Moments of distributions via the S-transform
18.338 Project Report

James Chen

November 2025

1 Introduction

It is known that many common laws encountered in random matrix theory possess densities that have moments that can be described in a combinatorial fashion. Some examples include the Wigner semicircle law, and the Marchenko-Pastur distribution, which have moments corresponding to Catalan numbers and Narayana numbers respectively.

The purpose of this project is to highlight a simple observation regarding computing such moments, in terms of the S-transform in free probability. We will use it to calculate the moments of distributions that arise naturally in random matrices, including the Wachter law (which can be thought of as a free analogue of the Beta distribution), and products/inverses of Wishart matrices. The combinatorial relations of the moments will also be explored briefly.

Our results will mostly be applied to expressions involving Wishart matrices. Recall that if X is a random $m \times n$ matrix of independent standard (real or complex) Gaussians, then $W = \frac{1}{n}X^*X$ has the **Wishart** distribution, denoted by $W \sim \mathcal{W}_n(m)$. We will assume for simplicity that $m \geq n$ so that the matrix X^*X is nonsingular almost surely.

It is well known [3] that if in the limit $m, n \to \infty$ such that $\frac{m}{n} \to \lambda \geq 1$, then the empirical spectral distribution of $\frac{1}{n}X^*X$ converges (weakly almost surely) to the **Marchenko-Pastur** distribution μ_λ, defined by the density

$$\rho_\lambda(x) = \frac{\sqrt{(x-a)(b-x)}}{2\pi x} \qquad a = (1-\sqrt{\lambda})^2, b = (1+\sqrt{\lambda})^2 \tag{1.1}$$

This distribution will form the basis of the investigations given in this report.

2 Free probability transforms and Lagrange inversion

As a brief introduction to free probability we will give its definition and some remarks about free independence.

Definition 2.1. A C^* free probability space is a pair $(\mathcal{A}, \varphi)$ where $\mathcal{A}$ is a unital C^* algebra and φ is a positive linear functional on $\mathcal{A}$ with $\varphi(1_\mathcal{A}) = 1$. Elements of $\mathcal{A}$ are called random variables, and φ is the expectation functional.

Examples include the C^* algebra of random matrices of size $n \times n$, where φ is given by $\varphi(M) = \frac{1}{n}\mathbb{E}(\mathrm{tr}(M))$ i.e. the expected value of the normalised trace.

For a bounded self adjoint random variable $a \in \mathcal{A}$ there exists a probability measure μ compactly supported on the real line which we call the distribution of a. Let m_n be the moments of μ, which are defined by

$$m_n := \int x^n \,\mathrm{d}\mu(x) = \varphi(a^n) \tag{2.1}$$

In the case of random matrices, this μ corresponds to the eigenvalue distribution. We will only consider measures μ which are absolutely continuous and thus have a probability density ρ.

Remark 2.0.1. There is the fundamental concept of random variables being *freely independent*. We will not need to go into detail what this means, other than noting that large random matrices with independent entries tend to be freely independent as the size of the matrix n tends to infinity.

Using the moments m_n, we will define several power series.

Definition 2.2. Let a be a (self adjoint) bounded random variable in a free probability space with distribution μ. Let m_n be the moments of μ.
The moment series $M_\mu(z)$ and Cauchy transform $G_\mu(z)$ are defined by

$$M_\mu(z) := \sum_{n=1}^{\infty} m_n z^n \qquad G_\mu(z) := \int \frac{1}{z-x}\,\mathrm{d}\mu(x) \tag{2.2}$$

Due to the compact support of μ, the Cauchy transform has the Laurent series expansion at ∞ of the form

$$G_\mu(z) = \frac{1}{z} + \frac{m_1}{z^2} + \frac{m_2}{z^3} + \dots \tag{2.3}$$

We can then define the S-transform, which is fundamental in free probability as it linearises multiplication of freely independent non-commuting random variables. This enables us to calculate the eigenvalue distribution of products of random matrices, given that they are freely independent in the large matrix limit. For simplicity, we will only consider the S-transform of measures supported on $(0,\infty)$.

