Determinantal correlations for classical projection processes

Peter J Forrester\(^1\) and Taro Nagao\(^2\)

\(^1\) Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia
\(^2\) Graduate School of Mathematics, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan
E-mail: pjforr@unimelb.edu.au and nagao@math.nagoya-u.ac.jp

Received 27 April 2011
Accepted 14 July 2011
Published 24 August 2011

Abstract. Recent applications in queuing theory and statistical mechanics have isolated the process formed by the eigenvalues of successive sub-matrices of the GUE. Analogous eigenvalue processes, formed in general from the eigenvalues of nested sequences of matrices resulting from random corank-1 projections of classical random matrix ensembles, are identified for the LUE and JUE. The correlations for all these processes can be computed in a unified way. The resulting expressions can then be analyzed in various scaling limits. At the soft edge, with the rank of the sub-matrices differing by an amount proportional to \(N^{2/3}\), the scaled correlations coincide with those known from the soft edge scaling of the Dyson Brownian motion model.

Keywords: correlation functions, random matrix theory and extensions
1. Introduction

Since the pioneering days of Wigner, Gaudin, Mehta and Dyson (see [33] for a collection of papers from this period), random matrix theory has shown itself to be perhaps the richest source of exact solutions for correlation and distribution functions of all statistical mechanical models. Many of the discoveries of this type have their motivation in new applications of random matrix theory, unknown in the pioneering days. A case in point is the recent work of Johansson and Nordenstam [25,30], who compute the exact form of the correlation functions for the coupled eigenvalue sequences obtained from the principal sub-matrices of Gaussian unitary ensemble (GUE) matrices.

The motivation for studying this so-called GUE minor process begins with a work of Baryshnikov [1]. Some years earlier Glynn and Whitt [18] had studied the problem of computing the distribution of exit times from a queuing system; in the limit the number of queues tends to infinity but the number of jobs remains finite. It was proved that for
general i.i.d. service times the scaled distributions $D_k$ of the exit time of the $k$th customer from the final queue could be written as

$$D_k = \sup_X \sum_{i=0}^{k-1} (B_i(t_{i+1}) - B_i(t_i)).$$

Here each $B_i$ denotes an independent standard Brownian motion, while the condition $X$ is that

$$0 = t_0 < t_1 < \cdots < t_k = 1.$$

By studying the particular case of exponential waiting times Baryshnikov was able to show that $\{D_k\}_{k=1,2,\ldots}$ could alternatively be specified as the joint distribution of $\{\mu_k\}$, where $\mu_k$ is the largest eigenvalue of the $k$th principal sub-matrix of an infinite GUE matrix $X$ with probability density function (PDF) proportional to $\exp(-X^2/2)$. Johansson and Nordenstam [25] give other occurrences in statistical mechanics of essentially the same process relating to the eigenvalues of sub-matrices of GUE matrices. These are in the specification of certain point processes relating to the boundary region (neighborhood of the frozen zone) of random domino tilings of the Aztec diamond [23] and to random lozenge tilings of a hexagon [24]. In the equivalent language of stepped surfaces, Okounkov and Reshetikhin [32] make similar observations. The recent work of Borodin, Ferrari and Sasamoto [5] encounters this process in the context of studying the dynamics of the asymmetric exclusion process.

The broader setting of the measures encountered in studies of the above statistical mechanical models relates to the Robinson–Schensted–Knuth (RSK) correspondence (see section 2.1) below. Consideration of the structures inherent therein [2, 17] identifies natural extensions of the GUE minor process. One such class of extensions replaces the Gaussian weight in the latter by the Laguerre or Jacobi weights, which can all be realized by a sequence of projections onto random complex hyperplanes, giving rise to so-called classical projection processes (the Gaussian, Laguerre and Jacobi weights are all classical from the viewpoint of the theory of orthogonal polynomials). Our main point in the present paper is that the correlations for the classical projection processes can be computed exactly in a unified way. The essential ingredient here is the Rodrigues formula for classical orthogonal polynomials. Moreover, known asymptotic formulae for the latter allow for the evaluation of scaling limits of the correlations.

We begin in section 2 by recalling how the RSK correspondence relates to a certain statistical mechanical model of last passage times. In section 3 the occurrence of special cases of the joint PDF in some random matrix setting is noted. The correlations for the classical projection process are calculated in section 4 and their scaling limits analyzed in section 5.

### 2. A joint probability density associated with RSK

#### 2.1. The case of general parameters

The Robinson–Schensted–Knuth (RSK) correspondence gives a bijection between $n_1 \times n_2$ non-negative integer matrices $[x_{i,j}]$ (rows counted from the bottom) with entry $(i,j)$ weighted $(a_i b_j)^{x_{i,j}}$, and pairs of weighted semi-standard tableaux (weights $\{a_i\}$, $\{b_j\}$)
of shape $\mu = (\mu_1, \ldots, \mu_n)$. One recalls that by definition of a semi-standard tableau the boxes in the shape $\mu$ (each row $i$ of length $\mu_i$ and left-justified) must be filled with integers from $\{1, 2, \ldots, n\}$, with $n$ referring to the content, so that the integers weakly increase along rows and strictly increase down columns. With each integer $k$ repeated $f_k$ times, and with weights $\{w_k\}_{k=1}^\infty$, the weight of the tableau is $\prod_{k=1}^N w_k^{f_k}$. Results from [21, 2] associate a probabilistic model to the RSK correspondence. Identify with each lattice site $(ij)$ a random non-negative integer variable $x_{i,j}$ chosen from the geometric distribution with parameter $a_i b_j$, so that

$$\Pr(x_{i,j} = k) = (1 - a_i b_j)(a_i b_j)^k.$$  \hspace{1cm} (2.1)

For given non-negative parameters $\{a_i\}$, $\{b_j\}$, and given $n_1, n_2 \in \mathbb{Z}^+$, a probabilistic quantity of interest is the sequence of last passage times:

$$L^{(l)}(n_1, n_2) = \max \sum_{(rd^*)^l} x_{i,j}, \quad l = 1, \ldots, \max(n_1, n_2).$$ \hspace{1cm} (2.2)

Here $(rd^*)^l$ denotes the set of $l$ disjoint (no common lattice points) $rd^*$ lattice paths, which in turn are defined as either a single point, or points connected by segments formed out of arbitrary positive integer multiples of steps to the right and steps up in the rectangle $1 \leq i \leq n_1$, $1 \leq j \leq n_2$.

A crucial feature of the RSK correspondence is that the length of the first row of the semi-standard tableaux pair is equal to (2.2) in the case $l = 1$, and thus $\mu_1 = L^{(1)}(n_1, n_2)$. More generally, all row lengths are determined by (2.2) according to [20]

$$\mu_l = L^{(l)}(n_1, n_2) - L^{(l-1)}(n_1, n_2)$$ \hspace{1cm} (2.3)

with $L^{(0)}(n_1, n_2) := 0$. Thus the distribution of the last passage times is fully determined by the distribution of $\{\mu_l\}$.

A well-known fact relating to the RSK correspondence is that the probability an $n_1 \times n_2$ non-negative matrix with elements chosen according to (2.1) corresponds to a pair of semi-standard tableaux with shape $\mu$, one of content $n_1$, the other of content $n_2$, is given by [27]

$$\Pr(\mu) = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} (1 - a_i b_j)s_\mu(a_1, \ldots, a_{n_1})s_\mu(b_1, \ldots, b_{n_2}),$$ \hspace{1cm} (2.4)

where $s_\mu$ denotes the Schur polynomial. A lesser known fact is that for $n_1 > n_2$ the joint probability that an $n_1 \times (n_2 + 1)$ non-negative integer matrix with elements chosen according to (2.1) corresponds to a pair of semi-standard tableaux with shape $\mu$, content $n_1$ and $n_2 + 1$, and the bottom left sub-matrix corresponds to a pair of semi-standard tableaux with shape $\kappa$, content $n_1$ and $n_2$ is [17]

$$\Pr(\mu \cap \kappa) = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2+1} (1 - a_i b_j)s_\mu(a_1, \ldots, a_{n_1})s_\kappa(b_1, \ldots, b_{n_2+1})b_{n_2+1}^{\sum_{j=1}^{n_2+1}(\mu_j - \kappa_j) + \mu_{n_2+1}}\chi(\mu > \kappa)$$ \hspace{1cm} (2.5)

where, with $\chi(A) = 1$ if $A$ is true, $\chi(A) = 0$ otherwise,

$$\chi(\mu > \kappa) := \chi(\mu_1 \geq \kappa_1 \geq \mu_2 \geq \kappa_2 \geq \cdots \geq \mu_{n_1} \geq \kappa_{n_1} \geq 0).$$ \hspace{1cm} (2.6)
It follows from (2.4) and (2.5) that given the pair of semi-standard tableaux corresponding to an $n_1 \times n_2$ matrix ($n_1 > n_2$) has shape $\kappa$, the probability of the pair of semi-standard tableaux corresponding to the $n_1 \times (n_2 + 1)$ matrix, obtained by adding an extra row to the existing matrix, having shape $\mu$ is

$$\Pr(\mu | \kappa) := \chi(\mu > \kappa) \prod_{i=1}^{n_1} (1 - a_i b_{n_2+1}) \frac{s_{\mu}(a_1, \ldots, a_{n_1})}{s_{\kappa}(a_1, \ldots, a_{n_1})} b_{n_2+1}^{\sum_{j=1}^{n_2}(\mu_j - \kappa_j) + \mu_{n_2+1}}. \tag{2.7}$$

Let us now seek the joint probability that with $n_1 \geq n_2 + p$, an $n_1 \times (n_2 + p)$ non-negative integer matrix with elements chosen according to (2.1) is such that the principal $n_1 \times (n_2 + s)$ sub-blocks ($s = 0, 1, \ldots, p$) correspond to pairs of semi-standard tableaux with shape $\mu^{(s)}$. This is computed from (2.4) and (2.7) according to

$$\Pr(\mu^{(0)}(0)) \prod_{s=0}^{p-1} \Pr(\mu^{(s+1)}|\mu^{(s)}) = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2 + p} (1 - a_i b_j) s_{\mu^{(s)}}(a_1, \ldots, a_{n_1}) s_{\mu^{(0)}}(b_1, \ldots, b_{n_2})$$

$$\times \prod_{s=1}^{p} b_{n_2+s}^{-\sum_{j=1}^{n_2+s}(\mu^{(s)}_j - \mu^{(s-1)}_j) + \mu_{n_2+s}} \chi(\mu^{(s)} > \mu^{(s-1)}). \tag{2.8}$$

### 2.2. Specializing the parameters

In [17] the joint probability (2.5) is specialized to the case of a geometrical progression of parameters

$$(a_1, \ldots, a_{n_1}) = (z, zt, zt^2, \ldots, zt^{n_1-1})$$

$$(b_1, \ldots, b_{n_2}) = (z, zt, zt^2, \ldots, zt^{n_2-1}). \tag{2.9}$$

This is a preliminary step for taking the so-called Jacobi limit, in which a joint probability of a type known from random matrix theory is obtained. In this subsection the parameters will be specialized according to (2.9) for the more general joint probability (2.8).

