

RESTRICTED MAXIMUM OF NON-INTERSECTING BROWNIAN BRIDGES

YAMIT YALANDA^{1,*}  AND NICOLÁS ZALDUENDO² 

Abstract. Consider a system of N non-intersecting Brownian bridges in $[0, 1]$, and let $\mathcal{M}_N(p)$ be the maximal height attained by the top path in the interval $[0, p]$, $p \in [0, 1]$. It is known that, under a suitable rescaling, the distribution of $\mathcal{M}_N(p)$ converges, as $N \rightarrow \infty$, to a one-parameter family of distributions interpolating between the Tracy-Widom distributions for the Gaussian Orthogonal and Unitary Ensembles (corresponding, respectively, to $p \rightarrow 1$ and $p \rightarrow 0$). It is also known that, for fixed N , $\mathcal{M}_N(1)$ is distributed as the top eigenvalue of a random matrix drawn from the Laguerre Orthogonal Ensemble. Here we show a version of these results for $\mathcal{M}_N(p)$ for fixed N , showing that $\mathcal{M}_N(p)/\sqrt{p}$ converges in distribution, as $p \rightarrow 0$, to the rightmost charge in a generalized Laguerre Unitary Ensemble, which coincides with the top eigenvalue of a random matrix drawn from the Antisymmetric Gaussian Ensemble.

Mathematics Subject Classification. 60K35, 60B20, 60J65.

Received October 24, 2023. Accepted April 1, 2024.

1. INTRODUCTION AND MAIN RESULTS

1.1. Model and motivation

The model of non-intersecting Brownian bridges consists in a collection of N Brownian bridges $(B_1(t), B_2(t), \dots, B_N(t))$, all starting from zero at time $t = 0$ and ending at zero at time $t = 1$, and conditioned (in the sense of Doob) to not intersect for $t \in (0, 1)$. We will always order the N paths increasingly, so that $B_1(t) < B_2(t) < \dots < B_N(t)$. This model and its many variants have attracted a lot of attention in the probability and statistical physics literature, in large part due to their connections with Random Matrix Theory (RMT). The most basic relation with RMT is that under a simple change of variables, non-intersecting Brownian bridges are mapped to the stationary Dyson Brownian motion for the Gaussian Unitary Ensemble (GUE), which is the evolution of the eigenvalues of an $N \times N$ Gaussian Hermitian matrix whose entries undergo independent (stationary) Ornstein-Uhlenbeck processes (see (2.4)). In particular, for fixed $t \in (0, 1)$ the distribution of $(B_1(t), \dots, B_N(t))$ coincides with the distribution of the GUE eigenvalues, which also means that the properly rescaled top position $B_N(t)$ converges to the Tracy-Widom GUE distribution F_{GUE} [1] which governs

Keywords and phrases: Non-intersecting Brownian bridges, KPZ universality class, random matrices, Airy2 process, antisymmetric Gaussian ensemble.

¹ Departamento de Ingeniería Matemática, Universidad de Chile, Santiago, Chile.

² Université de Lorraine, CNRS, Inria, IECL, UMR 7502, 54000 Nancy, France.

* Corresponding author: yamit@dim.uchile.cl

the asymptotic fluctuations of the top eigenvalue of a GUE matrix:

$$\mathbb{P}\left(\frac{B_N(t)}{\sqrt{2t(1-t)}} \leq \sqrt{2N} + \frac{1}{\sqrt{2}}N^{-1/6}r\right) \xrightarrow{N \rightarrow \infty} F_{\text{GUE}}(r). \tag{1.1}$$

Models of non-intersecting random paths also arise in the study of the Kardar-Parisi-Zhang (KPZ) universality class, a broad collection of one-dimensional random growth and related models which share a common fluctuation behavior after suitable space-time rescaling, governed by the KPZ fixed point [2]. The limiting distributions and spatial processes arising asymptotically for KPZ models depend on the choice of initial data. The simplest one corresponds to growth from a single point, referred to as droplet or narrow wedge initial condition, in which case the spatial fluctuations of the model converge to the Airy_2 process, $\mathcal{A}_2(x)$, minus a parabola, whose one point marginals are given by the TW-GUE distribution. The same Airy_2 process arises as the scaling limit of the top path of N non-intersecting Brownian bridges:

$$2N^{1/6}\left(B_N\left(\frac{1}{2}(1 + N^{-1/3}x)\right) - \sqrt{N}\right) \xrightarrow{N \rightarrow \infty} \mathcal{A}_2(x) - x^2, \tag{1.2}$$

in the sense of convergence in distribution in the topology of uniform convergence on compact sets. On the other hand, a famous result by Johansson [3] showed, using an argument based on a connection with KPZ models with flat initial data, that $\sup_{x \in \mathbb{R}} (\mathcal{A}_2(x) - x^2)$ is distributed as a Tracy-Widom GOE random variable [4], corresponding to the asymptotic fluctuations of the top eigenvalue of a Gaussian real symmetric random matrix. Putting these two facts together one concludes that the rescaled maximal height of $B_N(t)$ converges in distribution to a TW-GOE random variable: more precisely,

$$\mathbb{P}\left(2N^{1/6}\left(\max_{t \in [0,1]} B_N(t) - \sqrt{N}\right) \leq r\right) \xrightarrow{N \rightarrow \infty} F_{\text{GOE}}(4^{1/3}r), \tag{1.3}$$

a fact which by now has been proved in several ways in the literature (see for example [5, 6]). We refer the reader to [7] and to the introductions of [5, 8] and references therein for more background on the facts being discussed here. The relation between models of non-intersecting Brownian motions and other objects coming from RMT, integrable systems and the KPZ class has been studied intensively from many perspectives, see *e.g.* [9–18].

Consider now the maximal height of the top path restricted to an interval $[0, p]$, $p \in [0, 1]$:

$$\mathcal{M}_N(p) = \max_{t \in [0,p]} B_N(t). \tag{1.4}$$

Based on the same connection (1.2) between non-intersecting Brownian bridges and the Airy_2 process, together with known KPZ results, the following distributional limit for $\mathcal{M}_N(p)$ can be derived: letting $p(\alpha) = e^{2\alpha}/(1 + e^{2\alpha})$ we have, for each $\alpha \in \mathbb{R}$,

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(2N^{1/6}(\mathcal{M}_N(p(\alpha N^{-1/3})) \cosh(\min\{\alpha N^{-\frac{1}{3}}, 0\}) - \sqrt{N}) \leq r\right) = F_{2 \rightarrow 1}^{(\alpha)}(r) \tag{1.5a}$$

$$\text{with } F_{2 \rightarrow 1}^{(\alpha)}(r) \xrightarrow{\alpha \rightarrow \infty} F_{\text{GOE}}(4^{1/3}r), \quad F_{2 \rightarrow 1}^{(\alpha)}(r) \xrightarrow{\alpha \rightarrow -\infty} F_{\text{GUE}}(r). \tag{1.5b}$$

More concretely, $F_{2 \rightarrow 1}^{(\alpha)}$ corresponds to the marginal distribution of the $\text{Airy}_{2 \rightarrow 1}$ process arising as the scaling limit of KPZ models with half-flat initial data [19], whose distribution is known to interpolate between TW-GOE and TW-GUE. The $\alpha \rightarrow \infty$ limit to TW-GOE corresponds to (1.3). For the other case, $\alpha \rightarrow -\infty$ corresponds to the maximum in (1.4) being computed in an interval $[0, p]$ with $p \rightarrow 0$. For large N the top path is known to concentrate around the curve $2\sqrt{Nt(1-t)}$, and thus to first order one expects the maximum on $[0, p]$ to occur

near p as $p \rightarrow 0$ and to be close to $2\sqrt{Np(1-p)}$. The hyperbolic cosine factor in (1.5a) compensates for this decay, and the fluctuations then come from the value of B_N at the right edge of the interval, which has TW-GUE fluctuations. The derivation of (1.5a) is by now relatively standard. There are two possible routes. The first one is to use a version of (1.2) where the convergence holds uniformly in compact sets [20] together with estimates on the tails of the Airy₂ process to show that the limit equals $\mathbb{P}(\sup_{x \leq \alpha} (\mathcal{A}_2(x) - x^2) + \min\{\alpha, 0\}^2 \leq r)$, which was shown in [21] to be given by the marginals of the Airy_{2→1} process. The other one involves calculating the left hand side directly and checking that it converges to the known expressions for $F_{2 \rightarrow 1}^{(\alpha)}$ (this can be done using the formula derived Prop. 2.1 below, but we omit the details).