Remark 2.0.2. There is an analogous R-transform which linearises addition of freely independent random variables, but we will not use the R-transform in this report.

Definition 2.3. We define the S-transform of μ as the power series defined near 0 by

$$S_\mu(z) = \frac{z+1}{z} M_\mu^{-1}(z) \tag{2.4}$$

For freely independent random variables with distributions μ_a and μ_b supported on $(0,\infty)$, if we denote μ_{ab} to be the distribution of ab, then we have the relation [5]

$$S_{\mu_{ab}}(z) = S_{\mu_a}(z) S_{\mu_b}(z) \tag{2.5}$$

The above will be the only mention of free random variables and from this point on we will only work with the associated probability measures, giving interpretations in terms of random matrices. We will also drop the subscript when it is clear what distribution is in consideration.

Lemma 2.1 (Lagrange inversion)**.** *Let F be a formal power series in x satisfying the relation $F(x) = x\phi(F(x))$ for some analytic ϕ representable as a power series. Then we have for $n \geq 1$*

$$[x^n]F = \frac{1}{n}[x^{n-1}]\phi(x)^n \tag{2.6}$$

Theorem 2.2. *Let a be a random variable in a free probability space such that $\varphi(a) \neq 0$ and let S be the S-transform of a. Then we have for $n \geq 1$*

$$\varphi(a^n) = \frac{1}{n}[z^{n-1}]\left(\frac{z+1}{S(z)}\right)^n \tag{2.7}$$

Proof. Recall that $S(z) = \frac{z+1}{z}M^{-1}(z)$. Then by substituting z with $M(z)$, we obtain the relation

$$M(z) = z\frac{M(z)+1}{S(M(z))} = z\phi(M(z)) \qquad \phi(z) := \frac{z+1}{S(z)} \tag{2.8}$$

Lagrange inversion (Lemma 2.1) then yields the result. □

Example 2.1 (Marchenko Pastur moments)**.** The S-transform of the Marchenko pastur distribution with parameter λ is well known to be $S(z) = \frac{1}{z+\lambda}$. Thus, we can compute the moments as follows

$$m_n = \frac{1}{n}[z^{n-1}]((z+1)(z+\lambda))^n \tag{2.9}$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\left([z^k](z+1)^n\right)\left([z^{n-1-k}](z+\lambda)^n\right) \tag{2.10}$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\binom{n}{k}\binom{n}{n-1-k}\lambda^{k+1} = \frac{1}{n}\sum_{k=1}^{n}\binom{n}{k-1}\binom{n}{k}\lambda^k \tag{2.11}$$

which are exactly the Narayana polynomials $N_n(\lambda)$.

3 Free products and inverses of Marchenko-Pastur distributions

In the upcoming sections we will consider more complicated examples of expressions involving Wishart matrices and compute the moments of such random matrices using the S-transform.

3.1 Product of Wishart matrices

Consider N independent Wishart matrices $W_i \sim \mathcal{W}_n(m_i)$ for $i = 1, ..., N$. Since W_i are all random $n \times n$ matrices, we may consider the product

$$W = \prod_{i=1}^{N} W_i \tag{3.1}$$

Note that since the Wishart distribution is invariant under conjugation by Haar orthogonal/unitary matrices, we have that W_i are *asymptotically freely independent*. This means that in the limit as $n, m_1, ..., m_N \to \infty$ such that $\frac{m_i}{n} \to \lambda_i$, the limiting eigenvalue distribution of W which we denote μ is characterised by the S-transform:

$$S_\mu(z) = \prod_{i=1}^{N} S_{\mu_{\lambda_i}}(z) = \prod_{i=1}^{N} \frac{1}{z+\lambda_i} \tag{3.2}$$

We then may apply Theorem 2.2 to obtain that the moments of μ are given by

$$m_n = \frac{1}{n}[z^{n-1}][(z+1)(z+\lambda_1)...(z+\lambda_N)]^n \tag{3.3}$$

$$= \frac{1}{n}\sum_{1+l+k_1+,...,k_N=n}\binom{n}{l+1}\binom{n}{k_1},\ldots,\binom{n}{k_N}\lambda_1^{k_1},\ldots,\lambda_N^{k_N} \tag{3.4}$$

These can be recognised to be multivariate polynomials of $\lambda_1, \ldots, \lambda_N$ and are known as *multivariate Fuss-Narayana polynomials*, which are a generalisation of Narayana polynomials to N variables. This result on the moments can also be found in [2].