Let $\ell(\mu)$ refer to the number of non-zero parts in the partition $\mu$. Now for (2.8) to be non-zero we require $\ell(\mu^{(p)}) \leq n_2 + p$. Under this circumstance, we deduce from [17, equation (2.26)] with $\kappa \mapsto \mu^{(p)}$, $n_2 \mapsto n_1$, $n_1 \mapsto n_2 + p$, $r_j \mapsto h_j^{(p)} := \mu_j^{(p)} + n_2 + p - j$ that

$$s_{\mu^{(p)}}(1, t, \ldots, t^{n_1-1}) = t^{-\sum_{j=1}^{n_2-p}(j-1+n_2+p)} \prod_{l=1}^{n_1-1}(t; t)_l \prod_{i=1}^{n_1-1-n_2-p+1}(t; t)_l$$

$$\times \prod_{i=1}^{n_2+p} \frac{(t; t)_{k_i^{(p)}+n_1-n_2-p}}{(t; t)_{k_i^{(p)}}} \prod_{1 \leq i < j \leq n_2+p} (t^{k_j^{(p)}} - t^{k_i^{(p)}}), \tag{2.10}$$

where

$$(a; q)_l := \prod_{j=1}^{l} (1 - aq^{j-1}).$$

This makes explicit the first Schur polynomial factor in (2.8). In regards to the second Schur polynomial factor, making use of [17, equation (2.18)] with $n \mapsto n_2$, $n^* \mapsto n_2$,

$${\text{doi}}:10.1088/1742-5468/2011/08/P080115$$
\[ \lambda \mapsto \mu(0), \quad h_j \mapsto h_j(0) := \mu_j(0) + n_2 - j, \] gives
\[ s_{\mu(0)}(1, t, \ldots, t^{n_2-1}) = \frac{t^{-\sum_{j=1}^{n_2}(j-1)(n_2-j)}}{\prod_{l=1}^{n_2-1}(t; t)_l} \prod_{1 \leq i < j \leq n_2} (t^{h_j(0)} - t^{h_i(0)}). \tag{2.11} \]

Use of these results in (2.8), the fact that the Schur polynomial \( s_n \) is homogeneous of degree \( |k| \), and furthermore setting \( b_{n_2+s} = \alpha_s \) \((s = 1, \ldots, p)\), allows the following results to be deduced.

**Proposition 1.** Let \( k \in \mathbb{Z}_{\geq 0} \) and \( n_1 \geq n_2 + p \). On each site of the \( n_1 \times (n_2 + p) \) square lattice specify a random non-negative integer \( x_{i,j} \) according to the specification
\[ \Pr(x_{i,j} = k) = (1 - z^2t^{i+j-2})(z^2t^{i+j-2})^k, \quad j \leq n_2 \]
\[ \Pr(x_{i,n_2+s} = k) = (1 - \alpha_s z t^{-1})(\alpha_s z t^{-1})^k \quad (s = 1, \ldots, p). \]

Introduce the notations \( h_i^{(s)} := \mu_i^{(p)} + n_2 + p - i \) and
\[ \tilde{\chi}(h^{(p)}, h^{(p-1)}) := \chi(h_1^{(p)} \geq h_2^{(p-1)} > h_2^{(p)} \geq h_2^{(p-1)} > \cdots > h_{n_2+p-1}^{(p-1)} > h_{n_2+p}. \]

In this setting, the joint probability that the matrix \( [x_{i,j}]_{n_1 \times (n_2+p)} \) is such that the submatrices \( [x_{i,j}]_{n_1 \times (n_2+s)} \) \((s = 0, \ldots, p)\) correspond, under RSK, to pairs of tableaux of shape \( \mu^{(s)} \) has the explicit form
\[ K_{n_1,n_2,p}(\{\alpha_s\}, z, t) = z^{-2\sum_{j=1}^{n_2} (n_2-p+j)} \frac{\prod_{s=1}^{p} \alpha_s^{(n_2-s+1)}}{\prod_{1 \leq i < j \leq n_2} (t^{h_i(0)} - t^{h_j(p)})} \]
\[ \times \prod_{i=1}^{n_2+p} \frac{(t; t)_{h_i^{(p)}+n_1-n_2-p}}{(t; t)_{h_i^{(p)}}} \prod_{1 \leq i < j \leq n_2} (t^{h_j(0)} - t^{h_i(p)}) \tag{2.12} \]

with
\[ K_{n_1,n_2,p}(\{\alpha_j\}, z, t) = \frac{\prod_{s=1}^{p} \alpha_s^{(n_2-s+1)}}{t^{-\sum_{j=1}^{n_1-n_2-p} j(j-1)}} \]
\[ \times \frac{t^{-(n_2+p)} \prod_{j=1}^{n_1-n_2-p} (1-z^2t^{i+j-2})}{\prod_{l=1}^{n_2-1}(t; t)_l} \prod_{l=1}^{n_1-1}(t; t)_l \]
\[ \times \prod_{i=1}^{n_1-n_2-p} \prod_{j=1}^{n_2} (1 - z^2t^{i+j-2})(1 - \alpha_s z t^{-1}). \]

### 2.3. The Jacobi limit

The Jacobi limit of the setting of proposition 1 corresponds to each site of the \( n_1 \times (n_2 + p) \) square lattice specifying a non-negative continuous exponential random variable with site-dependent variance \( j \leq n_2 \) given by
\[ \Pr(x_{i,j} \in [y, y + dy]) = (i + j - 2 + 2a)e^{-y(i+j-2+2a)} dy, \quad j \leq n_2 \]
\[ \Pr(x_{i,n_2+s} \in [y, y + dy]) = (i + 1 + a + a_s)e^{-y(i+1+a+a_s)} dy, \quad (s = 1, \ldots, p). \tag{2.13} \]
This can be obtained from (2.12) by setting
\[ t = e^{-1/L}, \quad z = e^{-a_s/L}, \quad \alpha_s = e^{-a_s/L}, \quad k_j^{(s)} / L = x_j^{(s)} \] (2.14)
and taking the limit \( L \to \infty \) (see proposition 1 of [17] for a similar computation). The joint probability (2.12), scaled by multiplying by \( L^{(1+p)(n_2+p/2)} \), has the following limiting form.

**Corollary 1.** The PDF obtained by the limiting procedure (2.14) applied to (2.12) is equal to
\[
\hat{K}_{n_1,n_2,p}(\{a_j\}, a) e^{-a_s(\sum_{j=1}^{n_2+p} x_j^{(p)} + \sum_{j=1}^{n_2} x_j^{(0)})} \prod_{s=1}^{p} e^{-a_s(\sum_{j=1}^{n_2+s-1} x_j^{(s)} - \sum_{j=1}^{n_2+s-1} x_j^{(s-1)})} \chi(x(s) > x^{(s-1)})
\times \prod_{i=1}^{n_2+p} (1 - e^{-x_i^{(s)}})^{n_1-n_2-p} \prod_{1 \leq i < j \leq n_2+p} (e^{-x_j^{(p)}} - e^{-x_i^{(p)}}) \prod_{1 \leq i < j \leq n_2} (e^{-x_j^{(0)}} - e^{-x_i^{(0)}})
\]
where
\[
\hat{K}_{n_1,n_2,p}(\{a_j\}, a) = \frac{\prod_{i=1}^{n_1-n_2-p-1} \Gamma(\mathcal{a})}{(\prod_{i=1}^{n_1-n_2-p-1} \Gamma(\mathcal{a}))} \prod_{s=1}^{p} \frac{\Gamma(\mathcal{a}_s + a + n_1)}{\Gamma(\mathcal{a}_s + a)} \prod_{i=1}^{n_2} \frac{\Gamma(2\mathcal{a} + i + n_2 - 1)}{\Gamma(2\mathcal{a} + i - 1)}
\]
and
\[
\chi(x^{(s)}) > x^{(s-1)} := \chi(x_1^{(s)}) > x_1^{(s-1)} > \cdots > x_{n_2+s-1}^{(s)} > x_{n_2+s}^{(s)} > 0).
\]

Changing variables \( e^{-x_j^{(s)}} = y_j^{(s)} \) the PDF (2.15) is
\[
\hat{K}_{n_1,n_2,p}(\{a_j\}, a) \prod_{i=1}^{n_2+p} (y_i^{(p)})^{a_1-1} (1 - y_i^{(p)})^{n_1-n_2-p} \prod_{j=1}^{n_2} (y_j^{(0)})^{a_1} \prod_{s=1}^{p} \chi(y^{(s)} < y^{(s-1)})
\times \prod_{j=1}^{n_2+s-1} \frac{(y_j^{(s)})^{a_s}}{(y_j^{(s-1)})^{a_s}} \prod_{1 \leq i < j \leq n_2+p} (y_j^{(p)} - y_i^{(p)}) \prod_{1 \leq i < j \leq n_2} (y_j^{(0)} - y_i^{(0)})
\]
where
\[
\chi(y^{(s)} < y^{(s-1)}) := \chi(0 < y_1^{(s)} < y_1^{(s-1)} < \cdots < y_{n_2+s-1}^{(s-1)} < y_{n_2+s-1}^{(s)} < 1).
\]
In the special case \( a_s = a - j \ (s = 1, \ldots, p) \) this simplifies to the functional form
\[
\frac{1}{C} \prod_{l=1}^{n_2+p} w(y_l^{(p)}) \prod_{1 \leq i < j \leq n_2+p} (y_j^{(p)} - y_i^{(p)}) \prod_{1 \leq i < j \leq n_2} (y_j^{(0)} - y_i^{(0)}) \prod_{s=1}^{p} \chi(y^{(s)} < y^{(s-1)}),
\]
with \( C \) the normalization (\( C \) will be used generally below for this purpose and so its explicit value may vary from equation to equation) and \( w(y) = y^\alpha (1 - y)^\beta \) for certain \( \alpha \) and \( \beta \). The latter is the Jacobi weight, which in a functional form involving Vandermonde products is typical of a PDF arising in random matrix theory. Indeed, this functional form for each of the Gaussian, Laguerre and Jacobi weights can be obtained as eigenvalue PDFs.

Before turning to such random matrix interpretations, it should be pointed out that in the setting of proposition 1, choosing each site of the \( n_1 \times (n_2 + p) \) square lattice according to continuous exponential random variables
\[
\Pr(x_{i,j} \in [y, y + dy]) = e^{-y} dy
\]
(2.20)
leads to (2.19) with the particular Laguerre weight \( w(y) = y^{n_1-(n_2+p)}e^{-y} \). Further, rescaling the variables \( y_j^{\langle p \rangle} \mapsto n_1(1+y_j^{\langle p \rangle}/\sqrt{2/n_1}) \) therein, and taking \( n_1 \to \infty \) gives (2.19) back with the Gaussian weight \( w(y) = e^{-y^2} \). In the case \( p = 1 \) both of these facts are explicitly demonstrated in [17, propositions 4 and 5]; the extension to general \( p \) involves straightforward limiting procedures applied to (2.17).