The focus of this paper is the random variable $\mathcal{M}_N(p)$ for fixed $N \in \mathbb{N}$. In [5] it was shown that in the case $p = 1$ (*i.e.* when we look at the maximal height of the system of non-intersecting Brownian bridges over the whole interval $[0, 1]$), the square of $\mathcal{M}_N(1)$ is distributed as the largest eigenvalue of a random matrix drawn from the Laguerre Orthogonal Ensemble, which provides a finite N version of (1.3).

We will be interested in the other limiting case, *i.e.* the distribution of $\mathcal{M}_N(p)$ (after proper rescaling) as $p \rightarrow 0$. Some results available in the literature strongly suggest a connection between this distribution and that of the largest eigenvalue of a different random matrix family (see Sect. 1.3 for more details). The aim of this paper is to prove this connection and identify this limiting distribution which, as we prove, corresponds to the largest eigenvalue of a matrix of the Gaussian Antisymmetric Ensemble. To arrive at this result, we will exploit the relation between this family of matrices and the Generalized Laguerre Unitary Ensemble (see Thm. 1.1 below).

1.2. Main result

Let X be an $N \times m$ matrix, $m \geq N$, with independent real or complex standard Gaussian entries (in the complex case, the real and imaginary parts of each entry are independent with variance $1/2$). The $N \times N$ matrix $M = XX^*$ is sometimes referred to as a real or complex Wishart matrix, and plays a central role in multivariate statistics as the sample covariance matrix of a Gaussian population (see [22], Sect. 3.2). The eigenvalues $\lambda_1 \leq \dots \leq \lambda_N$, and particularly the largest eigenvalues of Wishart matrices, are of particular importance in statistical applications such as principal component analysis. The joint distribution of these eigenvalues can be computed explicitly, and is given by (see [22], p. 91)

$$\frac{1}{Z_N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^N \lambda_i^{\frac{\beta a}{2}} e^{-\beta \frac{\lambda_i}{2}} \tag{1.6}$$

with $\beta = 1$ in the real case, $\beta = 2$ in the complex case and $a = m - N + 1 - \frac{2}{\beta}$ (Z_N is a normalization constant).

The weights $\lambda_i^{\frac{\beta a}{2}} e^{-\beta \frac{\lambda_i}{2}}$ are associated to the generalized Laguerre orthogonal polynomials, and based on this the random matrix M is said to belong to the Laguerre Orthogonal Ensemble in the real case and to the Laguerre Unitary Ensemble in the complex case (the orthogonal and unitary names coming from the fact that the distribution of M is invariant under conjugation by fixed matrices from, respectively, the groups $O(N)$ and $U(N)$).

The distribution of the eigenvalues of an $N \times N$ LOE or LUE matrix depends on the parameter a , which is determined by the aspect ratio of the matrix X . Let $\bar{\lambda}_{L(O/U)E,a}$ denote the largest eigenvalue of such a matrix and let $F_{L(O/U)E,a}$ denote its distribution, *i.e.*

$$F_{L(O/U)E,N}^{(a)}(r) = \mathbb{P}(\bar{\lambda}_{L(O/U)E,a} \leq r). \tag{1.7}$$

Then [5] showed that

$$\mathbb{P}(\mathcal{M}_N(1) \leq r) = F_{LOE,N}^{(0)}(4r^2), \tag{1.8}$$

recall that $a = 0$ and $\beta = 1$ in this case. In other words, $4 \max_{t \in [0,1]} B_N(t)^2$ is distributed as the largest eigenvalue of an LOE matrix $M = XX^T$ with X of size $N \times (N + 1)$.

The largest eigenvalue of an LOE matrix is known to converge to a TW-GOE random variable [23] (more precisely, in our case we have $\lim_{N \rightarrow \infty} F_{\text{LOE},N}^{(1)}(4N + 4^{2/3}N^{1/3}r) = F_{\text{GOE}}(r)$), and hence (1.8) can indeed be regarded as a finite N analog of the first limit in (1.5b). It is natural then to wonder about a version of the second limit in (1.5b), corresponding to studying $\mathcal{M}_N(p)$ with $p \rightarrow 0$.

Of course, as $p \rightarrow 0$ the maximum of $B_N(t)$ for $t \in [0, p]$ goes to 0 too, so in order to see something interesting we need to rescale $\mathcal{M}_N(p)$ before taking the limit. $B_N(t)$ grows like \sqrt{t} for small t , so we will study the distributional limit

$$\hat{\mathcal{M}}_N(0) := \lim_{p \rightarrow 0} \frac{\mathcal{M}_N(p)}{\sqrt{p}}. \tag{1.9}$$

In view of (1.5) and (1.8), a natural guess is that this limit be related with the Laguerre Unitary Ensemble. Our main result, which we state next, confirms this, with the caveat that the parameter a has to be allowed to take non-integer values. Note that (1.6) defines a probability density for the vector $(\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N$ for any $a > -1$. In general the λ_i 's do not arise as eigenvalues of a naturally defined random matrix, and we think of (1.6) as defining a point process, or a Coulomb gas, in the real line. In any case, $F_{\text{L}(\text{O}/\text{U})\text{E},N}^{(a)}$ still makes sense as the distribution of the largest charge (largest particle in the point process) $\bar{\lambda}_{\text{L}(\text{O}/\text{U})\text{E},a}$ among the λ_i 's.

Theorem 1.1. *For every $N \in \mathbb{N}$ and $r \geq 0$,*

$$\lim_{p \rightarrow 0} \mathbb{P}(\mathcal{M}_N(p) \leq \sqrt{p}r) = \begin{cases} F_{\text{LUE},N/2}^{(1/2)}(r^2/2) & \text{if } N \text{ is even,} \\ F_{\text{LUE},(N+1)/2}^{(-1/2)}(r^2/2) & \text{if } N \text{ is odd.} \end{cases} \tag{1.10}$$

In other words, the distributional limit $\hat{\mathcal{M}}_N(0)$ in (1.9) is well defined, and $\hat{\mathcal{M}}_N(0)^2/2$ has the distribution of the largest charge in the generalized Laguerre Unitary Ensemble defined through (1.6) with size $\lfloor (N + 1)/2 \rfloor$, $\beta = 2$ and $a = \frac{1}{2}(-1)^N$.

Together, the identities (1.8) and (1.10) provide an analog of (1.5), which can be stated as follows:

$$\hat{F}_N^p(r) := \mathbb{P}(\mathcal{M}_N(p) \leq \sqrt{p}r) \longrightarrow \begin{cases} F_{\text{LOE},N}^{(0)}(4r^2) & \text{as } p \rightarrow 1, \\ F_{\text{LUE},\lfloor (N+1)/2 \rfloor}^{(\frac{1}{2}(-1)^N)}(r^2/2) & \text{as } p \rightarrow 0. \end{cases} \tag{1.11}$$

From results of [24] and [23] it is known that, under the scaling implied by (1.5), the LOE and LUE distributions on the right hand side of (1.11) converge to TW-GOE and TW-GUE, so these limits are consistent with (1.5b) (see the discussion after Theorem 1.2 in [5] for more details).

1.3. Matrix model for $\hat{\mathcal{M}}_0$

We have obtained the distributions $F_{\text{LUE},N}^{(\pm 1/2)}$ appearing in Theorem 1.1 as those of the largest charge in a generalized Laguerre Unitary Ensemble. This ensemble appears not to be related to some sort of generalized Wishart matrix but, remarkably, there is another simple random matrix model whose eigenvalues recover this distribution.

