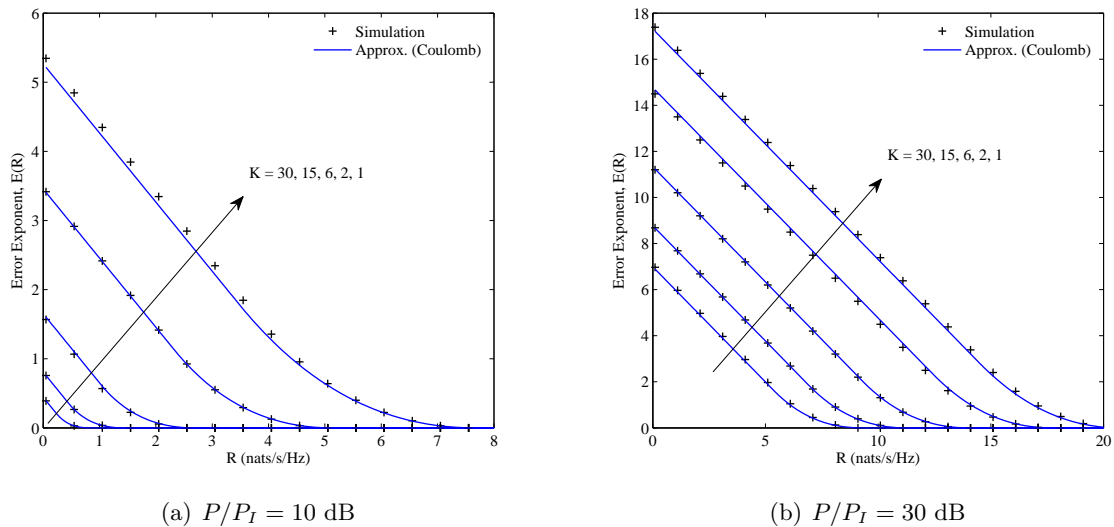


Figure 5: Error exponent for the multi-user MIMO scenario (deformed Jacobi case). Results shown for  $n_t = 4$  and  $n_r = 3$ . Similar to the single-user MIMO case, the Coulomb fluid approximation to the error exponent of multi-user MIMO channels is very accurate for both low and high signal-to-interference ratios.

is seemingly less significant than that seen previously for the single-user case.

Figure 5 compares the Coulomb fluid approximation for the error exponent, based on combining (3.99), (3.98) and (3.79), with the true error exponent computed via numerical simulation of (1.21) and (1.18). Again, we see that in all cases, the Coulomb fluid approximation is extremely accurate.

## 4 Beyond the Coulomb Fluid Approximation

In this section, we take a closer look at the Painlevé and Coulomb fluid representations of the mutual information distribution. We will focus on the Laguerre scenario (i.e., the single-user MIMO case), although the analysis can be extended to the Jacobi scenario also. Our main objective is to establish relationships between the exact characterization of the mutual information distribution via the Painlevé equation, and the simpler Coulomb fluid approximation. As a key result, we will show that for both the mean and the variance, the Coulomb fluid gives an *exact* representation to leading order in  $n$ . We also employ the Painlevé equation to establish the first-

order correction terms to the mean, variance, and third cumulant, which helps to characterize the deviation of the mutual information distribution from Gaussian.

## 4.1 Initial Study using Power Series Expansion

To get an intuitive feel for the problem, we start by looking for a  $1/t$  expansion in the Painlevé equation as  $t \rightarrow \infty$ . This is essentially the case where the SNR,  $P$ , is small. For simplicity, we consider the case  $m = n$ ; although the analysis can be extended to the case  $m \neq n$ .

### 4.1.1 Evaluating Cumulants from the Painlevé

To proceed, let

$$H_n(t) =: n\lambda + G_n(t) \tag{4.100}$$

such that

$$\mathcal{M}(\lambda) = \exp\left(\int_{\infty}^t \frac{G_n(x)}{x} dx\right). \tag{4.101}$$

Note also that

$$t \frac{d}{dt} \log \mathcal{M}(\lambda) = G_n(t). \tag{4.102}$$

From the Painlevé equation (2.12),  $G_n(t)$  satisfies

$$(tG_n'')^2 = (G_n'(t + 2n + \lambda) - G_n)^2 - 4(tG_n' - G_n + n^2)(G_n'^2 + \lambda G_n') \tag{4.103}$$

where the derivatives are with respect to  $t$ . Suppose that  $G_n(t)$  has a formal power series expansion in  $1/t$ :

$$\frac{G_n(t)}{n^2\lambda} = \sum_{k=1}^{\infty} \frac{b_k}{t^k} \tag{4.104}$$

where the coefficient  $b_k$  depends on  $n$  and  $\lambda$ . As a first step, we substitute the power series into (4.103) and find the first few  $b_k$ 's as follows:

$$\begin{aligned}
b_1 &= -1 \\
b_2 &= 2n - \lambda \\
b_3 &= -(1 + 5n^2 - 6n\lambda + \lambda^2) \\
b_4 &= 10n + 14n^3 - (5 + 29n^2)\lambda + 12n\lambda^2 - \lambda^3 \\
b_5 &= -(8 + 70n^2 + 42n^4 - (80n + 130n^3)\lambda + (15 + 95n^2)\lambda^2 - 20n\lambda^3 + \lambda^4) \\
b_6 &= 168n + 420n^3 + 132n^5 - (84 + 806n^2 + 562n^4)\lambda + (350n + 624n^3)\lambda^2 \\
&\quad - (35 + 235n^2)\lambda^3 + 30n\lambda^4 - \lambda^5 \\
b_7 &= -\{180 + 2121n^2 + 2310n^4 + 429n^6 - (2366n + 6510n^3 + 2380n^5)\lambda \\
&\quad + (469 + 4795n^2 + 3682n^4)\lambda^2 - (1120n + 2128n^3)\lambda^3 + (70 + 490n^2)\lambda^4 - 42n\lambda^5 + \lambda^6\}.
\end{aligned}$$

From these examples, it is clear that  $b_k$  takes the form

$$b_k = \sum_{\ell=0}^{k-1} \lambda^\ell \left[ B_{k,\ell} n^{k-\ell-1} + C_{k,\ell} n^{k-\ell-3} + O(n^{k-\ell-5}) \right], \quad (4.105)$$

where the coefficients  $B_{k,\ell}$  and  $C_{k,\ell}$  are *independent* of  $n$  and  $\lambda$ , and are computed via

$$B_{k,\ell} = \frac{1}{(k-\ell-1)! \ell!} \frac{d^{k-\ell-1}}{dn^{k-\ell-1}} \frac{d^\ell}{d\lambda^\ell} b_k \Big|_{\lambda=0, n=0}, \quad k = 1, 2, \dots, \infty, \quad \ell = 0, \dots, k-1 \quad (4.106)$$

and

$$C_{k,\ell} = \frac{1}{(k-\ell-3)! \ell!} \frac{d^{k-\ell-3}}{dn^{k-\ell-3}} \frac{d^\ell}{d\lambda^\ell} b_k \Big|_{\lambda=0, n=0}, \quad k = 3, 4, \dots, \infty, \quad \ell = 0, \dots, k-3 \quad (4.107)$$

respectively. Note that  $C_{k,\ell} = 0$  for  $k = 1$ ,  $k = 2$ , or  $\ell \geq k - 2$ . The  $B_{k,\ell}$  coefficients represent the leading order terms in  $n$ , whereas the  $C_{k,\ell}$  coefficients represent the next (non-zero) lower

order correction terms; for example,

$$B_{1,0} = -1, \quad B_{2,0} = 2, \quad B_{3,0} = -5, \quad B_{2,1} = -1, \quad B_{3,1} = 6, \quad B_{3,2} = -1$$

and

$$C_{3,0} = -1, \quad C_{4,0} = 10, \quad C_{5,0} = -70, \quad C_{3,1} = 0, \quad C_{4,1} = -5, \quad C_{5,1} = 80.$$

We aim to investigate the cumulants of the mutual information distribution, which in turn requires an expansion of the form

$$G_n(t) = \lambda g_1(t) + \lambda^2 g_2(t) + \dots \quad (4.108)$$

Together with (4.104) and (4.105), we then get

$$g_k(t) = g_k(n/P) = n^{2-k} \sum_{\ell=k}^{\infty} \left( B_{\ell,k-1} + \frac{C_{\ell,k-1}}{n^2} + O\left(\frac{1}{n^4}\right) \right) P^\ell, \quad k = 1, 2, \dots \quad (4.109)$$