3. Joint eigenvalue PDF for some nested sequences of random matrices

3.1. Gaussian unitary ensemble

By definition, matrices \( M_N \) from the \( N \times N \) GUE satisfy the recurrence

\[
M_{N+1} = \begin{bmatrix} M_N & \vec{w} \\ \vec{w}^\dagger & a \end{bmatrix}
\]

where \( a \sim N[0,1/\sqrt{2}] \) and each component \( w_j \) of \( \vec{w} \) has distribution \( w_j \sim N[0,1/2] + iN[0,1/2] \). Further, with \( U_N \) the unitary matrix which diagonalizes \( M_N \), and thus \( M_N = U_ND_NU_N^\dagger \), where \( D_N \) is the diagonal matrix of the eigenvalues of \( M_N \), one has

\[
\begin{bmatrix} U_N & \vec{0} \\ \vec{0}^T & 1 \end{bmatrix} \begin{bmatrix} M_N & \vec{w} \\ \vec{w}^\dagger & a \end{bmatrix} \begin{bmatrix} U_N & \vec{0} \\ \vec{0}^T & 1 \end{bmatrix}^\dagger \sim \begin{bmatrix} D_N & \vec{w} \\ \vec{w}^\dagger & a \end{bmatrix}.
\]

This bordered form is the key to studying the joint distribution of the eigenvalues of the sequence of GUE matrices \( \{M_j\}_{j=1,2,...} \), or equivalently that of the sequence of principal sub-matrices of a single infinite GUE matrix [1, 17, 14].

In particular, it follows from (3.2) that the characteristic polynomials \( p_N(\lambda), p_{N+1}(\lambda) \) for \( M_N, M_{N+1} \) are related by

\[
\frac{p_{N+1}(\lambda)}{p_N(\lambda)} = \lambda - a - \sum_{i=1}^{N} \frac{|w_i|^2}{\lambda - \lambda_i^{(N)}}
\]

where \( \{\lambda_i^{(N)}\} \) denotes the eigenvalues of \( M_N \) assumed ordered

\[
\lambda_1^{(N)} < \lambda_2^{(N)} < \cdots < \lambda_N^{(N)}.
\]

With \( \{\lambda_i^{(N)}\} \) regarded as given the PDF for the zeros of this random rational function, and thus the PDF for the distribution of the eigenvalues \( \{\lambda_i^{(N+1)}\} \) of \( M_{N+1} \), can be computed to be equal to

\[
e^{-\sum_{i=1}^{N+1} (\lambda_i^{(N+1)})^2} \prod_{1 \leq j < k \leq N+1} \frac{(\lambda_i^{(N+1)} - \lambda_j^{(N+1)})}{(\lambda_i^{(N+1)} - \lambda_k^{(N+1)})} \chi(\lambda^{(N+1)} < \lambda^{(N)})
\]

where \( \chi(\lambda^{(N+1)} < \lambda^{(N)}) \) is specified as in (2.18) except that the first and last inequalities are replaced by \(-\infty < \) and \( < \infty \), respectively. From this result, together with the fact that the eigenvalue PDF for \( N \times N \) GUE matrices is proportional to

\[
\prod_{i=1}^{N} e^{-(\lambda_i^{(N)})^2} \prod_{1 \leq j < k \leq N} (\lambda_k^{(N)} - \lambda_j^{(N)})^2
\]
it follows that the joint eigenvalue PDF for the sequence of GUE matrices \(\{M_n\}_{n=N,\ldots,N+p}\) is given by (2.19) with \(n_2 = N\) and \(w(y) = e^{-y^2}\), and the definition (2.16) of \(\chi(y^{(s)}) < y^{(s-1)}\) modified appropriately, and \(y_k^{(s)} := \lambda_k^{(s)}\).

### 3.2. Laguerre unitary ensemble

Matrices \(A_{(n)}\) from the \(N \times N\) LUE with parameter \(a = n - N\) are constructed from \(n \times N\) rectangular complex Gaussian matrices \(X_{(n)}\) with entries \(N[0, 1/\sqrt{2}] + iN[0, 1/\sqrt{2}]\) according to \(A_{(n)} = X_{(n)}^\dagger X_{(n)}\). Such matrices satisfy the recurrence

\[
A_{(n+1)} = A_{(n)} + \bar{x}x^\dagger, \quad A_{(0)} = [0]_{N \times N} \tag{3.3}
\]

where \(\bar{x}\) is an \(N \times 1\) column vector of complex Gaussians. Our interest is in the case \(n \leq N\) of \(A_{(n)}\), for which there are \(N - n\) zero eigenvalues. Then, after making use too of the invariance of \(\bar{x}x^\dagger\) by a unitary similarity transformation, the recursion (3.3) can be written in the equivalent form

\[
A_{(n+1)} = \text{diag}(a_1, \ldots, a_n, 0, \ldots, 0) + \bar{x}x^\dagger \tag{3.4}
\]

where \(\{a_i\}_{i=1,\ldots,n}\) are the non-zero eigenvalues of \(A_{(n)}\), assumed ordered so that \(0 < a_1 < \cdots < a_n\). From (3.4) it follows that the corresponding characteristic polynomials are such that

\[
\frac{\det(\lambda 1_N - A_{(n+1)})}{\det(\lambda 1_N - A_{(n)})} = 1 - \sum_{j=1}^{n} \frac{|x_j|^2}{\lambda - a_j} - \sum_{j=n+1}^{N} \frac{|x_j|^2}{\lambda} \tag{3.5}
\]

The conditional PDF for the zeros of this random rational function, and thus the conditional PDF for the eigenvalues of \(A_{(n+1)}\), can be computed exactly as [17, 14]

\[
\prod_{i=1}^{n+1} \chi_j^{N-(n+1)} \frac{1}{a_j} e^{-\sum_{j=1}^{n+1} \lambda_j + \sum_{j=1}^{n} a_j} \prod_{i=1}^{n+1} \frac{(\lambda_i - \lambda_j)}{(a_i - a_j)} \chi(\lambda < a)
\]

Recalling that the non-zero eigenvalue PDF for the matrices \(A_{(n)}\) is proportional to

\[
\prod_{j=1}^{n} a_j^{N-n} e^{-a_j} \prod_{i<k} (a_i - a_k)^2
\]

it follows that the joint eigenvalue PDF for the sequence of LUE matrices \(\{A_{(r)}\}_{r=n,\ldots,n+p}\) with \(n + p \leq N\) is given by (2.19) with \(n_2 = n\), \(w(y) = y^{N-(n+p)}e^{-y}\), and where \(y_k^{(s)}\) denotes the \(k\)th eigenvalue of \(A_{(s)}\).

### 3.3. Corank-1 random projections

Let \(A_n\) be an \(n \times n\) matrix with eigenvalues \(a_1 < a_2 < \cdots < a_n\), and let \(\bar{x}\) be an \(n \times 1\) random complex Gaussian normalized column vector. The matrix

\[
M_n := \Pi_n A_n \Pi_n, \quad \Pi_n := 1_n - \bar{x}\bar{x}^\dagger \tag{3.6}
\]

doi:10.1088/1742-5468/2011/08/P08011
then represents a corank-1 random projection of $A_n$. We know from [17] that in general $M_n$ has a single zero eigenvalue, while the non-zero eigenvalues $\{\lambda_j\}_{j=1,\ldots,n-1}$ have the conditional PDF
\begin{equation}
(n-1)! \prod_{i<j} (\lambda_j - \lambda_k)n \prod_{i<j} (a_j - a_k) \chi(a < \lambda).
\end{equation}

Introduce a nested sequence of matrices $\{A_i\}_{i=1,\ldots,n}$ by setting $A_{n-1}$ equal to the diagonal matrix formed from the non-zero eigenvalues of $M_n$, computing $M_{n-1}$ according to (3.6), setting $A_{n-2}$ equal to the diagonal matrix formed from the non-zero eigenvalues of $M_{n-1}$, and repeating. With the eigenvalues of $A_n$ having PDF proportional to
\begin{equation}
\prod_{i=1}^n w(a_i) \prod_{1 \leq j < k \leq n} (a_k - a_j)^2
\end{equation}
it follows immediately from (3.7) that the joint PDF for $\{A_i\}_{i=n-p,\ldots,n}$ is given by (2.19) with $n_2 + p = n$.

We remark (see, e.g., [11]) that the Laguerre and Jacobi cases can be realized by the matrix structure $A_n = X_n^T X_n$ for certain matrices (recall the discussion of section 3.2 for the Laguerre case). It follows that the above construction is then equivalent to applying a sequence of corank-1 projections directly to the matrix $X_n$.

4. Correlation for the classical projection process

4.1. Approach via a general formula

As a joint PDF, and with $p$ as a variable non-negative integer, we see that the most general form of the functional form (2.19) occurs with $n_2 = 1$. Thus the arbitrary $n_2 \in \mathbb{Z}^+$ case of (2.19) can be reclaimed from this by integrating over $y^{(0)}, y^{(1)}, \ldots, y^{(n_2-1)}$ and replacing $p \mapsto p + n_2 - 1, y^{(p)} \mapsto y^{(p-n_2)}$. Considering the case $n_2 = 1$, and after then changing notation by setting $p = N - 1, y^{(p)}_i \mapsto x^{(p+1)}_{p+2-i}$, (2.19) is
\begin{equation}
\frac{1}{C} \prod_{l=1}^{N} w(x_l^{(N)}) \prod_{1 \leq j < k \leq N} (x_j^{(N)} - x_k^{(N)}) \prod_{s=1}^{N-1} \chi(x^{(s+1)} > x^{(s)}).
\end{equation}

Here $\chi(x^{(s+1)} > x^{(s)})$ is defined as
\begin{equation}
\chi(x^{(s+1)} > x^{(s)}) = \chi(x^{(s+1)}_1 > x^{(s)}_1) > \cdots > x^{(s+1)}_s > x^{(s)}_s.
\end{equation}
(cf (2.16)) and $w(x)$ involves a factor $\chi(0 < x < 1)$. Hereafter we consider more general cases in which $w(x)$ does not always involve such a factor (see (4.2) below).