Given $n \in \mathbb{N}$, a $n \times n$ (purely imaginary) random matrix H is said to be drawn from the Antisymmetric Gaussian Ensemble if $H = \frac{i}{\sqrt{2}}(X - X^T)$ for X a real random matrix with independent standard Gaussian entries. The non-zero eigenvalues of this matrix come in pairs $\{\pm \lambda_j\}_{j=1, \dots, \lfloor n/2 \rfloor}$, and the joint density function for the

positive eigenvalues is given by (see [25–27])

$$\frac{1}{Z_n} \prod_{i=1}^{\lfloor n/2 \rfloor} \lambda_i^{1+(-1)^{n+1}} e^{-\lambda_i^2/2} \prod_{1 \leq j < k \leq \lfloor n/2 \rfloor} (\lambda_i^2 - \lambda_j^2)^2, \tag{1.12}$$

where Z_n is a normalization constant. Applying the change of variables $\lambda_j^2/2 \mapsto \lambda_j$ this joint density becomes that of the generalized LUE with $N = n - 1$ and $a = \frac{1}{2}(-1)^n$. In other words, if $\bar{\lambda}_{\text{AntiGE},n}$ denotes the largest eigenvalue of an $n \times n$ matrix from the Antisymmetric Gaussian Ensemble, then $\bar{\lambda}_{\text{AntiGE},N+1}^2 = 2\lambda_{\text{LUE},N}$, for $\lambda_{\text{LUE},N}$ the distribution of the largest charge in the generalized Laguerre Unitary Ensemble, where the equality holds in distribution.

Hence, this shows that

$$\hat{\mathcal{M}}_N(0) \stackrel{\text{dist}}{=} \bar{\lambda}_{\text{AntiGE},N+1}. \tag{1.13}$$

Our main motivation in this work was to study $\hat{\mathcal{M}}_N(0)$ by using the continuum statistics techniques, first employed in [28] in the context of the Airy₂ process, to obtain an explicit Fredholm determinant formula for the distribution of $\mathcal{M}_N(p)$ (see Prop. 2.1 below) and then investigating its limit. The last identity is also connected to earlier results in the literature which provide a different avenue for its proof. Although we originally thought that some results were missing for such a strategy to work, we learned that the necessary identities are known, which lead to a shorter, although rather indirect, proof of the result. We sketch it next.

Our model can be expressed in terms of GUE Dyson Brownian motion by employing a time change between Hermitian Brownian motions and bridges, which explicitly yields $\mathbb{P}(\mathcal{M}_N(p) \leq \sqrt{pr})$ equals to

$$\mathbb{P}(B_N(t) \leq \sqrt{pr} \text{ for all } t \in [0, p]) = \mathbb{P}((1-t)X_N(H(t/(1-t))) \leq \sqrt{pr} \text{ for all } t \in [0, p]),$$

where $X_N(H(t))$ is the largest eigenvalue of the GUE matrix $H(t)$ with entries undergoing (complex) Brownian motions (see [6]). If N is even (the odd N case is analogous), the result from [29] yields

$$\mathbb{P}\left(X_{N/2}^{(C)}\left(\frac{p}{1-p}\right) \leq \sqrt{pr}\right) \leq \mathbb{P}(\mathcal{M}_N(p) \leq \sqrt{pr}) \leq \mathbb{P}\left(X_{N/2}^{(C)}\left(\frac{p}{1-p}\right) \leq \frac{\sqrt{p}}{1-p}r\right),$$

where $X_1^{(C)}(t), \dots, X_{N/2}^{(C)}(t)$ is a Dyson Brownian motion of type C , which are the eigenvalues of an AntiGE matrix $J(t)$ with entries undergoing Brownian motions (see [27]).

2. FREDHOLM DETERMINANT FORMULA FOR THE DISTRIBUTION OF THE RESTRICTED MAXIMAL HEIGHT

The goal of this section is to derive Fredholm determinant formulas for the distribution of the random variables $\hat{\mathcal{M}}_N(0)$ and $\mathcal{M}_N(p)$. Let φ_n be the harmonic oscillator functions, or Hermite functions, defined by $\varphi_n(x) = e^{-x^2/2}p_n(x)$, with p_n the n -th Hermite polynomial (see *e.g.* [30], Sect. 18.3) normalized so that $\|\varphi_n\|_2 = 1$, and then define the Hermite kernel as

$$K_H^N(x, y) = \sum_{n=0}^{N-1} \varphi_n(x)\varphi_n(y). \tag{2.1}$$

We also define the reflection operator ϱ_r as well as projection and multiplication operators χ_r and E_r , acting on $L^2(\mathbb{R})$, through

$$\varrho_r f(x) = f(2r - x), \quad \chi_r f(x) = \mathbb{1}_{x>r} f(x), \quad \text{and} \quad E_r f(x) = e^{rx} f(x). \tag{2.2}$$

Let also $\bar{\chi}_r = I - \chi_r$.

Proposition 2.1. *Consider $N \in \mathbb{N}$ and $r \geq 0$. Fix $p \in (0, 1)$ and let $\alpha = \frac{1}{2} \log(\frac{p}{1-p})$. Then,*

$$\begin{aligned} \mathbb{P}(\sqrt{2}\mathcal{M}_N(p) \leq r) &= \det(I - K_H^N \chi_{r \cosh(\alpha)} K_H^N - K_H^N E_{r \sinh(\alpha)} \varrho_{r \cosh(\alpha)} E_{-r \sinh(\alpha)} \bar{\chi}_{r \cosh(\alpha)} K_H^N), \end{aligned} \tag{2.3}$$

where \det means the Fredholm determinant on the Hilbert space $L^2(\mathbb{R})$.

For the definition and properties of the Fredholm determinant we refer the reader to [31] or [7], Section 2.

Proof. It turns out to be convenient to express $\mathcal{M}_N(p)$ in terms of the stationary (GUE) Dyson Brownian motion, for which formulas are cleaner. This is the process which describes the evolution of the eigenvalues of an $N \times N$ complex Hermitian matrix whose entries evolve as independent stationary Ornstein-Uhlenbeck processes. In other words, each entry (both the real and the imaginary part) follows an independent copy of the solution of the SDE $dX_t = -X_t dt + \sigma dB_t$, where B is a standard Brownian motion with $\sigma = 1/\sqrt{2}$ for off-diagonal entries and $\sigma = 1$ on the diagonal. For each fixed t , the ordered eigenvalues $\lambda_1(t) \leq \dots \leq \lambda_N(t)$ are distributed, up to scaling, as the eigenvalues of an $N \times N$ GUE matrix, and if we let this stationary process be defined for $t \in \mathbb{R}$, then one has (see e.g. [17])

$$(B_i(s))_{i=1,\dots,N} \stackrel{\text{dist}}{=} \left(\sqrt{2s(1-s)} \lambda_i \left(\frac{1}{2} \log \left(\frac{s}{1-s} \right) \right) \right)_{i=1,\dots,N} \tag{2.4}$$

as processes defined for $s \in [0, 1]$. Thus, applying the change of variables $s \rightarrow \frac{e^{2t}}{e^{2t}+1}$ we have $\sqrt{2}\mathcal{M}_N(p) = \sup_{t \in (-\infty, \alpha]} \frac{\lambda_N(t)}{\cosh(t)}$ with α as in the statement of the result. Hence

$$\mathbb{P}(\sqrt{2}\mathcal{M}_N(p) \leq r) = \mathbb{P}(\lambda_N(t) \leq r \cosh(t) \ \forall t \leq \alpha). \tag{2.5}$$

We will focus now on the random variable on the right hand side.

Let D denote the operator $Df(x) = -\frac{1}{2}(f''(x) - (x^2 - 1)f(x))$. The Hermite functions φ_n are eigenfunctions for D , $D\varphi_n = n\varphi_n$ for $n \geq 0$, so that that K_H^N is the projection operator onto the space $\text{span}\{\varphi_0, \dots, \varphi_{N-1}\}$ associated to the first N eigenvalues of D . In particular, $e^{tD} K_H^N$ is well defined for all t , with integral kernel is given by

$$e^{tD} K_H^N(x, y) = \sum_{n=0}^{N-1} e^{tn} \varphi_n(x) \varphi_n(y). \tag{2.6}$$

It was shown in [32] that

$$\begin{aligned} \mathbb{P}(\lambda_N(t) \leq r \cosh(t) \ \forall t \in [-L, \alpha]) &= \det(I - K_H^N + \Theta_{L,\alpha}^r e^{(\alpha+L)D} K_H^N) \\ &= \det(I - K_H^N + e^{(\alpha+L)D} K_H^N \Theta_{L,\alpha}^r K_H^N) \end{aligned} \tag{2.7}$$

where $\Theta_{L,\alpha}^r$ is the solution operator for a certain boundary value problem involving D (see [5], Prop. 2.1 and Eq. (2.6) therein), and where in the second equality we used the facts that $e^{sD} K_H^N = K_H^N e^{sD} K_H^N$ and