Plugging this into (4.101) and integrating, the logarithm of the moment generating function becomes

$$\log \mathcal{M}(\lambda) = \sum_{k=1}^{\infty} \kappa_k \frac{\lambda^k}{k!} \quad (4.110)$$

where  $\kappa_k$  denotes the  $k$ th cumulant given by

$$\begin{aligned} \kappa_k &= -k! \int_0^P \frac{g_k(n/y)}{y} dy \\ &= -k! n^{2-k} \sum_{\ell=k}^{\infty} \frac{B_{\ell,k-1} + \frac{C_{\ell,k-1}}{n^2} + O\left(\frac{1}{n^4}\right)}{\ell} P^\ell. \end{aligned} \quad (4.111)$$

The mean is

$$\kappa_1 = n \left( P - P^2 + \frac{5}{3}P^3 - \frac{7}{2}P^4 + \dots \right) + \frac{1}{n} \left( \frac{1}{3}P^3 - \frac{5}{2}P^4 + \frac{70}{5}P^5 + \dots \right) + O\left(\frac{1}{n^3}\right)$$

the variance is

$$\kappa_2 = P^2 - 4P^3 + \frac{29}{2}P^4 - \frac{260}{5}P^5 + \dots + \frac{1}{n^2} \left( \frac{5}{2}P^4 - 32P^5 + \frac{806}{3}P^6 + \dots \right) + O\left(\frac{1}{n^4}\right)$$

the third cumulant is

$$\kappa_3 = \frac{1}{n} \left[ 2P^3 - 18P^4 + 114P^5 + \dots + \frac{1}{n^2} (18P^5 - 350P^6 + \dots) \right] + O\left(\frac{1}{n^5}\right)$$

and so on.

Importantly, this result demonstrates that as  $n$  grows large, the  $k$ th cumulant scales as

$$\kappa_k = O(n^{2-k}), \quad k = 1, 2, \dots \quad (4.112)$$

Thus, as expected, we see that the mean of the mutual information grows linearly with  $n$ , the variance converges to a constant, and all other cumulants disappear as  $n \rightarrow \infty$ . This reaffirms that the distribution becomes Gaussian for asymptotically large  $n$ .

Whilst in the analysis above we have substituted the numerical values of the constants  $B_{\ell,k}$ , explicit formulae can also be derived by directly using the Painlevé differential equation. We demonstrate the procedure by considering the coefficients in the mean summation (4.112), namely  $\{B_{\ell,0}\}$ . The same technique can be used to derive the formulae for the coefficients of the higher moments also.

Plugging the series (4.104) into the differential equation (4.103), keeping only the lowest order terms in  $\lambda$  and taking  $n$  large, the r.h.s. of (4.103) becomes

$$(Pn\lambda)^2 \left[ (4B_{1,0}^2 + 4B_{1,0}) + (12B_{1,0}B_{2,0} + 8B_{1,0}^2 + 8B_{2,0})P + \sum_{i=2}^{\infty} \chi_{i+2}P^i \right] \quad (4.113)$$

with

$$\chi_{i+2} = -4B_{i+1,0} - 8iB_{i,0} + \sum_{j=1}^{i-1} (2+j)B_{j+1,0} ((i-j+2)B_{i-j+1,0} + 4(i-j)B_{i-j,0}) . \quad (4.114)$$

Now consider the l.h.s. of (4.103). We have

$$\frac{(tG_n'')^2}{(n\lambda)^2} \sim \left( \sum_{i=1}^{\infty} \frac{i(i+1)B_{i,0}P^{i+1}}{n} \right)^2 \rightarrow 0 \quad (4.115)$$

as  $n \rightarrow \infty$ . Thus, dividing both the l.h.s. and r.h.s. by  $(Pn\lambda)^2$  and taking  $n$  large we get

$$(4B_{1,0}^2 + 4B_{1,0}) + (12B_{1,0}B_{2,0} + 8B_{1,0}^2 + 8B_{2,0})P + \sum_{i=2}^{\infty} \chi_{i+2}P^i = 0. \quad (4.116)$$

We calculate the  $B_{i,0}$ 's recursively, since every coefficient of  $P$  must equate to zero. Trivially, considering the constant and linear terms in  $P$ , we get

$$B_{1,0} = -1, \quad B_{2,0} = 2. \quad (4.117)$$

For higher order coefficients, we have the recurrence relation

$$B_{i+1,0} = -2iB_{i,0} + \frac{1}{4} \sum_{j=1}^{i-1} (2+j)B_{j+1,0} ((i-j+2)B_{i-j+1,0} + 4(i-j)B_{i-j,0}) \quad (4.118)$$

for  $i \geq 2$ . Based on this, the next few coefficients are evaluated as

$$B_{3,0} = -5, \quad B_{4,0} = 14, \quad B_{5,0} = -42. \quad (4.119)$$

It can be verified that this difference equation also admits the explicit solution

$$B_{\ell,0} = \frac{(-1)^\ell (2\ell)!}{\ell!(\ell+1)!}, \quad \ell = 1, 2, \dots \quad (4.120)$$

and therefore the mean of the mutual information takes the explicit form (to leading order of  $n$ )

$$\kappa_1 = -n \sum_{\ell=1}^{\infty} \frac{(-1)^\ell (2\ell)!}{\ell!(\ell+1)!} P^\ell + \mathcal{O}\left(\frac{1}{n}\right). \quad (4.121)$$

### 4.1.2 Comparison with the Coulomb Fluid

Now, consider the corresponding quantity derived based on the Coulomb fluid method:

$$\begin{aligned}\tilde{G}_n(1/P) &= t \frac{d}{dt} \log \mathcal{M}(\lambda) \\ &= \lambda \tilde{g}_1(1/P) + \lambda^2 \tilde{g}_2(1/P)\end{aligned}\tag{4.122}$$

where

$$\begin{aligned}\tilde{g}_1(1/P) &= P \frac{d}{dP} (\mathcal{S}_2^{\text{Comm.}}(1/P) + n \log P) \\ \tilde{g}_2(1/P) &= P \frac{d}{dP} \frac{\mathcal{S}_1^{\text{Comm.}}(1/P)}{2}.\end{aligned}\tag{4.123}$$

Here we have substituted the expression for  $\mathcal{M}(\lambda)$  given in (3.76), and used the fact that  $t = n/P$ .

From (3.84), we compute

$$\tilde{g}_1(1/P) = -n \left[ 1 + \frac{1}{2P} \left( 1 - \sqrt{1 + 4P} \right) \right]\tag{4.124}$$

which, after applying a Taylor expansion of  $\sqrt{1 + 4P}$  around zero, gives

$$\tilde{g}_1(1/P) = n \sum_{k=1}^{\infty} B_{k,0} P^k.\tag{4.125}$$

To leading order in  $n$ , this agrees precisely with  $g_1(t)$  in (4.109), thereby establishing that the Coulomb fluid method gives the *exact* value of the mean mutual information for large  $n$ .

From (3.86), we compute

$$\mathcal{S}_1^{\text{Comm.}}(1/P) = -P^2 + 4P^3 - \frac{29}{3}P^4 + 52P^5 - \frac{562}{3}P^6 + 680P^7 - \dots\tag{4.126}$$

giving

$$\tilde{g}_2(1/P) = -P^2 + 6P^3 - 29P^4 + 130P^5 - 562P^6 + \dots\tag{4.127}$$

This agrees precisely with  $g_2(t)$  in (4.109) to leading order in  $n$ , thereby confirming that the Coulomb fluid gives the correct asymptotic variance, in addition to the correct asymptotic mean.

In summary, we have

$$\begin{aligned}
\log \mathcal{M}(\lambda) &= \lambda \left( \mu_{\text{Coulomb}} + \frac{1}{n} \left( \frac{1}{3}P^3 - \frac{5}{2}P^4 + \frac{70}{5}P^5 + \dots \right) + \mathcal{O}\left(\frac{1}{n^3}\right) \right) \\
&+ \frac{\lambda^2}{2!} \left( \sigma_{\text{Coulomb}}^2 + \frac{1}{n^2} \left( \frac{5}{2}P^4 - 32P^5 + \frac{806}{3}P^6 + \dots \right) + \mathcal{O}\left(\frac{1}{n^4}\right) \right) \\
&+ \frac{\lambda^3}{3!} \left( \frac{1}{n} (2P^3 - 18P^4 + 114P^5 + \dots) + \frac{1}{n^3} (18P^5 - 350P^6 + \dots) + \mathcal{O}\left(\frac{1}{n^5}\right) \right) \\
&+ \sum_{k=4}^{\infty} \frac{\lambda^k}{k!} \mathcal{O}\left(\frac{1}{n^{k-2}}\right). \tag{4.128}
\end{aligned}$$

Here,  $\mu_{\text{Coulomb}}$  and  $\sigma_{\text{Coulomb}}^2$  represent the mean and variance respectively, calculated based on the Coulomb fluid method. All other terms represent correction terms, which essentially account for the deviation of the mutual information distribution from Gaussian for finite values of  $n$ .