It is convenient to regard (4.1) as a multi-species system, with species $s$ consisting of $s$ particles. Our interest is in the corresponding correlation functions, obtained by fixing some subset of particles of species $j$ at position $y_j$ ($j = 1, \ldots, r$) and integrating over all other particles. This will be done for the three classical weight functions:
\begin{align}
w(y) = \begin{cases} 
e^{-y^2}, & \text{Gaussian} \\
y^a e^{-y} \chi(y > 0), & \text{Laguerre} \\
y^a (1 - y)^b \chi(0 < y < 1), & \text{Jacobi}.
\end{cases}
\end{align}

\textbf{doi:10.1088/1742-5468/2011/08/P08011}
Crucial to the calculation of the correlation functions is the fact that (4.1) can be written as a product of determinants. First, by making use of the Vandermonde determinant we see
\[
\prod_{l=1}^{N-1} w(x_i^{(N)}) \prod_{1 \leq j < k \leq N} (x_j^{(N)} - x_k^{(N)}) \propto \det \left[ w(x_k^{(N)})p_{N-j}(x_k^{(N)}) \right]_{j,k=1,\ldots,N},
\]
with \( \{p_j(x)\}_{j=0,\ldots,N-1} \) a set of arbitrary polynomials, \( p_j(x) \) of degree \( j \). Furthermore, we know from [16, lemma 1] that
\[
\chi(x^{(s+1)} > x^{(s)}) = \det[\chi(x^{(s+1)} > x^{(s)})]_{j,k=1,\ldots,s+1}
\]
where \( x^{(s+1)} := -\infty \). Consequently (4.1) can be written in the form
\[
\frac{1}{C} \prod_{s=1}^{N-1} \det[\phi(x^{(s)}, x^{(s+1)})]_{j,k=1,\ldots,s+1} \det[\Psi^{N}_{N-j}(x^{(N)})]_{j,k=1,\ldots,N}
\]
with
\[
\phi(x, y) := \chi(y > x), \quad \Psi^{N}_{N-j}(x) := w(x)p_j(x).
\]

The general structure (4.4) is precisely that for which the correlations have been determined in the recent work [4, lemma 3.4]. To apply this result, with \((a * b)(x, y) := \int_{-\infty}^{\infty} a(x, z)b(z, y) \, dz\), it is necessary to compute the quantities
\[
\phi^{(n_1,n_2)}(x, y) := \underbrace{\phi * \cdots * \phi}_{n_2-n_1 \text{ times}}(x, y), \quad n_1 < n_2
\]
(for \( n_1 \geq n_2, \phi^{(n_1,n_2)}(x, y) := 0 \)) and
\[
\Psi_{n-j}^{n}(x) := (\phi^{(n,N)} * \Psi^{N}_{N-j})(x) \quad (1 \leq n < N, j = 1, \ldots, N).
\]

Use of (4.5) shows
\[
\phi^{(n_1,n_2)}(x, y) = \frac{1}{(n_2 - n_1 - 1)!} (y - x)^{n_2 - n_1 - 1} \chi(y > x)
\]
(with the convention that \( 1/((-p)!) = 0 \) for \( p \in \mathbb{Z}^+ \), this vanishes for \( n_1 \geq n_2 \)) and
\[
\Psi_{n-j}^{n}(x) = \frac{1}{(N - n - 1)!} \int_{x}^{\infty} w(y)p_{N-j}(y)(y - x)^{N-n-1} \, dy.
\]

To proceed further, we choose \( w(y) \) to be one of the classical weight functions (4.2). We further choose \( p_j(y) \) to be proportional to the corresponding orthogonal polynomials, as specified by their Rodrigues formulae
\[
p_j(y) = \frac{1}{e_j w(y)} \frac{d^j}{dy^j} (w(y)(Q(y))^j) = \begin{cases} H_j(y), & \text{Gaussian} \\ L_j^{(a)}(y), & \text{Laguerre} \\ P_j^{(a,b)}(1 - 2y), & \text{Jacobi} \end{cases}
\]

\[\text{doi:10.1088/1742-5468/2011/08/P08011}\]
with the quantities $e_j$ and $Q(y)$ defined in the various cases by the pairs

$$ (e_j, Q(y)) = \begin{cases} 
( ( -1)^j, 1), & \text{Gaussian} \\
( j!, y), & \text{Laguerre} \\
(2^j j!, y(1-y)), & \text{Jacobi.} 
\end{cases} \quad (4.9) $$

Substituting (4.8) in (4.7) and integrating by parts shows that for $j \geq 0$ ($n \neq N$)

$$ \Psi_j^n(x) = (-1)^{N-n} \frac{e_j}{e_{N-n+j}} \begin{cases} 
\frac{1}{N_j} H_j(x), & \text{Gaussian} \\
\frac{1}{N_j} L_j^{(a+N-n)}(x), & \text{Laguerre} \\
\frac{1}{N_j} P_j^{(a+N-n, b+N-n)}(1-2x), & \text{Jacobi}, 
\end{cases} \quad (4.10) $$

while for $j < 0$

$$ \Psi_j^n(x) = (-1)^{N-n+j} \frac{1}{e_{N-n+j}} \int_\infty^\infty (y-x)^{-j-1} w(y)(Q(y))^{N-n+j} dy. \quad (4.11) $$

As further required by [4, lemma 3.4], one introduces the polynomials $\{\Phi_j^n(x)\}_{j=0,\ldots,n-1}$, $n = 1, \ldots, N - 1$ by the orthogonality requirement

$$ \int_{-\infty}^{\infty} \Phi_j^n(x) \Phi_k^n(x) \, dx = \delta_{j,k}. $$

From (4.10) we see that

$$ \Phi_j^n(x) = (-1)^{N-n} \frac{e_{N-n+j}}{e_j} \begin{cases} 
\frac{1}{N_j} H_j(x), & \text{Gaussian} \\
\frac{1}{N_j} L_j^{(a+N-n)}(x), & \text{Laguerre} \\
\frac{1}{N_j} P_j^{(a+N-n, b+N-n)}(1-2x), & \text{Jacobi}, 
\end{cases} \quad (4.12) $$

where

$$ \mathcal{N}_j = \begin{cases} 
2^j j! \sqrt{\pi}, & \text{Gaussian} \\
\frac{\Gamma(j+a+1)}{\Gamma(j+1)}, & \text{Laguerre} \\
\frac{\Gamma(j+a+1)\Gamma(j+b+1)}{j!(2j+a+b+1)\Gamma(j+a+b+1)}, & \text{Jacobi.} 
\end{cases} \quad (4.13) $$

is the normalization associated with each polynomial respectively.

A crucial feature exhibited by (4.12) is that $\Phi^n_0(x)$ is a constant. Now, Assumption (A) of [4, lemma 3.4] requires that $\Phi^{n+1}_0(x) \propto \phi(x^{(n)}_{n+1}, x)$. Recalling from below (4.3) that $x^{(n)}_{n+1} := -\infty$, we see from the first definition in (4.5) that indeed $\phi(x^{(n)}_{n+1}, x)$ is similarly a constant. With Assumption (A) satisfied, [4, equation (3.25)] gives the sought explicit form of the correlation function, which itself is a determinant.
Proposition 2. Consider the joint PDF (4.1), with \( w(y) \) one of the three classical weights (4.2). The correlation function for eigenvalues of species \( s_j \) at positions \( y_j \) \((j = 1, \ldots, r)\) has the determinant form

\[
\rho_{(r)}(\{(s_j, y_j)\}_{j=1,\ldots,r}) = \det[K(s_j, y_j; s_k, y_k)]_{j,k=1,\ldots,r}. \tag{4.14}
\]

Here the kernel \( K \) is given in terms of the quantities \( \phi^{(m_1,m_2)}(x,y), \Phi^\alpha_j(x), \Phi^\alpha_j(y) \) specified by (4.6), (4.10), (4.11) and (4.12) according to

\[
K(s_j, y_j; s_k, y_k) = -\phi^{(s_j,s_k)}(y_j, y_k) + \sum_{l=1}^{s_k} \Psi^j_{s_j-l}(y_j) \Phi^s_k(y_k). \tag{4.15}
\]

4.2. Direct approach

The method of [4, lemma 3.4] leading to Proposition 2 relies on first discretizing the domain of the support of the eigenvalues. When the joint PDF is of the determinantal form (4.14). However, only when Assumption (A) holds can the explicit form (4.15) of the correlation kernel \( K \) be demonstrated. The significance of Assumption (A) is that it is a sufficient condition for the explicit computation of a certain matrix inverse occurring in the calculation.

In this section, a different viewpoint on the structure underlying (4.15) will be given. For this we adopt the method of [29], which in turn builds on the pioneering work of Dyson [10]. The basic idea is to express the joint eigenvalue PDF as a single determinant. The elements of this determinant must be such that the successive integrations required to obtain the correlation functions which have the effect of reducing the size of the matrix, but preserving those elements which remain.

The starting point is (4.4), rewritten in the form

\[
\frac{1}{C} \prod_{s=2}^{N} \det\begin{bmatrix}
1_{(N-s)\times(N-s)} & 0_{(N-s)\times s} \\
0_{s\times(N-s)} & \left[ \phi(x_j^{(s-1)}, x_k^{(s)}) - \kappa_s(x_j^{(s-1)}) \right]_{j=1,\ldots,s-1}^{[1]_{k=1,\ldots,s}}
\end{bmatrix}
\times \det[\Psi^N_j(x_k^{(N)})]_{j,k=1,\ldots,N}, \tag{4.16}
\]

so that all determinants are of the same dimension. The auxiliary function \( \kappa_s(x) \) is arbitrary, the value of the determinant being independent of \( \kappa_s(x) \).

In the notation for \( w(y), p_j(y) \) and \( N_j \) as defined by (4.2), (4.8) and (4.13), introduce the superscript \((s)\) to indicate that \( a \mapsto a + N - s \) (Laguerre case, original weight \( x^a e^{-x} \)), \( a \mapsto a + N - s, b \mapsto b + N - s \) (Jacobi case, original weight \( x^a(1 - x)^b \)). In the Gaussian case the superscript has no effect. With this meaning understood, expanding \( \phi(x, y) = \chi(y > x) \) in terms of \( \{p_j^{(s)}(y)\}_{j=0,1,\ldots} \) gives

\[
\phi(x, y) = \sum_{k=0}^{\infty} \frac{p_k^{(s)}(y)}{N_k^{(s)}} \int_x^\infty w(x)(t)p_k^{(s)}(t) \, dt. \tag{4.17}
\]

doi:10.1088/1742-5468/2011/08/P08011
Separating off the $k=0$ term, making use of the Rodrigues formula (4.8) and integrating by parts reduces this to

$$
\phi(x, y) = \frac{1}{\mathcal{N}_0^{(s)}} \int_x^\infty w^{(s)}(t) \, dt - w^{(s-1)}(x) \sum_{k=0}^{\infty} \frac{e_k}{e_{k+1} \mathcal{N}_{k+1}^{(s)}} p_k^{(s-1)}(x) p_{k+1}^{(s)}(y). \tag{4.18}
$$

Recalling from (4.16) that $\kappa_s(x)$ is to be subtracted from $\phi(x, y)$, the formula (4.18) suggests choosing

$$
\kappa_s(x) = \frac{1}{\mathcal{N}_0^{(s)}} \int_x^\infty w^{(s)}(t) \, dt. \tag{4.19}
$$

Making this choice, and with the notation

$$
\eta_k^{(s)}(x) = \left( \frac{w^{(s)}(x)}{\mathcal{N}_k^{(s)}} \right)^{1/2} p_k^{(s)}(x) \tag{4.20}
$$

(note that $\{\eta_k^{(s)}(x)\}_{k=0,1,\ldots}$ is a set of orthonormal functions) we obtain for (4.16) the expression

$$
\frac{1}{C} \prod_{s=2}^N \det \begin{bmatrix}
1_{(N-s)\times(N-s)} & 0_{(N-s)\times s} \\
0_{s\times(N-s)} & \tilde{\phi}_s(x_j^{(s-1)}, x_k^{(s)})
\end{bmatrix} \det[\tilde{\phi}_{j-1}^{(N)}(x_k^{(N)})]_{j,k=1,\ldots,N} \tag{4.21}
$$

where, with

$$
\gamma_j^{(t)} := e_j(\mathcal{N}_j^{(t)})^{1/2}, \tag{4.22}
$$

the quantity $\tilde{\phi}_s(x, y)$ is specified by

$$
\tilde{\phi}_s(x, y) := \sum_{k=0}^{\infty} \frac{\gamma_k^{(s-1)}}{\gamma_{k+1}} \eta_k^{(s-1)}(x) \eta_{k+1}^{(s)}(y)
$$

and for convenience the final row in the bottom right block of the first determinant has been moved to the first row.