3.2 Inverted Wishart matrices

Given the Marchenko-Pastur distribution with parameter $\lambda > 1$, we may consider the pushforward of associated probability measure with respect to the map $x \to \frac{1}{x}$, which one can show is the limiting eigenvalue density (in the sense of weak convergence almost surely) of W^{-1}, where $W \sim \mathcal{W}_n(m)$ as $\frac{m}{n} \to \lambda$. We will denote this measure by ν_λ.
Since we are only working with probability densities, we can denote this inverse Marchenko-Pastur measure by the density

$$\psi_\lambda(x) = \frac{1}{x^2}\rho_\lambda\left(\frac{1}{x}\right) = \frac{\sqrt{\left(\frac{1}{x}-a\right)\left(b-\frac{1}{x}\right)}}{2\pi x} \tag{3.5}$$

Theorem 3.1. *The S-transform of the inverse Marchenko-Pastur distribution with parameter $\lambda > 1$ is given by*

$$S(z) = \lambda - 1 - z \tag{3.6}$$

The moments of the distribution have the formula

$$m_n = (\lambda - 1)^{-n} S_n\left(\frac{1}{\lambda - 1}\right) \tag{3.7}$$

where we have defined the polynomials $S_n(x)$ as

$$S_n(x) := \sum_{k=0}^{n-1} S_{n,k} x^k \tag{3.8}$$

and the combinatorial numbers $S_{n,k}$ for $n \geq 1$, $0 \leq k \leq n-1$ as

$$S_{n,k} := \frac{1}{n}\binom{n}{k+1}\binom{n+k-1}{k} \tag{3.9}$$

Proof. The Cauchy transform of ν_λ can be computed with a change of variable from the known Marchenko-Pastur Cauchy transform to obtain

$$G_{\nu_\lambda}(z) = \frac{1}{z} - \frac{1}{z^2} G_{\mu_\lambda}\left(\frac{1}{z}\right) = \frac{1}{z} - \frac{\frac{1}{z} + 1 - \lambda - \sqrt{\left(\frac{1}{z} - a\right)\left(\frac{1}{z} - b\right)}}{2z} \tag{3.10}$$

By using the identity $M(z) = \frac{1}{z} G\left(\frac{1}{z}\right) - 1$, we obtain that (with the correct branch of the square root so that $M(0) = 0$)

$$\begin{aligned} M(z) &= \frac{1}{z}\left(z - \frac{z + 1 - \lambda + \sqrt{(z-a)(z-b)}}{\frac{2}{z}}\right) - 1 && (3.11) \\ &= -\frac{z + 1 - \lambda + \sqrt{(z-a)(z-b)}}{2} && (3.12) \end{aligned}$$

Using the fact that $a + b = 2\lambda + 2$ and $ab = (\lambda - 1)^2$, we can obtain the equation

$$(-2M(z) - z - 1 + \lambda)^2 = (z-a)(z-b) \tag{3.13}$$

from which we obtain the inverse formula

$$M^{-1}(z) = \frac{ab - (\lambda - 2z - 1)^2}{a + b - 2\lambda + 4z - 2} = z\frac{\lambda - 1 - z}{z + 1} \tag{3.14}$$