## 4.2 Refined Analysis for All $P$

We now present a more refined analysis, which does not require a power series representation of  $1/t$ . This analysis is based on evaluating a non-perturbative summation of the perturbation series in  $P$ , obtained from the terms of  $G_n(t)$  which are linear in  $\lambda$ , and to leading order in  $n$ . For this purpose, we substitute the series representation (4.108) into (4.103).

### 4.2.1 Analysis of the Mean

We start by considering the first cumulant (i.e., the mean). By comparing the coefficients of  $\lambda^k$  on the l.h.s. and r.h.s., we find that the coefficient of  $\lambda$  is identically equal to 0, and the coefficient of  $\lambda^2$  satisfies:

$$(g_1)^2 - 4n^2 g_1' - 2[t + 2n]g_1 g_1' + [t^2 + 4nt](g_1')^2 - t^2(g_1'')^2 = 0.$$



Note that

$$t = \frac{n}{P} =: nT.$$

We have introduced  $T$  so that the differential equation does not get too complicated. After this change of variable in  $t$  to  $nT$ , without introducing further notation in place of  $g_1$ , the differential equation becomes

$$(g_1)^2 - 4ng_1' - 2(T+2)g_1g_1' + [T^2 + 4T](g_1')^2 - \frac{T^2(g_1'')^2}{n^2} = 0$$

where now  $'$  denotes the derivative with respect to  $T$ . Letting

$$g_1(nT) = nY(T) \tag{4.129}$$

we find that

$$Y^2 - 4Y' - 2(T+2)YY' + [T^2 + 4T]Y'^2 - \frac{T^2(Y'')^2}{n^2} = 0. \tag{Y}$$

Taking  $n \rightarrow \infty$ ,  $Y$  is seen to satisfy

$$Y^2 - 4Y' - 2(T+2)YY' + [T^2 + 4T]Y'^2 = 0. \tag{Y0}$$

Now consider  $\tilde{g}_1(T)$ , the term analogous to  $g_1(nT)$  but derived based on the Coulomb fluid. From (4.124),

$$\tilde{g}_1(T) = nY_0(T) \tag{4.130}$$

where

$$Y_0(T) = -\frac{4+T - \sqrt{T(4+T)}}{4+T + \sqrt{T(4+T)}}. \tag{4.131}$$

This expression is found to satisfy Eq. (Y0) identically, thus confirming that the Coulomb fluid approach gives the exact value for the mean of the mutual information to leading order in  $n$  for

all values of  $P$ . Note that this result is stronger than that derived in the previous section, since it applies even for values of  $P$  (or values of  $1/t$ ) for which a formal convergent power series in  $1/t$  does not exist.

To compute the  $1/n$  correction to the previously obtained  $Y$ , we substitute

$$Y(T) = Y_0(T) + \frac{1}{n^2}Y_1(T) + \mathcal{O}\left(\frac{1}{n^4}\right) \quad (4.132)$$

into Eq. (Y), and then obtain  $Y_1$  by setting the coefficient of  $1/n^2$  equal to 0. This gives

$$2Y_0Y_1 - 2(2+T)Y_1Y_0' - 4Y_1' - 2(2+T)Y_0Y_1' + 2T(4+T)Y_0'Y_1' - T^2Y_0'' = 0$$

which is a first order linear equation in  $Y_1$ . However, the coefficient of  $Y_1'$  vanishes identically when we make use of  $Y_0$  from (4.131). The solution of the algebraic equation reads:

$$Y_1(T) = -\frac{1}{\sqrt{T}(4+T)^{5/2}}.$$

With these results, we can compute the asymptotic mean of the mutual information, including the first-order corrections, as

$$\begin{aligned} \kappa_1 &= -\int_0^P \frac{g_1(n/z)}{z} dz \\ &= \mu_{\text{Coulomb}} + \frac{1}{n}\mu_{\text{Correction}} + \mathcal{O}\left(\frac{1}{n^3}\right) \end{aligned} \quad (4.133)$$

where  $\mu_{\text{Coulomb}}$  is the mean value computed via the Coulomb fluid, which from (3.84) and (3.77) with  $\beta = 0$  is given by

$$\mu_{\text{Coulomb}} = n \left[ 2 \log \left( \frac{1 + \sqrt{1 + 4P}}{2} \right) - \frac{(1 - \sqrt{1 + 4P})^2}{4P} \right] \quad (4.134)$$

and  $\mu_{\text{Correction}}$  is the first order correction term given by

$$\mu_{\text{Correction}} = -\int_0^P \frac{Y_1(1/z)}{z} dz = \frac{1}{12} \left[ \frac{1 + 6P + 6P^2}{(4P + 1)^{3/2}} - 1 \right]. \quad (4.135)$$

Note that if we expand this expression for  $\kappa_1$  around  $P = 0$ , the series matches precisely with (4.111) as expected.

### 4.2.2 Analysis of the Variance

Now we consider the second cumulant. To this end, after substituting (4.108) into (4.103) and setting the coefficient of  $\lambda^3$  equal to 0, we get

$$2g_1'(1 + g_1')(g_1 - tg_1') - 2n^2(1 + 2g_1')g_2' + \{-g_1 + (2n + t)g_1'\}\{-g_2 + g_1' + (2n + t)g_2'\} - t^2g_1''g_2'' = 0. \quad (4.136)$$

Once again, applying the change of variable  $t = nT$ , but without introducing new notation for  $g_1$  and  $g_2$ , we find

$$\frac{2}{n}g_1' \left(1 + \frac{g_1'}{n}\right) (g_1 - Tg_1') - 2n \left(1 + 2\frac{g_1'}{n}\right) g_2' + \{-g_1 + (T + 2)g_1'\}\{-g_2 + g_1'/n + (2 + T)g_2'\} - \frac{1}{n^2}T^2 g_1''g_2'' = 0 \quad (4.137)$$

with ' denoting  $d/dT$ .

Let

$$g_2(nT) = Z_0(T) + \frac{1}{n^2}Z_1(T) + O\left(\frac{1}{n^4}\right). \quad (4.138)$$

We first compute  $Z_0(T)$ , the leading order term in  $n$ . To this end, substituting (4.138) along with (4.129) and (4.132) into (4.137), and then keeping only the leading order terms in  $n$  (the terms which are linear in  $n$ ), we obtain

$$2Y_0'(1 + Y_0')(Y_0 - TY_0') - 2(1 + 2Y_0')Z_0' + \{-Y_0 + (T + 2)Y_0'\}\{-Z_0 + (T + 2)Z_0' + Y_0'\} = 0. \quad (4.139)$$

Interestingly, if we plug in the expression for  $Y_0(T)$  given in (4.131), we find that the coefficient

of  $Z'_0$  is identically equal to 0. Thus, (4.139) reduces to a simple algebraic equation, whose solution is:

$$Z_0(T) = -\frac{1}{2} + \frac{1}{2}\sqrt{\frac{T}{4+T}} + \frac{1}{4+T}.$$

After making the substitution  $T = 1/P$ , it can be verified that this expression matches precisely with  $\tilde{g}_2(T)$  in (4.123), derived based on the Coulomb fluid method. This result confirms that the Coulomb fluid approach gives the exact value for the variance of the mutual information to leading order in  $n$  for *all values of  $P$* . Again, this result is stronger than that derived in the previous section, since it applies for values of  $P$  (or values of  $1/t$ ) for which a formal convergent power series in  $1/t$  does not exist.