Our next step is to introduce

$$
\eta_{j,l}^{(s)} := \begin{cases} 
\eta_j^{(s)}(x_l^{(s)}), & j \geq 0, \ l \geq 1 \\
\delta_{j,l-1}, & \text{otherwise},
\end{cases}
$$

extend the definition of $\gamma_j^{(t)}$ so that $\gamma_j^{(t)} := 1$ for $j < 0$, and in terms of these quantities to define

$$
A_{j,l}^{(s,t)} := \sum_{k=0}^{N-1} \frac{\gamma_k^{(s-N)}}{\gamma_{k+t-N}} \eta_k^{(s-N,j+s-N)} \eta_{k+t-N,l+t-N}^{(t)}
$$

$$
C_{j,l}^{(s,t)} := \sum_{k=0}^{\infty} \frac{\gamma_k^{(s-N)}}{\gamma_{k+t-N}} \eta_k^{(s-N,j+s-N)} \eta_{k+t-N,l+t-N}^{(t)} \tag{4.23}
$$

The importance of the two distinct ranges of summation exhibited in (4.23) for the calculation of dynamical correlations for determinantal PDFs has already been revealed in [29].

doi:10.1088/1742-5468/2011/08/P08011
We can write (4.21) in terms of \( \{A^{(s,t)}_{j,l}\} \), \( \{G^{(s,t)}_{j,l}\} \) so that it is

\[
\frac{1}{C} \det[A^{(N,1)}_{j,l}]_{j,l=1,...,N} \prod_{s=2}^{N} \det[G^{(s-1,s)}_{j,l}]_{j,l=1,...,N} = \frac{1}{C} \det \begin{bmatrix}
A^{(N,1)} & A^{(N,2)} & A^{(N,3)} & \ldots & A^{(N,N)} \\
0 & -G^{(1,2)} & -G^{(1,3)} & \ldots & -G^{(1,N)} \\
0 & 0 & -G^{(2,3)} & \ldots & -G^{(2,N)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & -G^{(N-1,N)}
\end{bmatrix}
\tag{4.24}
\]

where \( A^{(s,t)} := [A^{(s,t)}_{j,l}]_{j,l=1,...,N} \), \( G^{(s,t)} := [G^{(s,t)}_{j,l}]_{j,l=1,...,N} \). With \( \alpha^{(s,t)} \) the \( N \times N \) matrix such that \( \alpha^{(s,t)} A^{(t,u)} = A^{(s,u)} \), multiply row 1 by \( \alpha^{(j-1,N)} \) and add to row \( j \) \( (j = 2, \ldots, N) \) to rewrite this as

\[
\frac{1}{C} \det \begin{bmatrix}
A^{(N,1)} & A^{(N,2)} & A^{(N,3)} & \ldots & A^{(N,N)} \\
A^{(1,1)} & B^{(1,2)} & B^{(1,3)} & \ldots & B^{(1,N)} \\
A^{(2,1)} & A^{(2,2)} & B^{(2,3)} & \ldots & B^{(2,N)} \\
A^{(3,1)} & A^{(3,2)} & A^{(3,3)} & \ldots & B^{(3,N)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A^{(N-1,1)} & A^{(N-1,2)} & A^{(N-1,3)} & \ldots & B^{(N-1,N)}
\end{bmatrix}
\]

where \( B^{(s,t)} := A^{(s,t)} - G^{(s,t)} \). Moving the first block-row to the final block-row gives the structured formula

\[
\frac{1}{C} \det[F^{(s,t)}]_{s,t=1,...,N}, \quad F^{(s,t)} := \begin{cases} 
A^{(s,t)}, & s \geq t \\
B^{(s,t)}, & s < t.
\end{cases}
\tag{4.25}
\]

Moreover, from the definition of \( A^{(s,t)} \) and \( G^{(s,t)} \) we observe that for \( s \geq t \)

\[
F^{(s,t)} \propto \delta_{j,l}, \quad j \leq N - s \quad \text{or} \quad l \leq N - s
\]

while for \( s < t \)

\[
F^{(s,t)} = 0, \quad j \leq N - s \quad \text{or} \quad l \leq N - t.
\]

In this latter circumstance only one term contributes to each of \( A^{(s,t)} \) and \( G^{(s,t)} \) and it is the same for both. These special values allow the dimension of the block matrix in (4.25) to be reduced, giving for the joint PDF (4.1), \( p \) say,

\[
p(\bar{x}^{(1)}, \ldots, \bar{x}^{(N)}) = \frac{1}{C} \det[f^{(s,t)}]_{s,t=1,...,N}
\tag{4.26}
\]

where \( \bar{x}^{(j)} = (x^{(j)}_1, \ldots, x^{(j)}_j) \) and \( f^{(s,t)} \) is the \( s \times t \) matrix with entries

\[
f^{(s,t)}_{j,l} = f^{(s,t)}_{j-s+N, t-l+N} = \begin{cases} 
\sum_{k=1}^{t} \gamma^{(s)}_{k} \eta^{(s)}_{s-k}(x^{(s)}_j) \eta^{(t)}_{t-k}(x^{(t)}_l), & s \geq t \\
- \sum_{k=-\infty}^{t} \gamma^{(s)}_{k} \eta^{(s)}_{s-k}(x^{(s)}_j) \eta^{(t)}_{t-k}(x^{(t)}_l), & s < t.
\end{cases}
\tag{4.27}
\]
Determinantal correlations for classical projection processes

From the orthonormality of \( \{ \eta_k^{(s)}(x) \} \) it follows from (4.26) that

\[
\int_{-\infty}^{\infty} \frac{f_{j,l}^{(s,t)} f_{l,m}^{(t,u)} \, dx_l^{(t)}}{\frac{1}{2} \sum_{k=1}^{t} e_{t-k} \rho_{s-k}^{(s)}(x_j) \rho_{t-k}^{(t)}(x_l)} = \begin{cases} 
f_{j,m}^{(s,u)}, & s \geq t \geq u \text{ or } s < t < u \\
0, & \text{otherwise.} 
\end{cases} 
\] (4.28)

We seek to use the form (4.26), together with the property (4.28), to compute the correlation between eigenvalues of species \( s_j \) at positions \( y_j \) \((j = 1, \ldots, r)\). For this purpose, we group together the eigenvalues of distinct species in the correlation. Thus if the distinct species are \( \hat{s}_1, \ldots, \hat{s}_r \), with \( \hat{s}_1, \ldots, \hat{s}_s \in \{s_1, \ldots, s_r\} \), we write the positions being observed in species \( \hat{s} \) as \( \vec{x}^{(\hat{s})} := (x_1^{(\hat{s})}, \ldots, x_n^{(\hat{s})}) \) \((1 \leq n_\hat{s} \leq \hat{s})\). The correlation relating to \( \{\vec{x}^{(\hat{s})}\}_{\hat{s}=1,\ldots,\hat{r}} \) is specified in terms of the PDF \( p \) by

\[
\rho(\vec{r})(\{\vec{x}^{(\hat{s})}\}_{\hat{s}=1,\ldots,\hat{r}}) = \det \left[ \int_{S_{\hat{s} \setminus \{\hat{s}\}}} dx_1^{(\hat{s})} \cdots \int_{S_{\hat{s} \setminus \{\hat{s}\}}} dx_n^{(\hat{s})} \rho_{s-a}^{(a)}(\hat{s}_a - n_{\hat{s} - a}) \right] 
\times \int dx_{n_{\hat{s} - a} + 1} \cdots \int dx_{n_{\hat{s}}} p(\vec{x}^{(1)}, \ldots, \vec{x}^{(N)}).
\] (4.29)

Because of the structure (4.26) and the orthogonality relation (4.28), these integrals can all be computed by performing a Laplace expansion of the determinant (see, e.g., [29, 11]) to give

\[
\rho(\vec{r})(\{\vec{x}^{(\hat{s})}\}_{\hat{s}=1,\ldots,\hat{r}}) = \det \left[ \int_{J_{\hat{s}}} dx_1^{(\hat{s})} \cdots \int_{J_{\hat{s}}} dx_n^{(\hat{s})} \right] 
\times \int dx_{n_{\hat{s} - a} + 1} \cdots \int dx_{n_{\hat{s}}} p(\vec{x}^{(1)}, \ldots, \vec{x}^{(N)}).
\]

In the notation of (4.14) this can equivalently be written as

\[
\rho(\vec{r})(\{\vec{x}^{(\hat{s})}\}_{\hat{s}=1,\ldots,\hat{r}}) = \det [f_{j,k}^{(s,t)}]_{j,k=1,\ldots,r}
\] (4.30)

(note that the superscript on \( x_j^{(s,t)} \) is now redundant, and hence has been omitted) so to complete our task of rederiving (4.14) it is sufficient to show that

\[
f_{j,k}^{(s,t)} = a(s_j, x_j) K(s_j, x_j; s_k, x_k)
\] (4.31)

(the corresponding determinant is independent of the function \( a(s, x) \)).