Hence, the S-transform has the following simple expression, concluding the first part of the theorem:

$$S(z) = \lambda - 1 - z \tag{3.15}$$

Applying Theorem 2.2, we obtain that

$$\begin{aligned} m_n &= \frac{1}{n}[z^{n-1}]\left(\frac{1+z}{\lambda - 1 - z}\right)^n && (3.16) \\ &= \frac{1}{n}(\lambda - 1)^{-n}[z^{n-1}]\left[(1+z)^n \sum_{k=0}^{\infty}\binom{n+k-1}{k}\frac{z^k}{(\lambda-1)^k}\right] && (3.17) \\ &= \frac{1}{n}(\lambda - 1)^{-n}\sum_{k=0}^{n-1}\binom{n}{k+1}\binom{n+k-1}{k}\frac{1}{(\lambda-1)^k} && (3.18) \end{aligned}$$

which proves the second part.

∎

The numbers $S_{n,k}$ can be found in A088617. The notation $S_{n,k}$ for the numbers is due to the fact that they are related to the *large Schröder numbers*. The large Schröder numbers S_n are a sequence of integers analogous to the Catalan numbers, but instead of Dyck paths, they count paths that travel from $(0,0)$ to $(2n,0)$ using translating by $(1,1),(1,-1)$ plus an additional horizontal move $(2,0)$ and never pass below the x-axis. We will denote these as Schröder paths. The first few large Schröder numbers are $1,2,6,22,90,394,1806,8558$ and can be found in A006318.

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$n=1$	1							
$n=2$	1	1						
$n=3$	1	3	2					
$n=4$	1	6	10	5				
$n=5$	1	10	30	35	14			
$n=6$	1	15	70	140	126	42		
$n=7$	1	21	140	420	630	462	132	
$n=8$	1	28	252	1050	2310	2772	1716	429

Figure 1: Numbers $S_{n,k}$, whose rows sum to the large Schröder numbers. The $S_{n,k}$ entry counts the number of Schröder paths with k up-steps.

In Figure 1, we tabulate the integers $S_{n,k}$. We note that the sum of $S_{n,k}$ over all k is equal to S_n and $S_{n,k}$ count the number of Schröder paths that contain k up-steps, in a similar fashion to how Narayana numbers count the number of Dyck paths with additional constraints like number of valleys.

4 Combinatorics of Wachter moments

Using Wishart matrices, we can also define MANOVA matrices, named after multivariate analysis of variance from statistics. Consider two independently Wishart matrices $W_1 \sim \mathcal{W}_n(m_1)$ and $W_2 \sim \mathcal{W}_n(m_2)$. We define the matrix V

$$V = W_1(W_1 + W_2)^{-1} \tag{4.1}$$

Then we say that V has the **MANOVA** distribution, denoted by $V \sim \mathcal{M}_n(m_1, m_2)$ for convenience (though this is not standard notation). One may also consider the symmetric format $\tilde{V} = (W_1 + W_2)^{-\frac{1}{2}} W_1 (W_1 + W_2)^{-\frac{1}{2}}$ which exists as all matrices W_i are positive definite and have unique positive square root. It is clear that V and $\tilde{V}$ have the same eigenvalues, and that the eigenvalues of $\tilde{V}$ are real. Thus, they correspond to an empirical spectral distribution on the real line.
Just like in the case of the Marchenko-Pastur distribution, we have a similar result about the limiting distribution for large such matrices:

Proposition 4.1 ([6])**.** *Suppose that $m_1, m_2, n \to \infty$ such that $\frac{m_1}{n} \to a \geq 1$ and $\frac{m_2}{n} \to b \geq 1$. Then the empirical spectral distribution of V converges (weakly almost surely) to the **Wachter** distribution with parameters $a, b > 1$ which is given by the probability measure $\mu_{a,b}$ with density given by*

$$\rho_{a,b}(x) = (a+b)\frac{\sqrt{(x-\lambda_-)(\lambda_+ - x)}}{2\pi x(1-x)} \tag{4.2}$$

with $\lambda_\pm$ defined by

$$\lambda_\pm = \left(\sqrt{\frac{a}{a+b}\left(1 - \frac{1}{a+b}\right)} \pm \sqrt{\frac{1}{a+b}\left(1 - \frac{a}{a+b}\right)}\right)^2 \tag{4.3}$$

We now apply Theorem 2.2 to compute the moments of the Wachter distribution in a format that is analogous to the Marchenko-Pastur distribution.