Now consider the correction term,  $Z_1(T)$ , in (4.138). Again we substitute (4.138) along with (4.129) and (4.132) into (4.137). In this case, however, we extract only the terms of order  $1/n$ , which gives a rather large first order equation in the unknown  $Z_1(T)$ . Fortunately, we find that by plugging in the previously determined equations for  $Y_0(T)$ ,  $Y_1(T)$ , and  $Z_0(T)$ , the coefficient of  $Z'_1(T)$  vanishes identically, and so we are left with a *linear equation* in  $Z_1(T)$ . This is easily solved and we find

$$Z_1(T) = -\frac{8 + 16T + 20T^2 + 5T^3 + 6\sqrt{T(4+T)} + 10T^{3/2}\sqrt{4+T} + 5T^{5/2}\sqrt{4+T}}{T(4+T)^4 \left(2 + 4T + T^2 + 2\sqrt{T(4+T)} + T^{3/2}\sqrt{4+T}\right)}. \quad (4.140)$$

With these results, we can compute the asymptotic variance of the mutual information, including the first-order corrections, as

$$\begin{aligned} \kappa_2 &= -2 \int_0^P \frac{g_2(n/z)}{z} dz \\ &= \sigma_{\text{Coulomb}}^2 + \frac{1}{n^2} \sigma_{\text{Correction}}^2 + \mathcal{O}\left(\frac{1}{n^4}\right) \end{aligned} \quad (4.141)$$

where  $\sigma_{\text{Coulomb}}^2$  is the variance computed via the Coulomb fluid, which from (3.86) and (3.77) with  $\beta = 0$  is given by

$$\sigma_{\text{Coulomb}}^2 = 2 \log \left[ \frac{(4P+1)^{1/4} + (4P+1)^{-1/4}}{2} \right]. \quad (4.142)$$

and  $\sigma_{\text{Correction}}^2$  is the first order correction term given by

$$\sigma_{\text{Correction}}^2 = -2 \int_0^P \frac{Z_1(1/z)}{z} dz = \frac{1}{12} \left[ 1 - \frac{8P^3(1-3P)}{(4P+1)^3} - \frac{12P^3 + 30P^2 + 10P + 1}{(4P+1)^{5/2}} \right]. \quad (4.143)$$

Note that if we expand this expression for  $\kappa_2$  around  $P = 0$  we get a series which matches precisely with (4.111), as expected.

### 4.2.3 Analysis of the Third Cumulant

Now consider the third cumulant. After substituting (4.108) into (4.103), setting the coefficient of  $\lambda^4$  equal to 0, and then going through the same procedure as before (i.e., applying  $t = nT$ ) we get

$$\begin{aligned} & 4 \left( 1 + 2\frac{g'_1}{n} \right) (g_1 - Tg'_1) \frac{g'_2}{n} + \frac{4}{n} g'_1 \left( 1 + \frac{g'_1}{n} \right) (g_2 - Tg'_2) \\ & + \left( -g_2 + \frac{g'_1}{n} + (2+T)g'_2 \right)^2 + 2(-g_1 + (T+2)g'_1) \left( -g_3 + \frac{g'_2}{n} + (2+T)g'_3 \right) \\ & - 4n^2 \left[ \left( \frac{g'_2}{n} \right)^2 + \left( 1 + 2\frac{g'_1}{n} \right) \frac{g'_3}{n} \right] - \frac{T^2}{n^2} (g''_2 + 2g'_1 g''_3) = 0 \end{aligned} \quad (4.144)$$

where the derivatives are taken with respect to  $T$ . Now substitute (4.138) along with (4.129) and (4.132) and

$$g_3(T) = \frac{1}{n} X_0(T) + \frac{1}{n^3} X_1(T) + \mathcal{O}\left(\frac{1}{n^5}\right), \quad (4.145)$$

where  $X_0(T)$  and  $X_1(T)$  are  $n$  independent. We find that the highest order term in  $n$  is in fact  $n$  independent, and using the previously determined equations for  $Y_0(T)$ ,  $Y_1(T)$ ,  $Z_0(T)$ , and  $Z_1(T)$ , once again we find that the coefficient of  $X'_0(T)$  vanishes identically. The resulting linear equation in  $X_0(T)$  has solution

$$X_0(T) = \frac{1}{2} \left( \frac{2\sqrt{T}}{(4+T)^{5/2}} + \frac{T}{(4+T)^2} - \frac{\sqrt{T}}{(4+T)^{3/2}} \right). \quad (4.146)$$

Proceeding in a similar manner, we find

$$\begin{aligned}
X_1(T) = & -\frac{16}{T^{3/2}(4+T)^{11/2}} - \frac{30}{T^{1/2}(4+T)^{11/2}} - \frac{4\sqrt{T}}{(4+T)^{11/2}} + \frac{9T^{3/2}}{(4+T)^{11/2}} - \frac{T^{5/2}}{2(4+T)^{11/2}} \\
& + \frac{8}{T(4+T)^5} + \frac{10}{(4+T)^5} - \frac{10T}{(4+T)^5} + \frac{T^2}{2(4+T)^5} .
\end{aligned} \tag{4.147}$$

With these results, we can compute the asymptotic third cumulant of the mutual information as

$$\begin{aligned}
\kappa_3 &= -3! \int_0^P \frac{g_3(n/z)}{z} dz \\
&= \frac{1}{n} \kappa_{3,\text{Correction A}} + \frac{1}{n^3} \kappa_{3,\text{Correction B}} + \mathcal{O}\left(\frac{1}{n^5}\right)
\end{aligned} \tag{4.148}$$

where

$$\begin{aligned}
\kappa_{3,\text{Correction A}} &= -6 \int_0^P \frac{X_0(1/z)}{z} dz \\
&= 1 - \frac{1}{\sqrt{1+4P}} - \frac{3P}{1+4P} + \frac{P}{(1+4P)^{3/2}}
\end{aligned} \tag{4.149}$$

and

$$\begin{aligned}
\kappa_{3,\text{Correction B}} &= -6 \int_0^P \frac{X_1(1/z)}{z} dz \\
&= \frac{16P^6}{(4P+1)^4} - \frac{48P^6}{(4P+1)^5} - \frac{380P^5}{3(4P+1)^5} - \frac{610P^4}{3(4P+1)^5} - \frac{323P^3}{3(4P+1)^5} \\
&\quad - \frac{80P^2}{3(4P+1)^5} - \frac{10P}{3(4P+1)^5} - \frac{1}{6(4P+1)^5} + \frac{42P^5}{(4P+1)^{9/2}} + \frac{93P^4}{(4P+1)^{9/2}} \\
&\quad + \frac{71P^3}{(4P+1)^{9/2}} + \frac{21P^2}{(4P+1)^{9/2}} + \frac{3P}{(4P+1)^{9/2}} + \frac{1}{6(4P+1)^{9/2}} .
\end{aligned} \tag{4.150}$$

Note that if we expand this around  $P = 0$  we get a series which matches precisely with (4.111).

The correction terms to the mean, variance, and third cumulant are plotted in Figure 6. In particular, the ‘‘Mean’’ curves represent

$$n(E[I(\mathbf{x}, \mathbf{y})] - \mu_{\text{Correction}})$$

with  $\mu_{\text{Correction}}$  given by (4.135); the “Variance” curves represent

$$n^2(\text{Var}[I(\mathbf{x}, \mathbf{y})] - \sigma_{\text{Correction}}^2)$$

with  $\sigma_{\text{Correction}}^2$  given by (4.143); and the “ $\kappa_3$ ” curves represent

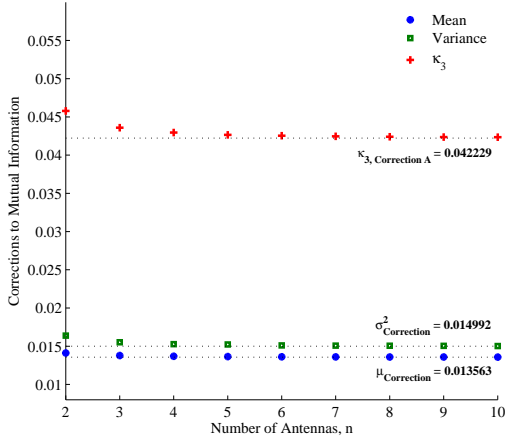
$$n(\kappa_3[I(\mathbf{x}, \mathbf{y})] - \kappa_{3,\text{Correction A}})$$

with  $\kappa_{3,\text{Correction A}}$  given by (4.149). Here, the mean  $E[I(\mathbf{x}, \mathbf{y})]$ , variance  $\text{Var}[I(\mathbf{x}, \mathbf{y})]$ , and third cumulant  $\kappa_3[I(\mathbf{x}, \mathbf{y})]$  of the mutual information were calculated using numerical integration procedures in Maple. From the figure, we can make some interesting observations. First, it is clearly evident that for low SNR  $P$ , all three correction terms converge very quickly to zero, confirming the near-Gaussian behavior of the distribution even for small  $n$ , as seen previously in Figure 2. However, as  $P$  increases, all three correction terms become much more significant. This, in turn, leads to a larger deviation from Gaussian, which again is in line with the numerics presented in Figure 2. It is also particularly interesting to note that the correction terms for the higher-order moments tend to *deviate quicker* than the lower order moments. To understand this phenomenon, it is useful to look closer at the correction terms as  $P$  grows large. This is the focus of the next subsection.