To verify (4.31) we begin by recalling (4.20) and (4.22) to see from the definition (4.27) that for \( s \geq t \)

\[
f_{j,l}^{(s,t)} = (w^{(s)}(x_j) w^{(t)}(x_l))^{1/2} \int_{-\infty}^{\infty} \sum_{k=1}^{t} e_{t-k} \rho_{s-k}^{(s)}(x_j) p_{t-k}^{(t)}(x_l) N_{t-k}^{(t)}
\]

and for \( s < t \)

\[
f_{j,l}^{(s,t)} = -(w^{(s)}(x_j) w^{(t)}(x_l))^{1/2} \int_{-\infty}^{\infty} \sum_{k=-\infty}^{0} e_{t-k} \rho_{s-k}^{(s)}(x_j) p_{t-k}^{(t)}(x_l) N_{t-k}^{(t)}.
\]

(4.33)

On the other hand, it follows from (4.8) and (4.10) that for \( j \geq 0 \) \((n \neq N)\)

\[
\Psi_j^{(n)}(x) = (-1)^{N-n} e_j^{(n)} w^{(n)}(x) p_j^{(n)}(x),
\]

\(\text{doi:10.1088/1742-5468/2011/08/P08011}\)
Determinantal correlations for classical projection processes

J. Stat. Mech. (2011) P08011

while according to (4.12)
\[
\Phi_j^a(x) = (-1)^{N-n} \frac{e_{N-n+j} P_j^{(n)}(x)}{e_j \mathcal{N}_j^{(n)}}. 
\]  

(4.35)

To proceed further, we must distinguish the cases \( s \geq t \) and \( s < t \). For \( s \geq t \), \( \phi^{(s,t)}(x, y) = 0 \), and this together with (4.34) and (4.12) substituted in (4.15) gives
\[
K(s, x_j; t, x_l) = (-1)^{s-t} w^{(s)}(x_j) \sum_{k=1}^{t} \frac{e_{s-k} p_{s-k}^{(s)}(x_j) p_{t-k}^{(t)}(x_l)}{\mathcal{N}_{t-k}^{(t)}}. 
\]  

(4.36)

The case \( s < t \) requires more calculation. We then have
\[
\phi^{(s,t)}(x_j, x_l) = \frac{1}{(t-s-1)!} (x_l - x_j)^{t-s-1} \chi(x_l > x_j). 
\]

Analogous to (4.17), we can expand \( \phi^{(s,t)} \) in terms of \( \{ p_k^{(t)}(y) \} \):
\[
\phi^{(s,t)}(x, y) = \frac{1}{(t-s-1)!} \sum_{k=0}^{\infty} \frac{p_k^{(t)}(y)}{\mathcal{N}_k^{(t)}} \int_x^\infty w^{(t)}(u)(u-x)^{t-s-1-k} du. 
\]

(4.37)

Proceeding now as in the derivation of (4.18) gives
\[
\phi^{(s,t)}(x, y) = (-1)^{t-s} w^{(s)}(x) \sum_{k=-\infty}^{s} \frac{e_{s-k} p_{s-k}^{(s)}(x) p_{t-k}^{(t)}(y)}{\mathcal{N}_{t-k}^{(t)}} 
\]
\[
+ \sum_{k=0}^{t-s-1} \frac{(-1)^k}{(t-s-k-1)!} \frac{p_k^{(t)}(y)}{e_k \mathcal{N}_k^{(t)}} \int_x^\infty w^{(t)}(u)(u-x)^{t-s-1-k} du. 
\]

(4.38)

We must also note from (4.11) that for \( j < 0 \)
\[
\Psi_j^{a}(x) = \frac{(-1)^{N-n+j}}{e_{N-n+j}} \frac{1}{(j-1)!} \int_x^\infty (u-x)^{-j-1} w^{(n-j)}(u) du. 
\]

This together with (4.35) allows us to conclude
\[
\sum_{p=s+1}^{t} \Psi_{s-p}^{a}(x) \Phi_{t-p}^{a}(y) = \sum_{k=-\infty}^{t-s-1} \frac{(-1)^k}{(t-s-k-1)!} \frac{p_k^{(t)}(y)}{e_k \mathcal{N}_k^{(t)}} \int_x^\infty (u-x)^{t-s-1-k} w^{(t-k)}(u) dy, 
\]

(4.39)

while it is immediate from (4.34) and (4.35) that
\[
\sum_{p=1}^{s} \Psi_{s-p}^{a}(x) \Phi_{t-p}^{a}(y) = (-1)^{s-t} w^{(s)}(x) \sum_{k=1}^{t} \frac{e_{s-k} p_{s-k}^{s}(x) p_{t-k}^{t}(y)}{\mathcal{N}_{t-k}^{(t)}}. 
\]

(4.40)

Subtracting (4.37) from the sum of (4.38) and (4.39) shows that for \( s < t \)
\[
K(s, x_j; t, x_l) = (-1)^{t-s-1} w^{(s)}(x_j) \sum_{k=-\infty}^{0} \frac{e_{s-k} p_{s-k}^{s}(x_j) p_{t-k}^{t}(x_l)}{\mathcal{N}_{t-k}^{(t)}}. 
\]

(4.41)

Comparing (4.32) with (4.36), and (4.33) with (4.40), we see that for general \( s, t \)
\[
f_{j, d}^{(s, t)} = (-1)^{s-t} \left( \frac{w^{(t)}(x_l)}{w^{(s)}(x_j)} \right)^{1/2} K(s, x_j; t, x_l), 
\]

thus verifying (4.31).
5. Scaling limits

5.1. Introductory comments

It is well known (see, e.g., [12, 11]) that the eigenvalue distributions for the joint PDF (3.8) with the classical weights (4.2) permit three distinct scalings as $n \to \infty$. These correspond to eigenvalues in the bulk of the spectrum, or in the neighborhood of the spectrum edge. There are two distinct cases of the latter—the soft edge and the hard edge. The hard edge is characterized by the eigenvalue density being strictly zero on one side. This occurs for $x < 0$ in the Laguerre ensemble, and for both $x < 0$ and $x > 1$ in the Jacobi ensemble. In contrast, the neighborhood of the largest eigenvalue of the Laguerre and Gaussian ensembles is such that the eigenvalue density is to leading order in $n$ zero, but at higher order it is non-zero. This is referred to as a soft edge.

For the projection process (4.1) with classical weights (4.2) we again expect these same three distinct scalings. However, the scaled correlations depend not just on position but also the species. In taking the bulk and hard edge limits the species will be chosen to differ from $N$ by a constant. However, this is not an appropriate prescription at the soft edge, as can be seen by consideration of the joint PDF (2.19). One observes that the lowest indexed species repel via a Vandermonde factor, with no restoring potential apart from the ordering constraint. Thus they will tend to cluster at the boundaries, which at the soft edge corresponds to the positions of the species ($N$), and lock into these positions along all lines a finite value from $N$.

Insight into the correct choice of soft edge scaling of the species can be obtained by recalling theory relating to the queuing process of Baryshnikov [1], or equivalently a lattice version of the last passage percolation model of Hammersley (see, e.g., [13]). For definiteness, let us consider the latter. With each site $(i, j)$ in the quadrant $\mathbb{Z}^+ \times \mathbb{Z}^+$ an exponential random variable $x_{ij}$ of density $e^{-t}$, $t > 0$, define the stochastic variable

$$l(m, n) = \max_{(1,1) u/r (m,n)} \sum x_{ij}. \quad (5.1)$$

Here the sum is over all lattice paths in $\mathbb{Z}^+ \times \mathbb{Z}^+$ which start at $(1,1)$ and finish at $(m,n)$ going either one lattice site up (u) or one lattice site to the right (r).

It is well known that with $x_n := l(n, n)$ the variables $\{x_1, \ldots, x_n\}$ have a joint PDF of the form (3.8) with $w(x) = e^{-x}$ [21]. Hence, as $n \to \infty$ the corresponding distributions permit a soft edge scaling describing the scaled distribution of $l(n, n)$. This scaling is fully described by the correlation function

$$\rho_{(n)}(y_1, \ldots, y_n) = \det[K_{\text{soft}}(y_j, y_k)]_{j,k=1,\ldots,n} \quad (5.2)$$

where

$$K_{\text{soft}}(x, y) := \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}(y)\text{Ai}'(x)}{x - y} = \int_0^\infty \text{Ai}(x + u)\text{Ai}(y + u) \, du \quad (5.3)$$

is the so-called Airy kernel. It is also known [22] that the sequence of stochastic variables $\{l(n + k, n - k)\}_{k=0,1,\ldots}$ with $k = O(n^{2/3})$ permit a scaling to a state specified by the
the case of the soft edge the semi-infinite interval $x$ which is uniformly bounded for the scaled variables $f$ replacing asymptotic expansions are known. The leading form of the sum is then obtained by functions—specifically gamma functions and orthogonal polynomials—for which uniform $N$ with $(\text{recall (4.32) and (4.36))}$. Our strategy to compute the various scaled limits (order sum is recognized as a Riemann sum, allowing the limiting scaled correlations to be

\[
\rho^{\text{soft}}_{(n)}((\tau_1, y_1), \ldots, (\tau_n, y_n)) = \det[K^{\text{soft}}((\tau_j, y_j), (\tau_k, y_k))]_{j, k = 1, \ldots, n} \tag{5.4}
\]

where

\[
K^{\text{soft}}((\tau_x, x), (\tau_y, y)) = \begin{cases} A^{(1)}_{\tau_y-\tau_x}(x, y), & \tau_y \geq \tau_x \\ A^{(2)}_{\tau_y-\tau_x}(x, y), & \tau_y < \tau_x, \end{cases}
\]

\[
A^{(1)}_{x}(x, y) := \int_{-\infty}^{\infty} e^{-\tau u} \text{Ai}(x + u) \text{Ai}(y + u) \, du \tag{5.5}
\]

\[
A^{(2)}_{x}(x, y) := -\int_{-\infty}^{0} e^{-\tau u} \text{Ai}(x + u) \text{Ai}(y + u) \, du.
\]

This is the so-called Airy process $A_2$, which underlies the distribution of the largest eigenvalue in the scaled limit of the Dyson Brownian motion model of the GUE [28, 15].

The significance of these facts with respect to the present study comes about upon recalling the theory revised in the sentence including (2.20). Thus one has that the eigenvalue in the scaled limit of the Dyson Brownian motion model of the GUE [28, 15].

$\rho \mapsto \rho^{\text{soft}}$ (recall (2.19) with $k = 0, n_1 = p = n$, $w(y) = e^{-y}$ are identical in distribution to the stochastic variables $\{l(n, n - k)\}_{k=0,1,\ldots}$. By analogy with the behavior of the stochastic variables $\{l(n+k, n-k)\}_{k=0,1,\ldots}$, one may anticipate that their distribution with $k = O(n^{2/3})$ is controlled by the Airy process $A_2$. We will find that, with the differences between the ranks (species) scaled to be of order $n^{2/3}$, then this is indeed the case, and that the same effect holds for the soft edge in the GUE minor process.

### 5.2. Method

For finite $N$, the correlation kernel $K$ has the functional form

\[
\sum_{k=1}^{N} g(k, N; x, s) \quad \text{or} \quad \sum_{k=-\infty}^{0} g(k, N; x, s) \tag{5.6}
\]

(recall (4.32) and (4.36)). Our strategy to compute the various scaled limits ($x, s$ scaled with $N$ and $N \to \infty$) is to take advantage of the summands in (5.6) being classical functions—specifically gamma functions and orthogonal polynomials—for which uniform asymptotic expansions are known. The leading form of the sum is then obtained by replacing $f$ by the leading term in its uniform asymptotic expansion, with an error term which is uniformly bounded for the scaled variables $x$ and $s$ within compact sets, or in the case of the soft edge the semi-infinite interval $x \in (x_0, \infty)$. Moreover, this leading order sum is recognized as a Riemann sum, allowing the limiting scaled correlations to be written as an integral.