Theorem 4.2. *The moments of the Wachter distribution with parameters $a, b > 0$ are given by*

$$m_n = \frac{1}{(a+b)^n} B_n\left(-\frac{a}{a+b}, a\right) \tag{4.4}$$

where $B_n(x,y)$ are polynomials with integer coefficients whose formula is given below:

$$B_n(x,y) := \sum_{k=0}^{n-1} \sum_{l=1}^{n-k} B_{n,k,l} x^k y^l \tag{4.5}$$

$$B_{n,k,l} := \frac{1}{n}\binom{n}{k+l}\binom{n}{l-1}\binom{n+k-1}{k} \tag{4.6}$$

Proof. By evaluating the Cauchy transform (we will not do this, since it is already in Alan's notes, but it can be done using contour integration), one can compute the S-transform of the Wachter distribution as

$$S(z) = \frac{z+a+b}{z+a} \tag{4.7}$$

Thus, by Theorem 2.2, the moments are given by

$$m_n = \frac{1}{n}[z^{n-1}]\left(\frac{(z+1)(z+a)}{z+a+b}\right)^n \tag{4.8}$$

We recall the power series expansion of $\frac{1}{(z+c)^n}$ as follows:

$$\frac{1}{(z+c)^n} = c^{-n}\sum_{k=0}^{\infty}\binom{n+k-1}{k}\frac{(-1)^k}{c^k}z^k \tag{4.9}$$

Using this, we can compute a closed form for the moments as follows:

$$m_n = \frac{1}{n}[z^{n-1}]\left[((z+1)(z+a))^n (a+b)^{-n}\sum_{k=0}^{\infty}\binom{n+k-1}{k}\frac{(-1)^k}{(a+b)^k}z^k\right] \tag{4.10}$$

$$= \frac{1}{n}(a+b)^{-n}\sum_{k=0}^{n-1}\left[\binom{n+k-1}{k}\frac{(-1)^k}{(a+b)^k}[z^{n-1-k}]((z+1)(z+a))^n\right] \tag{4.11}$$

$$= \frac{1}{n}(a+b)^{-n}\sum_{k=0}^{n-1}\binom{n+k-1}{k}\frac{(-1)^k}{(a+b)^k}a^k\sum_{l=1}^{n-k}\binom{n}{l+1}\binom{n}{k+l}a^l \tag{4.12}$$

from which the result holds. ☐

4.1 Combinatorial interpretations of Wachter moments

The form of Theorem 4.2 is written in a deliberate way to emphasise the polynomial structure of the moments, much like in the case of the Marchenko-Pastur distribution, which can be written as Narayana polynomials. It is natural to ask whether the numbers $B_{n,k,l}$ have a natural combinatorial interpretation. In Figure 2, we compute the first few such numbers and remark several patterns that can be seen, in increasing order of complexity.

- The Catalan numbers can be obtained on one edge of the pyramid i.e. $B_{n,n-1,0} = C_n$.
- The triangle formed by setting $k = 0$ yields the Narayana numbers i.e. $B_{n,0,l} = N_{n,l}$. This is clearly evident from the formula $B_{n,k,l}$ and as such, $B_{n,k,l}$ are a generalisation of Narayana numbers.
- The triangle formed by setting $l = 0$ is A088617. The row sums of this triangle are the large Schröder numbers, which have already been encountered in the previous section and Theorem 3.1.

$n=1$						1					
$n=2$					1		1				
						1					
$n=3$				1		3		1			
					3		3				
						2					
$n=4$			1		6		6		1		
				6		16		6			
					10		10				
						5					
$n=5$		1		10		20		10		1	
			10		50		50		10		
				30		75		30			
					35		35				
						14					
$n=6$	1		15		50		50		15		1
		15		120		225		120		15	
			70		315		315		70		
				140		336		140			
					126		126				
						42					

Figure 2: Wachter moment polynomial entries up to $n = 6$, where $k = 0, \ldots, n-1$ indicates the row within each triangular layer and $l = 1, \ldots, n-k$ indicates the entry in the current row.