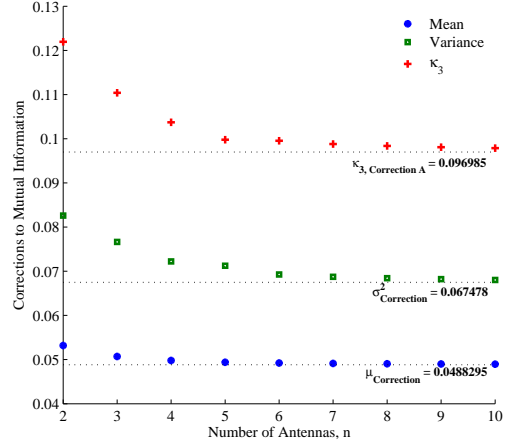
### 4.3 Analysis at Large $P$

For large  $P$ , the asymptotic mean (4.133), variance (4.141), and third cumulant (4.148) become

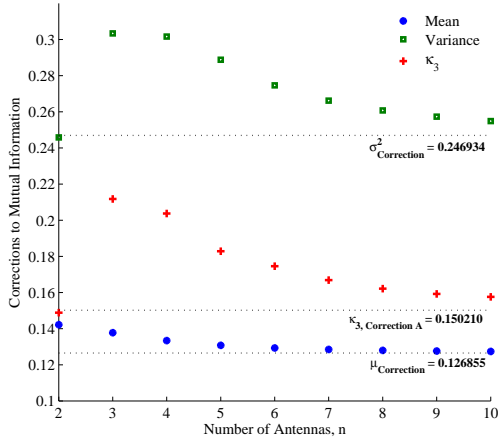
$$\begin{aligned} \mu &\sim n \log P + \frac{\sqrt{P}}{n} \left( \frac{1}{16} \right) \\ \sigma^2 &\sim \frac{\log(2P)}{2} + \frac{P}{n^2} \left( \frac{3}{8} \right) \\ \kappa_3 &\sim \frac{1}{n} \left( \frac{1}{4} \right) + \frac{P^2}{n^3} \left( \frac{1}{16} \right). \end{aligned} \tag{4.151}$$



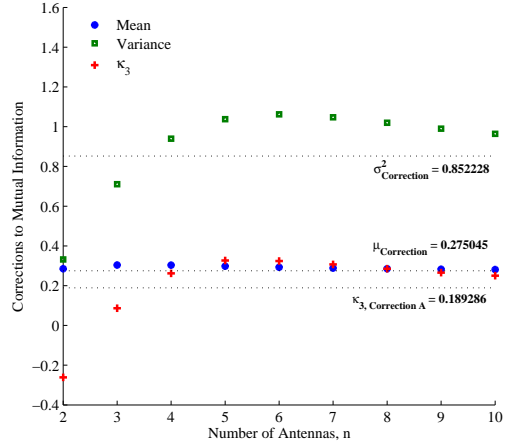
(a)  $P = 0$  dB



(b)  $P = 5$  dB



(c)  $P = 10$  dB



(d)  $P = 15$  dB

Figure 6: Correction terms to the mean and variance of the mutual information for the single-user MIMO scenario (deformed Laguerre case). Results are shown for  $n_r = n_t = n$ .



From these expressions, we notice that in all three cases, the correction terms (i.e., the second) are increasing in  $P$ . Since these terms give corrections to the Coulomb fluid Gaussian approximation, this implies that as  $P$  increases, the Coulomb fluid approximation loses accuracy, and the distribution deviates away from Gaussian, as our previous numerical results have indicated. We can also gain insights into the *relative* effect of  $P$  on each of the correction terms. Specifically, for the mean, we see that for the correction term to dominate the leading term, it must be at least  $O(n^4)$ . The variance, on the other hand, must only be  $O(n^2)$ , whilst the third cumulant is even more sensitive to  $P$ , and only requires  $P$  to be  $O(n)$ . This confirms that the higher moments are more sensitive to the variations in  $P$ , and moreover, it allows insights into the Gaussianity of the mutual information distribution in terms of *both*  $n$  and  $P$ .

We would like to mention, however, that some caution should be exercised in interpreting these results. In particular, since our analytical results which are asymptotic expansions in  $n$  are based on the Painlevé  $V$  representation of the moment generating function, they are formally valid for large  $n$  but *fixed*  $T$  or  $P$ . Therefore, it is to be expected that when  $P$  increases to the point where the corrections are comparable with and/or overwhelm the main terms, the Coulomb fluid approximation will break down. Providing a rigorous investigation of the situation where both  $n$  and  $P$  increase simultaneously is the subject of on-going work.

#### 4.4 Asymptotic Recurrence Coefficients

In addition to deriving the asymptotic moments, having determined the large  $n$  expansion of  $G_n(nT)$ , we can also compute the recurrence coefficients,  $\alpha_n$  and  $\beta_n$  for  $\alpha = 0$  and large  $n$  with  $t = nT$ . For this purpose, note that  $\alpha_n$  is easily related to  $R_n$  (see Eq. (2.20)) and  $R_n$  is in turn expressed in terms of  $H_n$  (or equivalently  $G_n$ ) and its derivatives (see Eq. (2.39)), while  $\beta_n$  related to  $H_n$  and its derivative through Eq. (2.35).

From our previous analysis, we have learned that for  $t = nT$  and  $n$  large,  $G_n(nT)$  has the expansion

$$G_n(nT) = \lambda \left( nY_0(T) + \frac{Y_1(T)}{n} + \dots \right) + \lambda^2 \left( Z_0(T) + \frac{Z_1(T)}{n^2} + \dots \right)$$

$$+ \lambda^3 \left( \frac{X_0(T)}{n} + \frac{X_1(T)}{n^3} + \dots \right) + \dots$$

From this series, after a straightforward if lengthy computation we find that

$$\begin{aligned} \alpha_n(nT) &= \left(2 - \frac{T}{2}\right) n + \lambda \left(1 - \sqrt{\frac{T}{4+T}}\right) \\ &+ \lambda \frac{\sqrt{T}}{(4+T)^2} \left[ \sqrt{4+T} + \lambda(\sqrt{4+T} - \sqrt{T}) \right] \frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right), \end{aligned} \quad (4.152)$$

and

$$\begin{aligned} \beta_n(nT) &= n^2 + \frac{4\lambda}{4+T + \sqrt{T(4+T)}} n \\ &- \frac{\left[2T + (2+T)(4+T)(-1 + \sqrt{T}/\sqrt{4+T})\right] \lambda^2}{2(4+T)^2} \\ &+ \lambda \frac{6 + 4T + \lambda^2(2T^2 + T^3 - T^{5/2}\sqrt{4+T} + 2T)}{\sqrt{T}(4+T)^{7/2}} \frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right). \end{aligned} \quad (4.153)$$

## 5 Concluding Remarks

This paper has focused on the computation of Hankel determinants which arise in the information-theoretic study of MIMO communication systems. We considered two practical scenarios; namely, single-user systems, in which case the determinants of interest are generated from a certain deformed Laguerre weight, and multi-user systems, in which case they are generated from a deformed Jacobi weight. In both cases, we derived new exact characterizations of the Hankel determinants in terms of classical Painlevé differential equations, as well as closed-form expressions which are formally valid for large matrix dimensions, but give accurate results for small dimensions also.

We also demonstrated that, in contrast to most prior work dealing with MIMO information theory, our exact and approximate results can be used together to derive explicit formula for the mean, variance, and higher order cumulants of the mutual information, as well as their corresponding correction terms. This is important, since it allows one to gain insights into the

Gaussianity of the mutual information distribution in terms of key system parameters, such as the number of antennas and the signal-to-noise ratio. For example, by focusing on the single-user MIMO scenario, we showed that the mutual information distribution may deviate strongly from Gaussian when the signal-to-noise ratio is relatively high, and we explicitly captured this effect through the cumulant formulas.