$K_H^N = e^{sD} K_H^N e^{-sD} K_H^N$ together with the cyclic property of the determinant. We use now the decomposition for $\Theta_{L,\alpha}^r$ given in [5]:

$$\Theta_{L,\alpha}^r = e^{-(\alpha+L)D} - e^{-(\alpha+L)D} \chi_r \cosh(\alpha) - R_{L,\alpha}^r \bar{\chi}_r \cosh(\alpha) - \Omega_{L,\alpha}$$

with $\Omega_{L,\alpha} = \chi_r \cosh(L) (e^{-(\alpha+L)D} - R_{L,\alpha}^r) \bar{\chi}_r \cosh(\alpha)$ and $R_{L,\alpha}^r$ an operator defined on $L^2(\mathbb{R})$ through its kernel defined for $x, y \in \mathbb{R}$ (after some rearranging) as

$$R_{L,\alpha}^r(x, y) = (\pi(e^{2\alpha} - e^{-2L}))^{-1/2} e^{\alpha - u_1 x^2 + u_2 x + u_3},$$

where u_1, u_2 and u_3 depend on y and are given by

$$u_1 = \frac{1+e^{2\alpha+2L}}{2(e^{2\alpha+2L}-1)}, \quad u_2 = \frac{2e^L(r+e^{2\alpha}r-e^\alpha y)}{e^{2\alpha+2L}-1}, \quad u_3 = -\frac{2(1+e^{2\alpha})(1+e^{2L})r^2-4e^\alpha(1+e^{2L})ry+(1+e^{2\alpha+2L})y^2}{2(e^{2\alpha+2L}-1)}.$$

Using this, the determinant in (2.7) is equal to

$$\det \left(I - K_H^N \chi_r \cosh(\alpha) K_H^N - e^{(\alpha+L)D} K_H^N R_{L,\alpha}^r \bar{\chi}_r \cosh(\alpha) K_H^N - e^{(\alpha+L)D} K_H^N \Omega_{L,\alpha} K_H^N \right).$$

$\Omega_{L,\alpha}$ is to be thought of as an error term, and in fact essentially the same proof as that of [5], Lemma 2.3, shows that $e^{(\alpha+L)D} K_H^N \Omega_{L,\alpha} K_H^N \xrightarrow{L \rightarrow \infty} 0$ in trace norm. The Fredholm determinant is continuous with respect to the trace norm, so this together with (2.5) and (2.7) shows that $\mathbb{P}(\sqrt{2}\mathcal{M}_N(p) \leq r) = \lim_{L \rightarrow \infty} \mathbb{P}(\lambda_N(t) \leq r \cosh(t) \forall t \in [-L, \alpha])$ equals

$$\lim_{L \rightarrow \infty} \det \left(I - K_H^N \chi_r \cosh(\alpha) K_H^N - e^{(\alpha+L)D} K_H^N R_{L,\alpha}^r \bar{\chi}_r \cosh(\alpha) K_H^N \right).$$

We will see now that the last kernel inside this last determinant does not depend on L , and equals the one appearing in the statement of the proposition. The contour integral representation of the Hermite function $\varphi_n(x) = (2^n n! \sqrt{\pi})^{-\frac{1}{2}} e^{-x^2/2} \frac{n!}{2\pi i} \oint dw \frac{e^{2wx-w^2}}{w^{n+1}}$ (the integral is over a simple contour around the origin) together with (2.6) imply that $e^{(\alpha+L)D} K_H^N R_{L,\alpha}^r(x, y)$ is equal to

$$\frac{e^\alpha}{\sqrt{\pi(e^{2\alpha} - e^{-2L})}} \sum_{n=0}^{N-1} e^{(\alpha+L)n} \varphi_n(x) (2^n n! \sqrt{\pi})^{-\frac{1}{2}} \frac{n!}{2\pi i} \oint dt \frac{e^{-t^2}}{t^{n+1}} \int_{\mathbb{R}} dz e^{-z^2/2+2tz-u_1 z^2+u_2 z+u_3}.$$

The z integral is a Gaussian integral, and computing it, changing variables $t \mapsto te^{(\alpha+L)}$ and regrouping terms in the exponent we get that $e^{(\alpha+L)D} K_H^N R_{L,\alpha}^r(x, y)$ equals

$$\sum_{n=0}^{N-1} \varphi_n(x) \left[(2^n n! \sqrt{\pi})^{-\frac{1}{2}} e^{-(e^{-\alpha}r+e^\alpha r-y)^2/2} \frac{n!}{2\pi i} \oint dt \frac{e^{-t^2+2t(e^{-\alpha}r+e^\alpha r-y)}}{t^{n+1}} \right] e^{r^2 \sinh(2\alpha) - 2r \sinh(\alpha)y}.$$

The term inside the brackets is just $\varphi_n(e^{-\alpha}r + e^\alpha r - y) = \varphi_n(2r \cosh(\alpha) - y)$ and therefore $e^{(\alpha+L)D} K_H^N R_{L,\alpha}^r = K_H^N e^{r \sinh(\alpha)\xi} \varrho_r \cosh(\alpha) e^{-r \sinh(\alpha)\xi}$, whence the result follows. \square

Corollary 2.2.

$$\lim_{p \rightarrow 0} \mathbb{P}(\mathcal{M}_N(p) \leq \sqrt{2pr}) = \det(I - K_H^N \chi_r K_H^N - K_H^N E_{-r} \varrho_r E_r \bar{\chi}_r K_H^N). \tag{2.8}$$

Proof. We need to compute the limit of the right hand side of (2.3) with r replaced by $2\sqrt{p}r$. Since for $\alpha = \frac{1}{2} \log(\frac{p}{1-p})$ one has $2\sqrt{p} \cosh(\alpha) \rightarrow 1$ and $2\sqrt{p} \sinh(\alpha) \rightarrow -1$, it is easy to see that the kernel inside the resulting Fredholm determinant converges pointwise to the kernel appearing on the right hand side of (2.8). In order to upgrade this to convergence of the Fredholm determinant itself we need to show that the convergence holds in trace norm. The proof of this is standard, and can be done by adapting the arguments used in [5], Appendix B; the present setting is simpler, and we omit the details. \square

3. PROOF OF THEOREM 1.1

Recall the definition of the Laguerre Unitary Ensemble in (1.6) (with $\beta = 2$). By standard methods in RMT and determinantal point processes (see e.g. [22, 33]), the distribution $F_{\text{LUE},m}^{(a)}$ of the rightmost charge λ_m in the size m generalized LUE can be expressed as a Fredholm determinant:

$$F_{\text{LUE},m}^{(a)}(r) = \det(I - K_L^{m,(a)} \chi_r K_L^{m,(a)}), \tag{3.1}$$

where the Laguerre kernel $K_L^{m,(a)}$ is defined as

$$K_L^{m,(a)} = \sum_{k=0}^{m-1} \psi_k^{(a)}(x) \psi_k^{(a)}(y),$$

for $\psi_k^{(a)}$ the generalized Laguerre functions which are defined as

$$\psi_k^{(a)}(x) = \frac{\sqrt{n!}}{\sqrt{\Gamma(n+a+1)}} x^{a/2} e^{-x/2} L_k^{(a)}(x) \mathbb{1}_{x \geq 0},$$

with $L_k^{(a)}$ the generalized Laguerre polynomial of degree k , satisfying $\int_0^\infty dx L_m^{(a)}(x) L_n^{(a)}(x) x^a e^{-x} = \frac{\Gamma(n+a+1)}{n!} \delta_{mn}$, i.e. the orthogonal polynomials with respect to the weight $x^a e^{-x}$, normalized so that $\|\psi_k^{(a)}\|_2 = 1$.

Let

$$U_N(r) = \det(I - K_H^N \chi_r K_H^N - K_H^N E_{-r} \varrho_r E_r \bar{\chi}_r K_H^N), \tag{3.2}$$

which is the right hand side of (2.8). Our goal is to prove:

Proposition 3.1. *For every $N \in \mathbb{N}$ and $r \geq 0$,*

$$U_N(r) = \det(I - K_L^{\lfloor (N+1)/2 \rfloor, \frac{1}{2}(-1)^N} \chi_{r^2} K_L^{\lfloor (N+1)/2 \rfloor, \frac{1}{2}(-1)^N}). \tag{3.3}$$

Theorem 1.1 follows directly from this and Corollary 2.2 after replacing r by $r/\sqrt{2}$. We turn now to the proof of the proposition.