- The triangle formed by summing over all l is A234950, known as the Borel triangle (despite never being studied by Borel, in fact this sequence was only mentioned rather recently [1][4]).

We will focus on combinatorial interpretations in terms of *marked Dyck paths*. For convenience, we recall the various terms to describe features of Dyck paths.

Definition 4.1 (Dyck Paths)**.** A Dyck path is a sequence of moves in $\mathbb{Z}^2$ from $(0,0)$ to $(2n,0)$ consisting of only moves of translating by $(1,1)$ or $(1,-1)$, such that the path never travels below the x-axis. It is convenient to denote these moves by letters U and D, so that a Dyck path may be interpreted as a sequence of letters.
We say that a move is at ground level if it starts or ends at a point with y-coordinate 0. As such, the first and last moves are always at ground level.
Given a Dyck path, we define a valley to be a pair of consecutive moves of the form DU. Figure 3 shows an example of such a path, ignoring the extra lines on some moves for now.

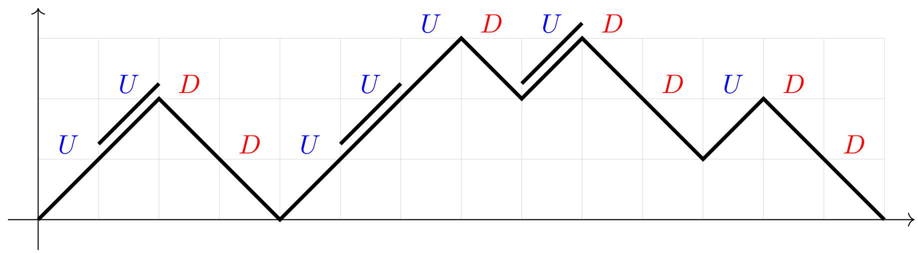

Figure 3: Dyck path of semi-length 7 with labeled moves. There are a total of 4 moves at ground level and 3 valleys in total. In addition, 3 up-steps have been marked.

It is well known that the Narayana numbers $N_{n,k}$ count the number of Dyck paths of semi-length n with $k-1$ valleys.

Definition 4.2 (Marked Dyck path)**.** Let P be a Dyck path of semi-length n. An *edge-marked Dyck path* is a pair (P, S) where S is a subset of moves of P. In Figure 3, we see an example of a marked Dyck path.

In [1], several combinatorial interpretations of the Borel triangle were proven in terms of *marked Catalan structures*, which are structures enumerated by Catalan numbers e.g. Dyck paths or rooted trees, with specific features distinguished.

Theorem 4.3 (Theorem 2 in [1])**.** *Let $B_{n,k}$ be the entries of the Borel triangle, starting with $n = 1$. Then $B_{n,k}$ of Borel's triangle counts the set of pairs (P, S) where P is a Dyck path of semi-length n and S consists of k up-steps of P, none of which is at ground level.*

Inspired by this, we prove the first combinatorial interpretation of the numbers $B_{n,k,l}$.

Theorem 4.4. *For $n \geq 1$, $0 \leq k \leq n-1$ and $0 \leq l \leq n-k-1$, the numbers $B_{n,k,l+1}$ count the number of marked Dyck paths of semi-length n with k marked up-steps not at ground level and l valleys that do not contain a marked up-step.*

As an example, in Figure 3, the marked Dyck path is of semi-length 7 with 3 marked up-steps not at ground level and 2 valleys that do not contain a marked up-step.