## 6 Appendix: Some Relevant Integral Identities

For the Coulomb fluid derivations, we will require the following integrals:

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}} dx = 2\pi \log\left(\frac{\sqrt{t+a} + \sqrt{t+b}}{2}\right) \quad (6.1)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)(x+t)}} dx = -\frac{2\pi}{\sqrt{(t+a)(t+b)}} \log\left(\frac{1}{2\sqrt{t+a}} + \frac{1}{2\sqrt{t+b}}\right) \quad (6.2)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)} x} dx = \frac{\pi}{\sqrt{ab}} \log\left(\frac{(\sqrt{ab} + \sqrt{(t+a)(t+b)})^2 - t^2}{(\sqrt{a} + \sqrt{b})^2}\right) \quad (6.3)$$

$$\int_a^b \frac{x \log(x+t)}{\sqrt{(b-x)(x-a)}} dx = \pi \frac{(\sqrt{a+t} - \sqrt{b+t})^2}{2} + \pi \frac{a+b}{2} \log\left(\frac{(\sqrt{(a+t)(b+t)} + t)^2 - ab}{4t}\right) \quad (6.4)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}(x-1)} dx = \pi \frac{\log\left(\frac{(\sqrt{1-a} + \sqrt{1-b})^2}{(t+1)^2 - (\sqrt{(t+a)(t+b)} - \sqrt{(1-a)(1-b)})^2}\right)}{\sqrt{(1-a)(1-b)}} \quad (6.5)$$

$$\int_a^b \frac{\log(1-x)}{\sqrt{(b-x)(x-a)}} dx = 2\pi \log\left(\frac{\sqrt{1-a} + \sqrt{1-b}}{2}\right) \quad (6.6)$$

$$\int_a^b \frac{\log(1-x)}{\sqrt{(b-x)(x-a)} x} dx = \frac{\pi}{\sqrt{ab}} \log\left(\frac{1 - (\sqrt{ab} - \sqrt{(1-a)(1-b)})^2}{(\sqrt{a} + \sqrt{b})^2}\right) \quad (6.7)$$

$$\int_a^b \frac{\log(1-x)}{\sqrt{(b-x)(x-a)}(x-1)} dx = \frac{2\pi}{\sqrt{(1-a)(1-b)}} \log\left(\frac{1}{2\sqrt{1-a}} + \frac{1}{2\sqrt{1-b}}\right) \quad (6.8)$$

$$\int_a^b \frac{\log(1-x)}{\sqrt{(b-x)(x-a)}(x+t)} dx = \pi \frac{\log\left(\frac{(t+1)^2 - (\sqrt{(t+a)(t+b)} - \sqrt{(1-a)(1-b)})^2}{(\sqrt{t+a} + \sqrt{t+b})^2}\right)}{\sqrt{(t+a)(t+b)}} \quad (6.9)$$

$$\mathcal{P} \int_a^b \frac{\sqrt{(b-y)(y-a)}}{(y-x)(y+t)} dy = \pi \left( \frac{\sqrt{(t+a)(t+b)}}{x+t} - 1 \right) \quad (6.10)$$

$$\mathcal{P} \int_a^b \frac{\sqrt{(b-y)(y-a)}}{(y-x)(y-1)} dy = \pi \left( \frac{\sqrt{(1-a)(1-b)}}{1-x} - 1 \right) \quad (6.11)$$

Before proving these results, we state for reference the following identities:

$$\int \frac{dx}{x\sqrt{ax^2+bx+c}} = -\frac{1}{\sqrt{c}} \log \frac{2\sqrt{c}\sqrt{ax^2+bx+c}+bx+2c}{x}, \quad c > 0, \quad (6.12)$$

$$\int \frac{dx}{\sqrt{(t+ax)(t+bx)}} = \frac{1}{\sqrt{ab}} \log \left[ 2\sqrt{ab}\sqrt{(t+a)(t+b)} + 2abx + (a+b)t \right] \quad (6.13)$$

$$\int \frac{dx}{x^2\sqrt{(x+a)(x+b)}} = \frac{a+b}{2(ab)^{3/2}} \log \left( \frac{(\sqrt{(x+a)(x+b)} + \sqrt{ab})^2 - x^2}{x} \right) - \frac{\sqrt{(x+a)(x+b)}}{abx} \quad (6.14)$$

$$\log(A+B) = \log A + \int_0^1 \frac{Bd\lambda}{A+\lambda B} \quad (6.15)$$

$$\int_a^b \frac{dx}{(x+t)\sqrt{(b-x)(x-a)}} = \frac{\pi}{\sqrt{(t+a)(t+b)}} \quad (6.16)$$

$$\int_a^b \frac{dx}{\sqrt{(b-x)(x-a)}} = \pi \quad (6.17)$$

$$\int_a^b \frac{xdx}{\sqrt{(b-x)(x-a)}} = \pi \frac{a+b}{2} \quad (6.18)$$

$$\mathcal{P} \int_a^b \frac{dy}{(x-y)\sqrt{(b-y)(y-a)}} = 0 \quad (6.19)$$

We start with some brief remarks concerning the identities (6.12)–(6.19), before focusing on the proofs of the main results (6.1)–(6.11). The integral (6.12) is [68, Eq. (2.266)], (6.13) is a minor variation of [68, Eq. (2.261)], and (6.14) follows from [68, Eq. (2.269.2)]. Note also that the r.h.s. of equation (6.15) is

$$\log A + \int_0^B \frac{dt}{A+t}$$

which obviously equals the l.h.s. To give an indication how (6.16)–(6.19) may be proved, first consider the analytic function

$$R(z) = \sqrt{(z-a)(z-b)}$$

defined in the complex plane slit along  $[a, b]$ . Here we assume that

$$0 < a < b < 1$$

without loss of generality. The branch of  $R(z)$  is chosen in such a way that

$$R(z) \rightarrow z, \quad \text{as } \Re z \rightarrow \infty.$$

Let  $F(x)$  be defined for  $x \in \mathbb{R}$  and extended to  $F(z)$ , a meromorphic function with poles. Let  $\Lambda$  be a “dog bone” contour traversed *clockwise* above and below the segment  $[a, b]$ , where the point of  $\infty$  is contained in the *interior* of  $\Lambda$ . Keeping in mind that

$$R_{\pm}(x) = \pm j \sqrt{(b-x)(x-a)}, \quad x \in (a, b),$$

where  $R_{\pm}(x)$  is defined to be the analytic continuation of  $R(z)$  to above and below the segment  $(a, b)$ , we have that

$$\int_a^b \frac{F(x)dx}{\sqrt{(b-x)(x-a)}} = \frac{j}{2} \int_{\Lambda} \frac{F(z)dz}{\sqrt{(z-a)(z-b)}},$$

and the r.h.s. of the above equation can be evaluated using residue calculus. The equation (6.16) follows immediately by computing the residue at  $t$ , while in computing (6.17) and (6.18) we should keep in mind the contributions from the residues at  $\infty$ . To compute the principal value integral (6.19), we first define

$$g(t) := \int_a^b \frac{dy}{(y+t)\sqrt{(b-y)(y-a)}} = \frac{\pi}{\sqrt{(t+a)(t+b)}}.$$

The principal value integral is then evaluated as

$$-\frac{1}{2} \lim_{\epsilon \rightarrow 0} [g(-x + j\epsilon) + g(-x - j\epsilon)], \quad x \in (a, b).$$

An easy computation gives (6.19).