The kernel K_H^N has finite rank, so the above Fredholm determinant can be expressed as the determinant of a finite matrix. In order to do so, let $J_1 : \ell^2(\{0, \dots, N-1\}) \rightarrow L^2(\mathbb{R})$ and $J_2 : L^2(\mathbb{R}) \rightarrow \ell^2(\{0, \dots, N-1\})$ be the kernels given by

$$J_1(x, n) = \varphi_n(x), \quad \text{and} \quad J_2(n, y) = \varphi_n(y),$$

where $\varphi_n(x) = C_n e^{-x^2/2} H_n(x)$ are the Hermite functions introduced above. Here H_n are the standard Hermite polynomials and $C_n = (\sqrt{\pi} 2^n n!)^{-\frac{1}{2}}$. We have $J_1 J_2 = K_H^N$ while $J_2 J_1$ equals the identity in $\ell^2(\{0, \dots, N-1\})$, so applying the cyclic property of the Fredholm determinant to the right hand side of (3.2) we deduce that

$U_N(r) = \det(I - J_2\chi_r J_1 - J_2 E_{-r} \varrho_r E_r \bar{\chi}_r J_1)$, which is now a Fredholm determinant on $\ell^2(\{0, \dots, N - 1\})$. As such, it can be rewritten as the determinant of an $N \times N$ matrix as desired:

$$U_N(r) = \det(I - M)$$

for M the matrix indexed by $\{0, \dots, N - 1\}$ and given by

$$\begin{aligned} M_{j,k} &= (J_2\chi_r J_1 + J_2 E_{-r} \varrho_r E_r \bar{\chi}_r J_1)(j, k) = \int_r^\infty dz \varphi_j(z) [\varphi_k(z) + \varphi_k(2r - z)e^{2r^2 - 2rz}] \\ &= C_j C_k \int_0^\infty dz H_j(r + z)[H_k(r + z) + H_k(r - z)] e^{-(r+z)^2} \\ &= C_j C_k \left[\sum_{l=0}^k \binom{k}{l} (4r)^{k-l} (-1)^l \int_r^\infty dz H_j(z) H_l(z) e^{-z^2} + \int_r^\infty dz H_j(z) H_k(z) e^{-z^2} \right], \end{aligned}$$

where in the second line we used the change of variable $z \mapsto z + r$ and in the third one the identity $H_n(x + y) = \sum_{k=0}^n \binom{n}{k} H_k(x)(2y)^{n-k}$ [34]. If we now define the matrices

$$F_{j,k} = \binom{k}{j} (4r)^{k-j} (-1)^j \mathbf{1}_{j \leq k} + \delta_{j,k}, \quad Q_{j,k} = C_j^2 \int_r^\infty dz H_j(z) H_k(z) e^{-z^2}, \tag{3.4}$$

then using the cyclic property of the determinant we obtain

$$U_N(r) = \det(I - QF). \tag{3.5}$$

Now let $G_N(r)$ denote the right hand side of (3.3), so that our goal is to prove $U_N = G_N$. In everything that follows we let $b = 1$ if N is odd and $b = 2$ if N is even. The Laguerre kernel appearing inside the Fredholm determinant in (3.3) can be expressed in terms of Laguerre polynomials using the relation (see [35])

$$L_n^{(\frac{1}{2}(-1)^b)}(u) = \frac{(-1)^n 2^{-2n-b+1}}{n!} u^{-(1+(-1)^b)/4} H_{2n+b-1}(u^{1/2}),$$

computing as we just did above using the cyclic property of the determinant leads to

$$G_N(r) = \det(I - A),$$

where the matrix A is indexed by $\{0, \dots, \lfloor \frac{N-1}{2} \rfloor\}$ and is given by $A_{j,k} = 2Q_{2j+b-1, 2k+b-1}$. Note that the matrices QF and A appearing within the determinants for U_N and G_N have different dimensions. In principle, this makes it hard to compare the two determinants. The key is that M , which is of size N , actually has rank $\lfloor \frac{N+1}{2} \rfloor$, which is the size of A . In order to see, and use this, we first introduce the matrices

$$S_{j,t} = \frac{t!}{2^{2j}(t - N + 2j + b)!} H_{t-N+2j+b}(r), \quad j = 0, \dots, \frac{1}{2}(N - b), \quad t = 0, \dots, N - 1, \tag{3.6}$$

$$T_{u,k} = \frac{(-1)^{\frac{N-b}{2}+k} (-i)^u 2^{2k}}{u!(N - 2k - b - u)!} H_{N-2k-b-u}(ir), \quad u = 0, \dots, N - 1, \quad k = 0, \dots, \frac{1}{2}(N - b). \tag{3.7}$$

Here, and in everything that follows, we use the convention

$$H_m(x) = 0 \quad \text{and} \quad \frac{1}{m!} = 0 \quad \forall m < 0.$$

We note that $H_n(x)$ has the same parity as n , and thus $i^n H_n(ir)$ is real, which shows that T is a real matrix. Proposition 3.1 will follow from the following two results:

Proposition 3.2. *The matrices S and T satisfy:*

- (1) $2TS = F$ (with F as defined in (3.4)).
- (2) $ST = I$, the identity matrix of size $\lfloor (N + 1)/2 \rfloor$.

Property (1) implies that F , and thus QF , have rank $\lfloor (N + 1)/2 \rfloor$ as claimed.

Proposition 3.3. $2SQT = A$.

Proof of Proposition 3.1. (3.3) is now a simple consequence of the above identities and the cyclic property of the determinant:

$$G_N(r) = \det(I - A) = \det(I - 2SQT) = \det(I - 2QTS) = \det(I - QF) = U_N(r). \quad \square$$

We turn to the proofs of the two propositions. They depend on the following two lemmas, whose proofs are deferred to the end of the section.

Lemma 3.4. *Let m be a non-negative integer.*

- (1) *If $m > 0$ then*

$$\sum_{s \geq 0} \binom{m}{2s+1} (-1)^s H_{m-2s-1}(ir) H_{2s+1}(r) = \frac{1}{2} i^{m-1} (4r)^m, \tag{3.8}$$

$$\sum_{s \geq 0} \binom{m}{2s} (-1)^s H_{m-2s}(ir) H_{2s}(r) = \frac{1}{2} i^m (4r)^m. \tag{3.9}$$

- (2) $\sum_{t=0}^m (-i)^t \binom{m}{t} H_t(r) H_{m-t}(ir) = \delta_{m,0}$.

- (3) *For all integers d and $m \geq 0$ such that $d + m \geq 0$,*

$$\sum_{u=0}^m \binom{m}{u} (-i)^u H_{m-u}(ir) H_{u+m+d}(r) = \frac{i^m 2^m (m+d)!}{d!} H_d(r).$$

Lemma 3.5. *For all integers $m \geq 0$, $t \geq 0$ and $d \geq m$:*

- (i) *If $t \geq m + 1$ then $\sum_{l \geq 0} (-1)^l \binom{t}{l} \binom{t-l-1}{m-l} = (-1)^m$.*

- (ii) $\sum_{u=0}^k (-1)^u \frac{(t+u)!}{u!(k-u)!(t+u-h)!} = (-1)^k \frac{h!}{k!} \binom{t}{h-k}$.