Let the scaling of coordinates be given by $x_j \mapsto \alpha_N + a_N X_j =: A_N(x_j), s_j \mapsto b_N \tau_j$ and append as a superscript to the finite system correlations the symbol $(\cdot)$ so they are $\rho^{(N)}_{(k)}$. The uniform convergence with respect to the scaled variables $x$ within an interval $I$ allows one to conclude that

\[
\lim_{N \to \infty} a_N^k \int_I dx_1 \cdots \int_I dx_k \rho^{(N)}_{(k)}((b_1 \tau_1, A_N(X_1)), \ldots, (b_N \tau_k, A_N(X_k)))
\]

\[
= \int_I dx_1 \cdots \int_I dx_k \rho_{(k)}((\tau_1, X_1), \ldots, (\tau_k, X_k)).
\]
This implies convergence of the corresponding spacing distributions (see [34], [6, proposition 2]).

5.3. Fixed differences between species

In keeping with section 5.1, the two regimes of interest in this setting are the bulk and edge scalings.

**Bulk scaling**

Explicit details will be worked out only in the Gaussian case, as this case is typical. In particular the same scaled correlations result from the Laguerre and Jacobi cases, as is typical in random matrix theory (a form of universality).

The bulk scaling is obtained by the change of variables

\[ y_i = \frac{\pi Y_i}{\sqrt{2N}}, \tag{5.7} \]

which makes the mean particle density in the neighborhood of the origin unity, and choosing the species to differ from \( N \) by a constant:

\[ s_i = N - c_i. \tag{5.8} \]

**Proposition 3.** For the bulk scaling specified by (5.7) and (5.8)

\[ y_i = \frac{\pi Y_i}{\sqrt{2N}}, \tag{5.9} \]

and with \( s_i \) specified in terms of \( c_i \) by (5.8):

\[
\lim_{N \to \infty} \frac{b_N(c_i)}{b_N(c_j)} \frac{\pi}{\sqrt{2N}} K(s_j, y_j; s_l, y_l) = B((c_j, Y_j), (c_l, Y_l)),
\tag{5.10}
\]

where \( b_N(c) := (2N)^{-c/2} \) and

\[
B((\tau_x, x), (\tau_y, y)) := \begin{cases} 
\int_0^1 s^{\tau_y - \tau_x} \cos(\pi s(x - y) + \pi(\tau_x - \tau_y)/2) \, ds, & \tau_y \geq \tau_x \\
- \int_1^\infty s^{\tau_y - \tau_x} \cos(\pi s(x - y) + \pi(\tau_x - \tau_y)/2) \, ds, & \tau_y < \tau_x
\end{cases}.
\tag{5.11}
\]

Consequently

\[
\lim_{N \to \infty} \left( \frac{\pi}{\sqrt{2N}} \right)^r \rho(r)(\{(s_j, y_j)\}_{j=1,\ldots,r}) = \det[B((c_j, Y_j), (c_l, Y_l))]_{j,l=1,\ldots,r}.
\tag{5.12}
\]

**Proof.** Substituting the Gaussian case of (4.2)–(4.9) and (4.13) in (4.36) shows that for \( s_j \geq s_l \) (\( c_j \leq c_l \))

\[ K(s_j, y_j; s_l, y_l) = e^{-y_j^2} \sum_{k=1}^{s_l} \frac{1}{2^{n-k}(s_l-k)!} H_{s_j-k}(y_j)H_{s_l-k}(y_l). \tag{5.13} \]
In keeping with the discussion in section 5.2 above, to analyze the sum we make use of the known uniform asymptotic expansion [35]:

$$\frac{\Gamma(n/2 + 1)}{\Gamma(n + 1)} e^{-x^2/2} H_n(x) = \cos(\sqrt{2n + 1}x - n\pi/2) + O(n^{-1/2})O(1),$$

and a simple trigonometric identity to deduce that

$$K(s_j, y_j; s_l, y_l) = \frac{1}{2\sqrt{\pi}} \sum_{k=1}^{N-c_l} \frac{1}{2^{N-c_l-k}} \frac{(N - c_j - k)!}{((N - c_j - k)/2)!(N - c_l - k)/2!} \times \cos\left(\pi \sqrt{\frac{N - k}{N}} (Y_j - Y_l) + \frac{\pi}{2}(c_j - c_l)\right) \left(1 + O(N^{-1/2})O(1)\right).$$

Here the main contribution to the sum comes from \((N - k)/N = O(1)\). Expanding the ratio of factorials in this setting using Stirling’s formula gives

$$K(s_j, y_j; s_l, y_l) = \frac{2(c_l-c_j)/2}{\sqrt{2\pi}} N^{c_l-c_j-1/2} \sum_{k=1}^{s_l} \frac{N - k}{N} \frac{(c_l-c_j-1)/2}{\sqrt{2\pi}} \times \cos\left(\pi \sqrt{\frac{N - k}{N}} (Y_j - Y_l) + \frac{\pi}{2}(c_j - c_l)\right) \left(1 + O(N^{-1/2})O(1)\right).$$

(5.14)

Recognizing the sum as the Riemann sum approximation to a definite integral in the variable \((N - k)/N = t\), then changing variables \(t = s^2\) in the definite integral gives (5.10) for \(c_j \leq c_l\).

It remains to study the case \(s_j < s_l\) (\(c_j > c_l\)) as specified by (4.15). As the only difference between (4.40) and (4.36) is in the range of summation, the working leading to (5.14) again applies, so this asymptotic formula remains valid but with \(k\) summed from \(-\infty\) to 0. Crucially, because \(c_j > c_l\) this sum is convergent and is furthermore a Riemann sum approximation to the same definite integral as found for the case \(c_j \leq c_l\), but on \((-\infty, 0]\) instead of \([0, 1]\), hence implying the second formula in (5.10).

\(\square\)

5.4. Hard edge scaling

Hard edge scaling is possible for both the Laguerre and Jacobi cases; here the details will be given in the Laguerre case only, as the limiting correlations are the same in both cases. For the \(N \times N\) LUE the hard edge scaling results from the change of variables

$$x_i = \frac{X_i}{4N},$$

(5.15)

which makes the inter-eigenvalue spacings in the neighborhood of the hard edge \(x = 0\) of order unity. We seek the limiting correlations with the scaling (5.15) and the species specified by (5.8).

\textbf{Proposition 4.} For the hard edge scaling (5.15) and with \(s_i\) specified in terms of \(c_i\) by (5.8)

$$\lim_{N \to \infty} \frac{1}{4N} \frac{h_N(c_l)}{h_N(c_j)} K(s_j, x_j; s_l, x_l) = H((c_j, X_j); (c_l, X_l)).$$

(5.16)

doi:10.1088/1742-5468/2011/08/P08011
where $h_N(c) := (2N)^{-c}$ and

$$
H((\tau_x, x), (\tau_y, y)) := \begin{cases}
\frac{1}{4} \int_{0}^{1} s^{(\tau_x - \tau_y)/2} J_{a+\tau_x}((sx)^{1/2}) J_{a+\tau_y}((sy)^{1/2}) \, ds, & \tau_y \geq \tau_x \\
-\frac{1}{4} \int_{1}^{\infty} s^{(\tau_y - \tau_x)/2} J_{a+\tau_x}((sx)^{1/2}) J_{a+\tau_y}((sy)^{1/2}) \, ds, & \tau_y < \tau_x,
\end{cases} \tag{5.17}
$$

with the convergence uniform for $(c_j, X_j)$, $(c_l, X_l)$ within compact sets. Consequently

$$
\lim_{N \to \infty} \left( \frac{1}{4N} \right)^r \rho(r) \left( \{(s_j, x_j)\}_{j=1, \ldots, r} \right) = \det[H(c_j, x_j; c_l, x_l)]_{j,l=1, \ldots, N}. \tag{5.18}
$$

Proof. Substituting the explicit form of the Laguerre case of the quantities in (4.36) gives

$$
K(N - c_j, x_j; N - c_l, x_l) = x_j^{a+c_j} e^{-x_j} \sum_{k=1}^{N-c_l} \frac{\Gamma(N - c_j - k + 1)}{\Gamma(N - k + a + 1)} L_{N-c_j-k}^{(a+c_j)}(x_j) L_{N-c_l-k}^{(a+c_l)}(x_l), \tag{5.19}
$$

valid for $c_l \geq c_j$. As $x_j, x_l$ are scaled according to (5.15), it is appropriate to make use of the uniform asymptotic expansion [35]

$$
e^{-x^2/2} x^{a/2} L_n^a(x) = n^{a/2} J_a(2nx^{1/2}) + \begin{cases}
x^{5/4} O(n^{a/2-3/4}), & cn^{-1} < x < \omega \\
x^{a/2+2} O(n^a), & 0 < x < cn^{-1}.
\end{cases}
$$

Using this, and expanding the ratio of gamma functions with $(N-k)/N = O(1)$ using Stirling’s formula, we deduce

$$
K(N - c_j, x_j; N - c_l, x_l) = (2N)^{a-c_j} \sum_{k=1}^{N-c_l} \left( \frac{N-k}{N} \right)^{(c_l-c_j)/2} J_{a+c_l} \left( \left( \frac{N-k}{N} \right)^{1/2} X_j \right) J_{a+c_l} \left( \left( \frac{N-k}{N} \right)^{1/2} X_l \right) \left( 1 + O \left( \frac{1}{N} \right) O(1) \right). \tag{5.20}
$$

This is a Riemann sum, and the result (5.16) in the case $c_l \geq c_j$ follows.

The expression (5.19) is also valid for $c_l < c_j$, provided the summation is now made over $k \in \mathbb{Z}_{\leq 0}$. Following the above working through again gives (5.20), but with the summation over $k \in \mathbb{Z}_{\leq 0}$. Because $c_l < c_j$ the sum is convergent, and its leading form given by the definite integral is made explicit in (5.17). \qed

5.5. Soft edge scaling with difference between species $O(N^{2/3})$

As anticipated from the viewpoint of last passage percolation, the soft edge scaling permits well-defined correlations with the species separated by $O(N^{2/3})$. The details can be worked out for both the Gaussian and Laguerre cases, although the limiting correlations correspond to the Airy process $A_2$ and so are independent of the particular case.