We now come to the main integrals, (6.1)–(6.11). Start by considering (6.1)–(6.5). Of these, we will explicitly derive (6.2); the other integrals are evaluated in a similar way with the help of the properties (6.12)–(6.14). Using (6.15) along with (6.16), we obtain

$$\begin{aligned} & \int_a^b \frac{\log(x+t)dx}{(x+t)\sqrt{(b-x)(x-a)}} \\ &= \frac{\pi \log t}{\sqrt{(t+a)(t+b)}} + \pi \int_0^1 \frac{1}{\lambda-1} \left( \frac{1}{\sqrt{(t+a)(t+b)}} - \frac{1}{\sqrt{(t+a\lambda)(t+b\lambda)}} \right) d\lambda \\ &= \frac{\pi \log t}{\sqrt{(t+a)(t+b)}} + \pi \int_0^1 \frac{1}{x} \left( \frac{1}{\sqrt{(-ax+t+a)(-bx+t+b)}} - \frac{1}{\sqrt{(t+a)(t+b)}} \right) dx \\ &= \frac{\pi \log t}{\sqrt{(t+a)(t+b)}} + \pi \lim_{\epsilon \rightarrow 0} \int_\epsilon^1 \frac{1}{x} \left( \frac{1}{\sqrt{(-ax+t+a)(-bx+t+b)}} - \frac{1}{\sqrt{(t+a)(t+b)}} \right) dx, \end{aligned}$$

where we have made the substitution  $x = 1 - \lambda$  and have replaced  $\int_0^1 \dots$  by  $\lim_{\epsilon \rightarrow 0} \int_\epsilon^1 \dots$ , so that we may invoke (6.12). The integration is now completed as

$$\begin{aligned} & \int_a^b \frac{\log(x+t)}{(x+t)\sqrt{(b-x)(x-a)}} dx = \frac{\pi \log t}{\sqrt{(t+a)(t+b)}} \\ & \quad + \lim_{\epsilon \rightarrow 0} \frac{\pi}{\sqrt{(t+a)(t+b)}} \left( -\log \frac{\epsilon}{4(t+a)(t+b)} - \log t [2\sqrt{(t+a)(t+b)} + 2t + a + b] - \log(1/\epsilon) \right) \\ &= \frac{\pi}{\sqrt{(t+a)(t+b)}} \log \left( \frac{4(t+a)(t+b)}{2\sqrt{(t+a)(t+b)} + 2t + a + b} \right) \end{aligned} \tag{6.20}$$

$$\sim \pi \frac{\log t}{t}, \quad t \rightarrow \infty. \tag{6.21}$$

Note that the correct large  $t$  behavior is reproduced in (6.21). Some trivial algebra yields (6.2).

Now consider (6.6)–(6.9). We will explicitly derive (6.9); the integral (6.8) is then obtained by the analytical continuation of (6.9) to  $t = -1$ , whereas the integrals (6.6) and (6.7) are obtained by taking  $t \rightarrow \infty$  and  $t \rightarrow 0$  in (6.9) respectively. With the Schwinger parametrization

(6.15) and the partial fraction decomposition

$$\frac{x}{(x - 1/\lambda)(x + t)} = \frac{1/\lambda}{(x - \lambda)(t + 1/\lambda)} + \frac{t}{(t + 1/\lambda)(x + t)},$$

the integral becomes

$$\begin{aligned} \int_a^b \frac{\log(1-x)}{(x+t)\sqrt{(b-x)(x-a)}} dx &= \int_0^1 \frac{d\lambda}{1+\lambda t} \int_a^b \left( \frac{1/\lambda}{x-1/\lambda} + \frac{t}{x+t} \right) \frac{dx}{\sqrt{(b-x)(x-a)}} \\ &= \pi \int_0^1 \left( -\frac{1}{\sqrt{(1-\lambda a)(1-\lambda b)}} + \frac{t}{\sqrt{(t+a)(t+b)}} \right) \frac{d\lambda}{1+\lambda t}. \end{aligned}$$

The last equation was obtained by invoking (6.16) and taking the analytic continuation of  $t$  to  $-1/\lambda$ , together with the implicit assumption that  $\Re t > 0$  in (6.16). From a further change of variable  $1 + \lambda t = x$ , we have

$$\begin{aligned} \int_a^b \frac{\log(1-x)}{(x+t)\sqrt{(b-x)(x-a)}} dx &= \pi \int_1^{1+t} \left( \frac{1}{\sqrt{(t+a)(t+b)}} - \frac{1}{\sqrt{(t+a-ax)(t+b-bx)}} \right) \frac{dx}{x} \\ &= \frac{\pi}{\sqrt{(t+a)(t+b)}} \left[ \log(1+t) \right. \\ &\quad \left. + \log \frac{2\sqrt{(t+a)(t+b)}\sqrt{(t+a-ax)(t+b-bx)} - [t(a+b) + 2ab]x + 2(t+a)(t+b)}{x} \right]_{x=1}^{x=1+t} \\ &= \frac{\pi}{\sqrt{(t+a)(t+b)}} \log \frac{2\sqrt{(t+a)(t+b)}\sqrt{(1-a)(1-b)} + [2-a-b]t + a+b - 2ab}{2\sqrt{(t+a)(t+b)} + 2t+a+b} \\ &\sim \frac{\pi}{t} \log \frac{2\sqrt{(1-a)(1-b)} + 2-a-b}{4}, \quad t \rightarrow \infty. \end{aligned}$$

Simple algebra yields (6.9).

Finally, consider the principal value integrals (6.10) and (6.11). These results are obtained by taking the square root to the denominator of the integrand, followed by performing a partial fraction decomposition and invoking (6.18) and (6.19).

## References

- [1] E. TELATAR, Capacity of multi-antenna Gaussian channels, *European Transactions on Telecommunications* 10:585–595 (1999).
- [2] G. J. FOSCHINI and M. J. GANS, On the limit of wireless communications in fading environment when using multiple antennas, *Wireless Personal Communications* 6:311-335 (1998).
- [3] M. CHIANI, M. Z. WIN and A. ZANELLA, On the capacity of spatially correlated MIMO Rayleigh-fading channels, *IEEE Transactions on Information Theory* 49:2363-2371 (2003).
- [4] P. J. SMITH, S. ROY and M. SHAFI, Capacity of MIMO systems with semicorrelated flat fading, *IEEE Transactions on Information Theory* 49:2781-2788 (2003).
- [5] H. SHIN and J. H. LEE, Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole, *IEEE Transactions on Information Theory* 49:2636-2647 (2003).
- [6] M. R. MCKAY and I. B. COLLINGS, General capacity bounds for spatially correlated Rician MIMO channels, *IEEE Transactions on Information Theory* 51:3121-3145 (2005).
- [7] M. R. MCKAY, P. J. SMITH, H. A. SURAWEERA and I. B. COLLINGS, On the mutual information distribution of OFDM-based spatial multiplexing: Exact variance and outage approximation, *IEEE Transactions on Information Theory* 54:3260-3278 (2008).
- [8] X. GAO, B. JIANG, X. LI, A. B. GERSHMAN and M. R. MCKAY, Statistical eigenmode transmission over jointly-correlated MIMO channels, *IEEE Transactions on Information Theory* 55:3735-3750 (2009).
- [9] A. GRANT, Rayleigh fading multi-antenna channels, *EURASIP Journal on Applied Signal Processing* 3:316-329 (2002).
- [10] A. LOZANO and A. M. TULINO, Capacity of multiple-transmit multiple-receive antenna architectures, *IEEE Transactions on Information Theory* 48:3117-3128 (2002).



- [11] A. M. TULINO and S. VERDÚ, Random matrix theory and wireless communications, Foundations and Trends in Communications and Information Theory 1:1-163 (2004).
- [12] R. G. GALLAGER, Information Theory and Reliable Communication, John Wiley and Sons, Inc., New York, 1968.
- [13] W. C. BAULDRY, Estimate of the asymmetric Freud polynomials on the real line, J. Approx. Theory 63:225-237 (1990).
- [14] S. BELMEHDI and A. RONVEAUX, Laguerre-Frued's equations for the recurrence coefficients of semi-classical orthogonal polynomials, J. Approx. Theory 76:351-268 (1994).
- [15] S. BONAN and D. S. CLARK, Estimates of the orthogonal polynomials with weight  $\exp(-x^m)$ ,  $m$  an even positive integer, J. Approx. Theory 46:408-410 (1986).
- [16] S. BONAN and D. S. CLARK, Estimates of the Hermite and Freud polynomials, J. Approx. Theory 63:210-224 (1990).
- [17] S. BONAN, D. S. LUBINSKY and P. NEVAI, Orthogonal polynomials and their derivative. II, SIAM J. Math. Anal. 18:1163-1176 (1987).
- [18] S. BONAN and P. NEVAI, Orthogonal polynomials and their derivatives. I., J. Approx. Theory 40:134-147 (1984).
- [19] Y. CHEN and M. E. H. ISMAIL, Ladder operator and differential equations for orthogonal polynomials, J. Phys. A. 30:7817-7829 (1997).
- [20] Y. CHEN and M. E. H. ISMAIL, Jacobi polynomials from compatibility conditions, Proc. Amer. Math. Soc. 133:465-472 (2005).
- [21] E. BASOR and Y. CHEN, Painlevé V and the distribution function of a discontinuous linear statistics in the Laguerre unitary ensembles, J. Phys. A. 42, (2009).
- [22] Y. CHEN and G. PRUESSNER, Orthogonal polynomials with discontinuous weight, J. Phys. A. 38:L191-L198 (2005).