Proof of Proposition 3.2. We have

$$(2TS)_{j,k} = 2(-1)^{\frac{N-b}{2}} (-i)^j \frac{k!}{j!} \sum_{s=0}^{\frac{N-b}{2}} (-1)^s \frac{H_{N-2s-b-j}(ir)}{(N-2s-b-j)!} \frac{H_{k-N+2s+b}(r)}{(k-N+2s+b)!}.$$

Note that the right hand side vanishes whenever $N - 2s - b - j < 0$ or $k - N + 2s + b < 0$, and thus for $j > k$ we have $(2TS)_{j,k} = 0 = F_{j,k}$, which proves (1) in this case. Now take $j \leq k$. Using the same argument we can replace the summation over $0 \leq s \leq \frac{N-b}{2}$ by a summation over $s \geq \frac{N-b}{2} - \ell$ with $\ell = \lfloor \frac{k}{2} \rfloor$ and then change variables to rewrite the above expression as

$$(2TS)_{j,k} = 2(-1)^{-\ell} (-i)^j \binom{k}{j} \sum_{s=0}^{\infty} (-1)^s \binom{k-j}{k-2\ell+2s} H_{2\ell-j-2s}(ir) H_{k-2\ell+2s}(r)$$

In the case $k = j$, when k is even the only term surviving from the sum is $s = 0$ and the sum equals 1 (using $H_0 \equiv 1$), while if k is odd all terms in the sum vanish, so using this we get $(2TS)_{k,k} = (-1)^k + 1 = F_{k,k}$. Assume now that $j < k$. If k is even $(2TS)_{j,k}$ equals

$$2(-1)^{\frac{k}{2}} (-i)^j \binom{k}{j} \sum_{s=0}^{\infty} (-1)^s \binom{k-j}{2s} H_{k-j-2s}(ir) H_{2s}(r) = 2(-1)^{\frac{k}{2}} (-i)^j \binom{k}{j} \frac{1}{2} i^{k-j} (4r)^{k-j}$$

by (1) from Lemma 3.4. The right hand side equals $\binom{k}{j} (-1)^j (4r)^{k-j}$, which is $F_{j,k}$ as desired. The case where k is odd is analogous. This finishes proving (1).

Next we consider (2). We have

$$\begin{aligned} (ST)_{j,k} &= (-1)^{\frac{N-b}{2}+k} 2^{2(k-j)} \sum_{t=0}^{N-1} (-i)^t \frac{H_{t-N+2j+b}(r)}{(t-N+2j+b)!} \frac{H_{N-2k-b-t}(ir)}{(N-2k-b-t)!} \\ &= (-1)^{\frac{N-b}{2}+k} 2^{2(k-j)} \sum_{t=N-b-2j}^{N-a-2k} (-i)^t \frac{H_{t-N+2j+b}(r)}{(t-N+2j+b)!} \frac{H_{N-2k-b-t}(ir)}{(N-2k-a-t)!} \\ &= (-1)^{\frac{N-b}{2}+k} 2^{2(j-k)} (-i)^{N-b-2j} \frac{1}{(2(j-k))!} \sum_{t=0}^{2(k-j)} (-i)^t \binom{2(j-k)}{t} H_t(r) H_{2(j-k)-t}(ir), \end{aligned}$$

where in the second equality we used again that if either $t - N + 2j + b < 0$ or $N - 2k - b - t < 0$ then the sum vanishes. In particular, if $j < k$ then $(ST)_{j,k} = 0$. For $j \geq k$, we apply (2) from Lemma 3.4 to obtain

$$(ST)_{j,k} = (-1)^{\frac{N-b}{2}+k} 2^{2(k-j)} (-i)^{N-b-2j} \frac{1}{(2(j-k))!} \delta_{k-j,0} = \delta_{k-j,0},$$

and the proof is complete. □

Proof of Proposition 3.3. Using the standard Hermite polynomial identities $H_l(z) = 2zH_{l-1}(z) - 2(l-1)H_{l-2}(z)$ and $H'_l(z) = 2lH_{l-1}(z)$ we have

$$\int_r^\infty dz H_l(z) e^{-z^2} = H_{l-1}(r) e^{-r^2},$$

for all $l \geq 1$. On the other hand, the product of two Hermite polynomials can be expressed [35] as

$$H_j(z)H_k(z) = \sum_{l=0}^{\min\{j,k\}} 2^l l! \binom{j}{l} \binom{k}{l} H_{j+k-2l}(z). \tag{3.10}$$

Using these facts in (3.4) we get

$$Q_{j,k} = C_j^2 \sum_{l=0}^{\min\{j,k\}} 2^l l! \binom{j}{l} \binom{k}{l} H_{j+k-2l-1}(r) e^{-r^2} \quad \text{for } j \neq k, \tag{3.11}$$

$$Q_{0,0} = \frac{1}{2} \operatorname{erfc}(r), \quad Q_{jj} = \frac{1}{2} \operatorname{erfc}(r) + C_j^2 \sum_{l=0}^{j-1} 2^l l! \binom{j}{l} \binom{j}{l} H_{2j-2l-1}(r) e^{-r^2} \quad \text{for } j > 0, \tag{3.12}$$

where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty dz e^{-z^2}$ is the complementary error function. Now define matrices

$$\tilde{Q} = Q - \frac{1}{2} \operatorname{erfc}(r) I \quad \text{and} \quad \tilde{A} = A - \frac{1}{2} \operatorname{erfc}(r) \tilde{I}$$

where I and \tilde{I} are the identity matrices of sizes N and $(N - a + 1)/2$. Note that $\tilde{A}_{j,k} = 2\tilde{Q}_{2j+b-1, 2k+b-1}$. Thanks to (2) of Proposition 3.2, in order to prove the result it is enough to show that

$$2S\tilde{Q}T = \tilde{A}.$$

We have

$$\begin{aligned} \sqrt{\pi} e^{r^2} (S\tilde{Q}T)_{j,k} &= (-1)^{\frac{N-b}{2}+k} 2^{2(k-j)} \sum_{t=N-2j-b}^{N-1} \sum_{l=0}^t \frac{1}{(t-N+2j+b)!} 2^{l-t} \binom{t}{l} \frac{1}{(N-2k-b-l)!} \\ &\times H_{t+2j+b-N}(r) (-i)^l \sum_{u=0}^{N-2k-b-l} \binom{N-2k-b-l}{u} (-i)^u H_{N-2k-b-u-l}(ir) H_{t+u-l-1}(r), \end{aligned} \tag{3.13}$$

where we have applied the change $u \mapsto u - l$. We will focus on the case $j \geq k$ and explain at the end how the same arguments work for the case $j < k$. Consider first the sum in t restricted to $N - 2k - b < t < N$. Using (3) from Lemma 3.4 with $m = N - 2k - b - l \geq 0$ and $d = t + 2k + b - 1 - N \geq 0$, the u sum equals $\frac{i^{N-2k-b-l} 2^{N-2k-b-l} (t-l-1)!}{(t+2k+b-1-N)!} H_{t+2k+b-1-N}(r)$ for $l < t$ and is zero for $l = t$, and thus (3.13) with the t sum restricted to $N - 2k - b < t < N$ is equal to

$$\begin{aligned} 2^{N-2j-b} \sum_{t=N-2k-b+1}^{N-1} \frac{H_{t+2j+b-N}(r) H_{t+2k+b-1-N}(r)}{2^t (t-N+2j+b)!} \sum_{l=0}^{t-1} (-1)^l \binom{t}{l} \binom{t-l-1}{N-2k-b-l} \\ = 2^{N-2j-b} \sum_{t=N-2k-b+1}^{N-1} \frac{H_{t+2j+b-N}(r) H_{t+2k+b-1-N}(r)}{2^t (t-N+2j+b)!}, \end{aligned} \tag{3.14}$$

where we have used (i) in Lemma 3.5 with $m = N - 2k - b$ (which is even). Applying the change of variable $t \mapsto t - (N - 2k - b + 1)$, using (3.10) and changing the order of the sum, the previous expression is equal to

$$\begin{aligned} \frac{1}{2^{2j-2k+1}} \sum_{h=0}^{2k+b-2} \sum_{t=h}^{2k+b-2} \frac{1}{2^{t-h} (t+2j-2k+1-h)!} \binom{t}{h} H_{2t+2j-2k-1-2h}(r) \\ = \frac{1}{2^{2j-2k+1}} \sum_{t=0}^{2k+b-2} \frac{1}{2^t (t+2j-2k+1)!} H_{2t+2j-2k-1}(r) \sum_{h=0}^{2k+b-2-t} \binom{t+h}{t}, \end{aligned} \tag{3.15}$$

where in the second line we have changed $t \mapsto t + h$ and interchanged the order of summation. Using the identity $\sum_{i=t}^s \binom{i}{t} = \binom{s+1}{t+1}$, the h sum above equals $\binom{2k+b-1}{2k+b-2-t}$, and then changing $t \mapsto 2k + b - 2 - t$ we get

$$\frac{1}{2^{2j+b-1}(2j+b-1)!} \sum_{t=0}^{2k+b-2} 2^t t! \binom{2j+b-1}{t} \binom{2k+b-1}{t} H_{2j+2k+2b-2t-3}(r). \quad (3.16)$$