Proof of Theorem 4.4. Let $f_{n,k,l}$ be the quantity stated in the theorem and $F(x, y, z)$ be the corresponding generating function defined by

$$F(x, y, z) = \sum_{n,k,l\geq 0} f_{n,k,l} x^n y^k z^l \tag{4.13}$$

It is well known that a non-degenerate Dyck path P may be uniquely decompose into $P = UP_1DP_2$ where P_1 and P_2 are Dyck paths. We will refine this decomposition for Dyck path with marked up-steps not at ground level. Given a non-degenerate Dyck path with marked up-steps not at ground level, we may uniquely decompose it as

$$P = UP_1DP_2 \qquad P_1 = Q_0Q_1Q_2, \dots Q_n \tag{4.14}$$

where P_2 and Q_0 are (possibly degenerate) Dyck paths with no marked up-steps at ground level and $Q_i, i \geq 1$ are non-degenerate marked Dyck paths with no up-steps at ground level except the first up-step, which is always marked.
Denote the number of valleys in a marked Dyck path P without a marked up-step $V(P)$. We use the decomposition $P = UQ_0, ..., Q_nDP_2$ to determine contributions to the number of valleys without a marked up-step.

- Each term Q_i, $i = 0, ..., n$ and P_2 contribute $V(Q_i)$ and $V(P_2)$ valleys without a marked up-step respectively.
- Since each Q_i for $i \geq 1$ begins with a marked up-step, there are no contributions from the valleys in between consecutive Q_iQ_{i+1} terms.

- An additional valley without a marked up-step is created between the D and P_2 terms if and only if P_2 is non-degenerate.

Hence the total contributions can be written in the form

$$V(P) = V(P_2) + \mathbf{1}_{\text{length}(P_2)>0} + \sum_{i=0}^{n} V(Q_i) \tag{4.15}$$

We now formulate the implicit relation the generating function F satisfies.

- The UD pair increases the semi-length by 1 and nothing else, hence contributes an x term.
- The Q_0 contributes an F term and P_2 contributes an $z(F-1)+1$ term due to (4.15).
- The Q_i for $i \geq 1$ contribute a geometric series term of the form

$$1 + y(F-1) + (y(F-1))^2 + \cdots = \frac{1}{1-y(F-1)} \tag{4.16}$$

Hence, we obtain the implicit relation

$$F = 1 + xF\frac{z(F-1)+1}{1-y(F-1)} \tag{4.17}$$

Denoting $\tilde{F} = F - 1$, we may write this in the form

$$\tilde{F} = x(\tilde{F}+1)\frac{z\tilde{F}+1}{1-y\tilde{F}} \tag{4.18}$$

Lagrange inversion (Lemma 2.1) then tells us that

$$[x^n]\tilde{F} = \frac{1}{n}[x^{n-1}]\left(\frac{(x+1)(zx+1)}{1-yx}\right)^n \tag{4.19}$$

$$= \frac{1}{n}[x^{n-1}]\left((x+1)^n(zx+1)^n\sum_{k=0}^{\infty}\binom{n+k-1}{k}y^k x^k\right) \tag{4.20}$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\binom{n+k-1}{k}y^k\sum_{l=0}^{n-1-k}\binom{n}{n-1-k-l}\binom{n}{l}z^l = \sum_{k=0}^{n-1}\sum_{l=0}^{n-1-k}B_{n,k,l+1}y^k z^l \tag{4.21}$$

which yields the result. □

References

[1] Yue Cai and Catherine Yan. Counting with borel's triangle. *Discrete Mathematics*, 342(2):529–539, 2019.

[2] Romuald Lenczewski and Rafal Salapata. Multivariate fuss-narayana polynomials and their application to random matrices, 2012.

[3] V A Marčenko and L A Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, apr 1967.

[4] Jeffrey B. Remmel. Consecutive up–down patterns in up–down permutations. *Electronic Journal of Combinatorics*, 21(3):P3.2, 2014.

[5] Dan Voiculescu. Multiplication of certain non-commuting random variables. *Journal of Operator Theory*, 18(2):223–235, 1987.

[6] Kenneth W. Wachter. The strong limits of random matrix spectra for sample matrices of independent elements. *The Annals of Probability*, 6(1):1–18, 1978.