- [23] Y. CHEN and M. V. FEIGIN, Painlevé IV and degenerate Gaussian unitary ensembles, *J. Phys. A* 30:12381-12393 (2006).
- [24] Y. CHEN and A. R. ITS, Painlevé III and a singular linear statistics in Hermitian random matrix ensembles, I, *J. Approx. Theory* 162:270-297 (2010).
- [25] Y. CHEN and L. ZHANG, Painlevé VI and the unitary Jacobi ensembles, *Studies in Applied Mathematics*, Published Online: Mar 2010. DOI:10.1111/j.1467-9590.2010.00483.x.
- [26] E. BASOR, Y. CHEN and T. EHRHARDT, Painlevé V and time-dependent Jacobi polynomials, *J. Phys. A* 43 (2010) 015204.
- [27] D. DAI and L. ZHANG, Painlevé VI and Hankel determinants for the generalized Jacobi weight, *J. Phys. A: Math. Theor.* 43 (2010) 055207.
- [28] Y. CHEN and N. D. LAWRENCE, On the linear statistics of Hermitian random matrices, *J. Phys. A: Math. Gen.* 31:1141-1152 (1998).
- [29] F. J. DYSON, Statistical theory of energy levels of complex systems I-III, *J. Math. Phys.* 3:140-175 (1962).
- [30] Y. CHEN and M. E. H. ISMAIL, Thermodynamic relations of the Hermitian matrix ensembles, *J. Phys. A* 6:6633-6654 (1997).
- [31] Y. CHEN and S. M. MANNING, Asymptotic level spacing of the Laguerre ensemble: A Coulomb fluid approach, *J. Phys. A.: Math. Gen.* 27:3615-3620 (1994).
- [32] G. SZEGÖ, *Hankel forms*, vol. 1 (Basle: Birkhauser), p.111, English translation of *A Hankel-fèle forma krol: Collected papers*.
- [33] E. BASOR, Y. CHEN and H. WIDOM, Hankel determinants as Fredholm determinants: *Random Matrix Models and Their Applications* vol 40 (MSRI publications), Cambridge University Press, Cambridge, 21-29, 2001. <http://www.msri.org/publications/books/Book40/contents.html>

- [34] E. BASOR, Y. CHEN and H. WIDOM, Determinants of Hankel matrices, *J. Funct. Anal.* 179:214-234 (2001).
- [35] E. BASOR and Y. CHEN, Perturbed Hankel determinants, *J. Phys. A.: Math. Gen.* 38:10101-10106 (2005).
- [36] P. KAZAKOPOULOS, P. MERTIKOPOULOS, A. L. MOUSTAKAS and G. CAIRE, Living at the edge: A large deviations approach to the outage MIMO capacity, submitted to the *IEEE Transactions on Information Theory*, available at <http://arxiv.org/abs/0907.5024>.
- [37] Z. D. BAI and J. W. SILVERSTEIN, CLT of linear spectral statistics of large dimensional sample covariance matrices, *Annals of Probability* 32:553-605 (2004).
- [38] Y. CHEN and S. M. MANNING, Distribution of linear statistics in random matrix models (metallic conductance fluctuations), *J. Phys.: Cond. Matter* 6:3039-3044 (1994).
- [39] E. A. JORSWEICK and H. BOCHE, Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback, *IEEE Transactions on Wireless Communications* 3:1543-1553 (2004).
- [40] A. M. TULINO, A. LOZANO and S. VERDÚ, Capacity-achieving input covariance for single-user multi-antenna channels, *IEEE Transactions on Wireless Communications* 5:662-671 (2006).
- [41] C. CHUAH, D. TSE, J. KAHN and R. VALENZUELA, Capacity scaling in dual-antenna-array wireless systems, *IEEE Transactions on Information Theory* 48:637-650 (2002).
- [42] H. WIDOM, On the relation between orthogonal, symplectic and unitary matrix ensembles, *J. Stat. Phys.* 94:347-363 (1999).
- [43] M. KANG and M.-S. ALOUINI, Capacity of MIMO Rician channels, *IEEE Transactions on Wireless Communications* 5:112-122 (2006).
- [44] Z. WANG and G. B. GIANNAKIS, Outage mutual information of space-time MIMO channels, *IEEE Transactions on Information Theory* 50:657-662 (2004).

- [45] J. G. ANDREWS, W. CHOI and R. W. HEATH JR., Overcoming interference in spatial multiplexing MIMO cellular networks, *IEEE Communications Magazine* 14:95-104 (2007).
- [46] M. KANG, M.-S. ALOUINI and L. YANG, Outage probability and spectrum efficiency of cellular mobile radio systems with smart antennas, *IEEE Transactions on Communications* 50:1871-1877 (2002).
- [47] M. CHIANI, M. Z. WIN and A. ZANELLA, Error probability for optimum combining of  $M$ -ary PSK signals in the presence of interference and noise, *IEEE Transactions on Communications* 51:1949-1957 (2003).
- [48] R. J. MUIRHEAD, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc. 1982.
- [49] M. KANG, L. YANG and M.-S. ALOUINI, Capacity of MIMO Rician channels with multiple correlated Rayleigh co-channel interferers, *Proc. IEEE Global Communications Conference* 1119-1123 (2003).
- [50] G. SZEGÖ, *Orthogonal Polynomials*, 4th ed., in: American Mathematical Society Colloquium Publications, 23 American Mathematical Society, Providence, R. I., 1975.
- [51] A. MAGNUS, Painlevé-type differential equations for the recurrence coefficients of semiclassical orthogonal polynomials, *J. Comput. Appl. Math.* 57:215-237 (1995).
- [52] C. A. TRACY and H. WIDOM, Fredholm determinants, differential equations and matrix models, *Commun. Math. Phys.* 163:33-72 (1994).
- [53] P. J. FORRESTER and N. S. WITTE, The distribution of the first eigenvalue at the hard edge of the Laguerre unitary ensemble, *Kyushu Math. J.* 61:457-526 (2007).
- [54] P. J. FORRESTER and N. S. WITTE, Application of  $\tau$ -function theory of Painlevé equations to random matrices: PV, PIII, the LUE, JUE, and CUE, *Comm. Pure Appl. Math.* 55:679-727 (2002).

- [55] M. TODA, Theory of Non-Linear Lattices, Springer-Verlag, Berlin, 1989.
- [56] J. L. GAMMEL and J. NUTTALL, Note on generalized Jacobi polynomials, in The Riemann Problem, Complete Integrability and Arithmetic Applications, Bures-sur-Yvette/New York, Lecture Notes in Math., Vol. 925, pp 258–270, Berlin:Springer, 1979/80.
- [57] M. TSUJI, Poential Theory in Modern Function Theory, Maruzen, Tokyo, 1959.
- [58] F. D. GAKHOV, Boundary Value Problems, Pergamon, 1966.
- [59] S. G. MIKHLIN, Integral Equations and their Applications to Certain Problems in Mechanics, Mathematical Physics and Technology, Pergamon, 1957.
- [60] H. WIDOM, Lectures on Integral Equations, Van Nostrand Mathematical Studies 17, 1969.
- [61] N. I. MUSKHELISHVILI, Some Basic Problems of the Mathematical Theory of Elasticity, Kluwer, 1977.
- [62] H. WIDOM, Singular integral equations in  $L_P$ , Trans. Amer. Math. Soc. 97:131-160 (1960).
- [63] V. A. MARČENKO and L. A. PASTUR, Distribution of eigenvalues in certain sets of random matrices, Mat. Sb. (N.S.) 72:507-536.
- [64] F. J. DYSON, Distribution of eigenvalues for a class of real symmetric matrices, Rev. Mex. Fis. 20:231 (1971).
- [65] P. B. RAPAJIC and D. POPESCU, Information capacity of a random signature multiple-input multiple-output channel, IEEE Transactions on Communications 48:1245-1248 (2000).
- [66] J. SCHWINGER, Particle, Sources and Fields vol. 2, Advanced Books Classics, 1998.
- [67] L. ZHENG and D. N. C. TSE, Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels, IEEE Transactions on Information Theory 49:1073-1096 (2003).
- [68] I. S. GRADSHTEYN and I. M. RYZHIK, Table of Integrals, Series and Products, 7<sup>th</sup> ed. Elsevier, Singapore, 2007.