Now we compute the sum (3.13) in the region $N - 2j - b \leq t \leq N - 2k - b$. Consider first the terms with $0 \leq l < t - 1$. For the sum in u we use (3) of Lemma 3.4 with the same choices of m and d . Now $m \geq 0$ and $m + d = t - l - 1 \geq 0$ but $d < 0$, so the whole sum vanishes. This means that we are only left with the term $l = t$, namely

$$\begin{aligned} (-1)^{\frac{N-b}{2}+k} 2^{2(k-j)} \sum_{t=N-2j-b}^{N-2k-b} \frac{1}{(t-N+2j+b)!} \frac{1}{(N-2k-b-t)!} H_{t+2j+b-N}(r) \\ \times \sum_{u=0}^{N-2k-b-t} (-i)^{u+t} \binom{N-2k-b-t}{u} H_{N-2k-b-u-t}(ir) H_{u-1}(r). \end{aligned} \quad (3.17)$$

The $u = 0$ term vanishes since $H_{-1}(r) = 0$. Then changing variables $t \mapsto t + N - 2j - b$, the above becomes

$$(-1)^{k-j} 2^{2(k-j)} \sum_{u=1}^{2j-2k} \frac{(-i)^u H_{u-1}(r)}{u!(2j-2k-u)!} \sum_{t=0}^{2j-2k-u} \binom{2j-2k-u}{t} (-i)^t H_t(r) H_{2j-2k-u-t}(ir).$$

By (2) of Lemma 3.4 with the last sum is equal to $\frac{1}{2^{2j-2k} \frac{1}{(2j-2k)!}} H_{2j-2k-1}(r)$, which completes the term for $t = 2k + b - 1$ for the sum in (3.16), then taking the change of variables $t \rightarrow 2k + b - 1 - t$ we obtain $\frac{1}{2} \sqrt{\pi} e^{r^2} \tilde{A}_{j,k}$ and the result follows.

The case $j < k$ is simpler. In computing (3.13) we now have $t \geq N - 2j - b + 1 > N - 2k - b + 1$, so only the first t region needs to be considered. The same argument now leads to (3.13) being equal to

$$\frac{1}{2^{2j+b-1}(2j+b-1)!} \sum_{t=0}^{2j+b-1} 2^t t! \binom{2j+b-1}{t} \binom{2k+b-1}{t} H_{2j+2k+2b-2t-3}(r).$$

□

It only remains to prove Lemmas 3.4 and 3.5.

Proof of Lemma 3.4. Let $H_n^{[\alpha]}(x) = \left(\frac{\alpha}{2}\right)^{n/2} H_n(x/\sqrt{2\alpha})$. From [36], equation (4.2.1) (note that the book uses a different Hermite function normalization) we have $\sum_{s=0}^n \binom{n}{s} H_s^{[\alpha]}(x) H_{n-s}^{[-\alpha]}(y) = (x+y)^n$, and therefore

$$\sum_{s \geq 0} \binom{m}{s} H_{m-s}^{[-1/2]}(-r) H_s^{[1/2]}(r) = 0,$$

while, since the $H_k(x)$ is even for even k and odd for odd k ,

$$\sum_{s \geq 0} (-1)^s \binom{m}{s} H_{m-s}^{[-1/2]}(-r) H_s^{[1/2]}(r) = \sum_{s \geq 0} \binom{m}{s} H_{m-s}^{[-1/2]}(-r) H_s^{[1/2]}(-r) = (-2r)^m.$$

Adding and subtracting these two equations we obtain

$$\sum_{s \geq 0} \binom{m}{2s} H_{m-2s}^{[-1/2]}(-r) H_{2s}^{[1/2]}(r) = \frac{1}{2} (-2r)^m, \text{ and } \sum_{s \geq 0} \binom{m}{2s+1} H_{m-2s-1}^{[-1/2]}(-r) H_{2s+1}^{[1/2]}(r) = -\frac{1}{2} (-2r)^m.$$

Since $H_n(r) = 2^n H_n^{[1/2]}(r)$ and $H_n(ir) = (-i)^n 2^n H_n^{[-1/2]}(-r)$, the identities in (1) follow.

Next we consider (2). The case $m = 0$ is straightforward (since $H_0 \equiv 1$). For $m > 0$ we use the identity

$$H_n(\gamma r) = \sum_{l=0}^{\lfloor \frac{n}{2} \rfloor} \gamma^{n-2l} (\gamma^2 - 1)^l \binom{n}{2l} \frac{(2l)!}{l!} H_{n-2l}(r), \quad (3.18)$$

which is valid for all $\gamma \in \mathbb{C}$ [34]. Using it with $\gamma = i$ we get

$$\begin{aligned} \sum_{t=0}^m (-i)^t \binom{m}{t} H_t(r) H_{m-t}(ir) &= i^m \sum_{t \geq 0} \sum_{l \geq 0} (-1)^t \binom{m}{t} 2^l \binom{m-t}{2l} \frac{(2l)!}{l!} H_t(r) H_{m-t-2l}(r), \\ &= i^m m! \sum_{l \geq 0} \sum_{h \geq 0} 2^{l+h} \frac{1}{l!h!} H_{m-2l-h}(r) \frac{1}{(m-2l-2h)!} \sum_{t \geq 0} (-1)^t \binom{m-2l-2h}{t-h} = 0, \end{aligned} \quad (3.19)$$

where in the second equality we used (3.10) and the third one follows because the last sum equals $(-1)^h \sum_{t=0}^{m-2l-2h} (-1)^t \binom{m-2l-2h}{t}$, which vanishes by the Binomial Theorem.

For (3) we write use (3.18) and then (3.10) to write the left hand side as

$$i^m \sum_{u=0}^m \sum_{l=0}^{\lfloor \frac{m-u}{2} \rfloor} (-1)^u \binom{m}{u} \binom{m-u}{2l} \frac{2^l (2l)!}{l!} H_{m-u-2l}(r) H_{m+u+d}(r) \quad (3.20)$$

$$= m! i^m \sum_{u=0}^m \sum_{l=0}^{\lfloor \frac{m-u}{2} \rfloor} \sum_{h=0}^{m-u-2l} \frac{(-1)^u (m+u+d)! 2^{l+h}}{l!h!u!(m-2l-h-u)!(m+u+d-h)!} H_{2m+d-2l-2h}(r) \quad (3.21)$$

$$= m! i^m \sum_{l=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{h=0}^{m-2l} \frac{2^{l+h} H_{2m+d-2l-2h}(r)}{l!h!} \sum_{u=0}^{m-2l-h} \frac{(-1)^u (m+u+d)!}{u!(m-2l-h-u)!(m+u+d-h)!} \quad (3.22)$$

$$= m!(m+d)! i^m \sum_{l=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{h=0}^{m-2l} \frac{(-1)^{m-h} 2^{l+h} H_{2m+d-2l-2h}(r)}{l!(m-2l-h)!(2h+2l-m)!(2m+d-2l-2h)!}, \quad (3.23)$$

where the third line is obtained by the Fubini's theorem and then we have used Lemma 3.5 part (ii) with $k = m - 2l - h$ and $t = m + d$. Let $g(r)$ denote the last expression. We want to prove that $g(r) = \frac{i^m 2^m (m+d)!}{d!} H_d(r)$ which, by the orthogonality of the Hermite polynomials ($\int_{\mathbb{R}} dr H_n(r) H_k(r) e^{-r^2} = C_n^{-2} \delta_{n,k}$) is equivalent to proving that, for each $n \geq 0$,

$$\int_{\mathbb{R}} dr g(r) H_n(r) e^{-r^2} = \frac{i^m 2^m (m+d)!}{C_d^2 d!} \delta_{n,d}.$$

Passing the integration inside the sums defining $g(r)$ yields a factor $C_n^{-2} \delta_{n,2m+d-2l-2h}$ inside the sum, and thus the integral can only be non-zero if d and n have the same parity, in which case the only non-zero term comes from $h = m - l + \frac{d-n}{2}$. In other words, the left hand side above equals

$$\frac{m!(m+d)! i^m (-1)^{(n-d)/2} 2^{m+(d-n)/2}}{C_n^2 (m-(n-d))! (\frac{1}{2}(n-d))! n!} \sum_{l=0}^{\lfloor \frac{m}{2} \rfloor} (-1)^l \binom{\frac{1}{2}(n-d)}{l}.$$

The prefactor vanishes unless $0 \leq (n-d) \leq m$, and in this case we have $\lfloor m/2 \rfloor \geq (n-d)/2$, so the last sum only ranges over $0 \leq l \leq (n-d)/2$ and then if $n > d$ the sum vanishes by the Binomial Theorem. The only possibility is then $n = d$, which forces $l = 0$ and $h = m$, turning the above expression into $\frac{i^m 2^m (m+d)!}{C_d^2 d!}$, as desired. \square

Proof of Lemma 3.5. Using the Binomial Theorem we have, for $x \in \mathbb{R}$ with $|x| < 1$,

$$\sum_{m \geq 0} \sum_{l \geq 0} (-1)^l \binom{t}{l} \binom{t-l-1}{m-l} x^m = \sum_{l \geq 0} (-x)^l \binom{t}{l} \sum_{m \geq 0} \binom{t-l-1}{m} x^m = (1+x)^{t-1} \sum_{l \geq 0} \binom{t}{l} \left(\frac{-x}{1+x} \right)^l,$$

which equals $\frac{1}{1+x} = \sum_{m \geq 0} (-1)^m x^m$. (i) now follows by equating coefficients.

For (ii) it is enough to rearrange the identity as $\sum_{u=0}^k (-1)^u \binom{k}{u} \binom{t+u}{h} = (-1)^k \binom{t}{h-k}$, which can be proved by induction in k using Pascal's rule $\binom{k+1}{u} = \binom{k}{u} + \binom{k}{u-1}$. \square

ACKNOWLEDGEMENTS

Both authors received support from Centro de Modelamiento Matemático Basal Funds FB210005 from ANID-Chile. Y.Y.'s work was funded in part by the National Agency for Research and Development (ANID)/DOCTORADO NACIONAL/2019-21191420. Both authors thank Professor Daniel Remenik for many helpful discussions and comments and for his exceptional support and guidance through the development of this article.

DATA AVAILABILITY STATEMENT

No new data/code were created or analyzed in this study.

REFERENCES

- [1] C.A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159** (1994) 151–174.
- [2] K. Matetski, J. Quastel and D. Remenik, The KPZ fixed point. *Acta Math.* **227** (2021) 115–203.
- [3] K. Johansson, Discrete polynuclear growth and determinantal processes. *Commun. Math. Phys.* **242** (2003) 277–329.
- [4] C.A. Tracy and H. Widom, On orthogonal and symplectic matrix ensembles. *Commun. Math. Phys.* **177** (1996) 727–754.
- [5] G.B. Nguyen and D. Remenik, Non-intersecting Brownian bridges and the Laguerre orthogonal ensemble. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 2005–2029.
- [6] W. FitzGerald and J. Warren, Point-to-line last passage percolation and the invariant measure of a system of reflecting Brownian motions. *Probab. Theory Related Fields* **178** (2020) 121–171.
- [7] J. Quastel and D. Remenik, Airy Processes and Variational Problems. Vol. 69 of *Springer Proc. Math. Stat.*. Springer, New York (2014) 121–171.
- [8] G.B. Nguyen and D. Remenik, Extreme statistics of non-intersecting Brownian paths. *Electron. J. Probab.* **22** (2017) Paper No. 102, 40.
- [9] M. Adler, J. Delépine and P. van Moerbeke, Dyson's nonintersecting Brownian motions with a few outliers. *Commun. Pure Appl. Math.* **62** (2009) 334–395.
- [10] M. Adler, P.L. Ferrari and P. van Moerbeke, Airy processes with wanderers and new universality classes. *Ann. Probab.* **38** (2010) 714–769.
- [11] P.J. Forrester, S.N. Majumdar and G. Schehr, Non-intersecting Brownian walkers and Yang–Mills theory on the sphere. *Nuclear Phys. B* **844** (2011) 500–526.
- [12] K. Liechty, Nonintersecting Brownian motions on the half-line and discrete Gaussian orthogonal polynomials. *J. Stat. Phys.* **147** (2012) 582–622.
- [13] K. Liechty, G.B. Nguyen and D. Remenik, Airy process with wanderers, KPZ fluctuations, and a deformation of the Tracy–Widom GOE distribution. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 2250–2283.
- [14] K. Liechty and D. Wang, Nonintersecting Brownian bridges between reflecting or absorbing walls. *Adv. Math.* **309** (2017) 155–208.
- [15] G. Schehr, S.N. Majumdar, A. Comtet and J. Randon-Furling, Exact distribution of the maximal height of p vicious walkers. *Phys. Rev. Lett.* **101** (2008) 150601, 4.

- [16] C.A. Tracy and H. Widom, Differential equations for Dyson processes. *Commun. Math. Phys.* **252** (2004) 7–41.
- [17] C.A. Tracy and H. Widom, Nonintersecting Brownian excursions. *Ann. Appl. Probab.* **17** (2007) 953–979.
- [18] J. Warren, Dyson’s Brownian motions, intertwining and interlacing. *Electron. J. Probab.* **12** (2007) 573–590.
- [19] A. Borodin, P.L. Ferrari and T. Sasamoto, Transition between Airy₁ and Airy₂ processes and TASEP fluctuations. *Commun. Pure Appl. Math.* **61** (2008) 1603–1629.
- [20] I. Corwin and A. Hammond, Brownian Gibbs property for Airy line ensembles. *Invent. Math.* **195** (2014) 441–508.
- [21] J. Quastel and D. Remenik, Supremum of the airy 2 process minus a parabola on a half line. *J. Statist. Phys.* **150** (2013) 442–456.
- [22] P.J. Forrester, Log-gases and Random Matrices. Vol. 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ (2010).
- [23] I.M. Johnstone, On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** (2001) 295–327.
- [24] K. Johansson, Shape fluctuations and random matrices. *Commun. Math. Phys.* **209** (2000) 437–476.
- [25] A. Edelman and M. La Croix, The singular values of the GUE (less is more). *Random Matrices Theory Appl.* **4** (2015) 1550021, 37.
- [26] M.L. Mehta, Random Matrices. Vol. 142 of *Pure and Applied Mathematics (Amsterdam)*, 3rd edn. Elsevier/Academic Press, Amsterdam (2004).
- [27] P. Sosoe and U. Smilansky, On the spectrum of random anti-symmetric and tournament matrices. *Random Matrices Theory Appl.* **5** (2016) 1650010, 33.
- [28] I. Corwin, J. Quastel and D. Remenik, Continuum statistics of the Airy₂ process. *Commun. Math. Phys.* **317** (2013) 347–362.
- [29] A. Borodin, P.L. Ferrari, M. Prähofer, T. Sasamoto and J. Warren, Maximum of Dyson Brownian motion and non-colliding systems with a boundary. *Electron. Commun. Probab.* **14** (2009) 486–494.
- [30] NIST Digital Library of Mathematical Functions. <http://dlmf.nist.gov/>. Online companion to [37].
- [31] B. Simon, Trace Ideals and their Applications. Vol. 120 of *Mathematical Surveys and Monographs*, 2nd edn. American Mathematical Society (2005).
- [32] A. Borodin, I. Corwin and D. Remenik, Multiplicative functionals on ensembles of non-intersecting paths. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2015) 28–58.
- [33] K. Johansson, Random matrices and determinantal processes, in *Mathematical Statistical Physics*. Elsevier B.V., Amsterdam (2006) 1–55.
- [34] Z.-y. Huang and J.-a. Yan, Introduction to Infinite Dimensional Stochastic Analysis. Vol. 502 of *Mathematics and its Applications*, Chinese edn. Kluwer Academic Publishers, Dordrecht; Science Press Beijing, Beijing (2000).
- [35] G.E. Andrews, R. Askey and R. Roy, Special Functions. Vol. 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1999).
- [36] S. Roman, The Umbral Calculus. Vol. 111 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York (1984).
- [37] F.W.J. Olver, D.W. Lozier, R.F. Boisvert and C.W. Clark, editors, NIST handbook of mathematical functions. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge (2010). With 1 CD-ROM (Windows, Macintosh and UNIX).

Please help to maintain this journal in open access!



This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.