**Proposition 5.** In the Gaussian case, scale $s_l$ according to

$$
s_i = N - 2c_i N^{2/3}, \tag{5.21}
$$

$$
doi:10.1088/1742-5468/2011/08/P08011
$$
and scale $y_i$ according to

$$y_i = (2s_i)^{1/2} + \frac{Y_i}{\sqrt{2s_i^{1/6}}}.$$  \tag{5.22}

For large $N$

$$\lim_{N \to \infty} \frac{1}{\sqrt{2N^{1/6}}} \frac{\alpha_N(c_l, Y_i)}{\alpha_N(c_j, Y_j)} K(s_j, y_j; s_i, y_i) = \begin{cases} A^{(1)}_{c_j-c_l}(Y_j, Y_i), & c_j \geq c_l \\ A^{(2)}_{c_j-c_l}(Y_j, Y_i), & c_j < c_l \end{cases}$$  \tag{5.23}

where $A^{(1)}$, $A^{(2)}$ are given by (5.5) and $\alpha_N(c, Y) := e^{-N^{1/3}Y} - e^{-N^{2/3}} e^{N^{1/3}Y} e^{2\sqrt{3}Y}$, with the convergence uniform for $Y_j, Y_k$ in a semi-infinite interval $(x_0, \infty)$. Consequently

$$\lim_{N \to \infty} \left( \frac{1}{\sqrt{2N^{1/6}}} \right)^r \rho(r)\{(s_j, y_j)\}_{j=1,\ldots,r} = \det \left[ K^{\text{soft}} ((c_j, Y_j), (c_k, Y_k)) \right]_{j,k=1,\ldots,r}.$$  \tag{5.24}

In the Laguerre case, scale $s_i$ according to

$$s_i = N - \check{s}_i, \quad \check{s}_i := 2c_3(2N)^{2/3}$$  \tag{5.25}

and scale $y_i^{(s_i)}$ according to

$$y_i = 4s_i + 2(a + N - s_i) + 2(2N)^{1/3}Y_i.$$  \tag{5.26}

For large $N$

$$\lim_{N \to \infty} 2(2N)^{2/3} \frac{\beta_N(c_l, Y_i)}{\beta_N(c_j, Y_j)} K(s_j, y_j; s_i, y_i) = \begin{cases} A^{(1)}_{c_j-c_l}(Y_j, Y_i), & c_l \geq c_j \\ A^{(2)}_{c_j-c_l}(Y_j, Y_i), & c_l < c_j \end{cases}$$  \tag{5.27}

with $\beta_N(c, Y) = e^{-(2N)^{1/3}Y} - e^{-(2N)^{2/3}} e^{2(2N)^{1/3}Y} e^{2\sqrt{3}Y}$, and the convergence is uniform for $Y_j, Y_k$ in a semi-infinite interval $(x_0, \infty)$. Consequently

$$\lim_{N \to \infty} (2N)^{2/3} \rho(r)\{(s_j, y_j)\}_{j=1,\ldots,r} = \det \left[ K^{\text{soft}} ((c_j, Y_j), (c_k, Y_k)) \right]_{j,k=1,\ldots,r}.$$  \tag{5.28}

**Proof.** The derivation is very similar in both cases, so we’ll be content with presenting the details in the Laguerre case only. Reading off from (5.19) we have

$$K(N - \check{s}_j, y_j; N - \check{s}_l, y_l) = y_j^{a_i} e^{-y_j} \sum_{k=1}^{N-k_l} \frac{\Gamma(N - \check{s}_j - k + 1)}{\Gamma(N - k + a + 1)} L^{(a+i_\check{s}_j)}_{N-\check{s}_j-k}(y_j) L^{(a+i_\check{s}_l)}_{N-\check{s}_l-k}(y_l).$$  \tag{5.29}

This formula is valid for $\check{s}_j \leq \check{s}_l$ (for $\check{s}_j > \check{s}_l$ the RHS is to be modified by multiplying by $-1$ and changing the summation terminals to $k \in \mathbb{Z}_{\leq 0}$; this modification does not change the working below in any essential way, and so will not be considered explicitly).

We seek the asymptotic form of (5.19) upon the scalings (5.26) and (5.25). Our chief tool is the uniform asymptotic expansion [26]:

$$x^{a/2} e^{-x/2} L_n^a(x) = (-1)^n (2n)^{-1/3} \sqrt{(n+a)!/n!} \times \begin{cases} O(e^{-X}), & X \geq 0 \\ O(1), & X < 0 \end{cases}$$  \tag{5.30}

doi:10.1088/1742-5468/2011/08/P08011 23
where

\[ x = 4n + 2a + 2(2n)^{1/3}X \]  

(5.31)

(this form allows for \( a = O(n) \); the classical Plancherel–Rotach-type formula given in, for example, [35] requires \( a \) to be fixed and correspondingly has \( \sqrt{(n+a)!/n!} \) replaced by \( n^{a/2} \)). Use of this formula, rewritten to be

\[ x^{n/2}e^{-x/2}L_{n-k}^a(x) = (-1)^{n-k}(2n)^{-1/3}(n-k+a)!/(n-k)! \times \left( \text{Ai}(X + \frac{2k}{(2n)^{1/3}}) + O(n^{-2/3}) \begin{cases} O(e^{-X-k/n^{1/3}}), & k \geq 0 \\ O(1), & k < 0 \end{cases} \right) \]

(5.32)

with \( n = N - \tilde{s}_j \) shows that for large \( N \)

\[ K(N - \tilde{s}_j, y_j; N - \tilde{s}_l, y_l) = e^{-(2N)^{1/3}(Y_j - Y_l)(2N)^{-2/3} \sum_{k=1}^{N-\tilde{s}_j} \left( \frac{(N - \tilde{s}_j - k)!}{(N - \tilde{s}_l - k)!} \right)^{1/2} \times \left( \text{Ai}(Y_j + 2k/(2N)^{1/3})\text{Ai}(Y_l + 2k/(2N)^{1/3}) + O(N^{-2/3})O(e^{-Y_j - Y_l}) \right) \}. \]

The leading order contribution to the summation comes from \( k \) of order \( N^{1/3} \). Using this fact, noting from Stirling’s formula that for large \( s \)

\[ \left( \frac{(N - k - \tilde{s}_j)!}{(N - k - \tilde{s}_l)!} \right)^{1/2} = s^{(k_j - k_l)/2} \exp \left( \frac{(k_j^2 - k_l^2)}{4N} + \frac{(k_j^3 - k_l^3)}{12N^2} + O(N^{-1}) \right) \],

(5.33)

where \( k_i = k + \tilde{s}_i \) \((i = j, l)\) and the remainder term is uniform in \( k/N^{1/3} \), the sum is recognized as the Riemann sum approximation to \( A^{(1)} \) as defined in (5.5), implying the result (5.27) in the case \( c_i \geq c_j \).

\[ \square \]

6. Discussion

As pointed out to us by Borodin, replacing (5.11) by

\[
\begin{cases}
\frac{1}{2} \int_{-1}^{1} (is)^{\tau_y - \tau_x} e^{is\pi(x-y)} \, ds, \quad \tau_y > \tau_x \\
-\frac{1}{2} \int_{\mathbb{R} \setminus [-1,1]} (is)^{\tau_y - \tau_x} e^{is\pi(x-y)} \, ds, \quad \tau_y < \tau_x
\end{cases}
\]

(6.1)

leaves the determinant (5.12) unchanged. Further changing scale \( x \mapsto x/\pi, y \mapsto y/\pi \) removes the \( \pi \)'s from the exponents, and multiplies each integrand by a factor of \( 1/\pi \) to account for the corresponding scaling of the correlation function. The significance of this form is that it is identical to the \( \gamma = 0 \) case of the correlation kernel for the so-called bead model [8, theorem 2]. In fact, the bead model was already known to be closely related to the GUE minor process [8, section 4.1]. The form (6.1) can also be obtained as a limit of the incomplete beta kernel of Okounkov and Reshetikhin [31, section 3.1.7] (write the parameter \( z \) as \( z = 1 + ia \), change variables \( w = 1 + ias \), rescale the space variable \( l \) by \( a^{-1} \) and take \( a \to 0 \)). The recent work [19] obtains the incomplete beta kernel in the context of a study of random lozenge tilings. Further the kernels of [3, theorem 4.4] permit degeneracies to (6.1).

doi:10.1088/1742-5468/2011/08/P08011
Another discussion point is in relation to consistency between the present results and results from [16]. In [16] the correlations for the PDF

\[
\frac{1}{C} \prod_{j=1}^{n} e^{-(x_j+y_j)/2}e^{A(x_y-y_j)/2} \prod_{1 \leq j < k \leq n} (x_j - x_k)(y_j - y_k)\chi(x_1 > y_1 > \cdots > x_n > y_n) \tag{6.2}
\]

were computed, along with the scaled limits at the soft and hard edges, and in the bulk. The PDF \((6.2)\) with \(A = -1\) is identical to the PDF \((2.19)\) with \(w(y) = e^{-y}, n_2 = n, p = 1\) and \(y_{n+1}^{(1)} = 0\). Setting \(y_{n+1}^{(1)} = 0\) would not be expected to alter the soft edge and bulk scaling limits, so it should be that the scaled correlations in [16] contain as special cases the results \((5.2)\) and \((5.12)\) for \(|c_j - c_l| = 0, 1\).

To see that this is indeed the case, we recall from [16] that with

\[
A \mapsto \left\{ \begin{array}{ll}
\sqrt{n} \alpha/\pi, & \text{bulk} \\
\alpha/(2n)^{1/3}, & \text{soft edge} \\
\end{array} \right. \tag{6.3}
\]

the scaled correlation for \(k_1\) variables of species type \(x\), and \(k_2\) variables of species type \(y\) was calculated to equal

\[
\rho_{(k_1,k_2)}(X_1, \ldots, X_{k_1}, Y_1, \ldots, Y_{k_2}) = \det \left[ \begin{array}{cc}
[K_{\text{scaled}}(X_j, X_l)]_{j=1, \ldots, k_1, l=1, \ldots, k_2} & [K_{\text{scaled}}(X_j, Y_l)]_{j=1, \ldots, k_1, l=1, \ldots, k_2} \\
[K_{\text{scaled}}(Y_j, X_l)]_{j=1, \ldots, k_1, l=1, \ldots, k_2} & [K_{\text{scaled}}(Y_j, Y_l)]_{j=1, \ldots, k_1, l=1, \ldots, k_2} \\
\end{array} \right] \tag{6.4}
\]

where

\[
K_{\text{scaled}}(Y, Y') = K^{\text{soft}}(Y, Y')
\]

\[
K_{\text{scaled}}(Y, X) = -e^{\alpha(X-Y)}\chi(X > Y) + e^{\alpha X/2} \int_{-\infty}^{X} e^{-\alpha v/2}K^{\text{scaled}}(v, Y) \, dv
\]

\[
K_{\text{scaled}}(X, Y) = -e^{-\alpha X/2} \frac{\partial}{\partial X} \left( e^{\alpha X/2}K^{\text{scaled}}(X, Y) \right)
\]

\[
K_{\text{scaled}}(X, X') = -e^{\alpha(X-X')/2} \frac{\partial}{\partial X} \left( e^{\alpha X/2} \int_{-\infty}^{X'} e^{-\alpha v/2}K^{\text{scaled}}(X, v) \, dv \right). \tag{6.5}
\]

In the soft edge case \(K^{\text{scaled}} = K^{\text{soft}}\) as specified by \((5.3)\), while in the bulk \(K^{\text{scaled}} = K^{\text{bulk}}\), where \(K^{\text{bulk}}(X, Y) = \frac{\sin \pi(X - Y)}{\pi(X - Y)} = \int_{0}^{1} \cos \pi(X - Y)t \, dt\). \tag{6.6}

We see from \((6.3)\) that \(A = -1\) corresponds to \(\alpha = 0\) in the bulk and \(\alpha \to -\infty\) at the soft edge. We see from \((6.5)\) that for \(\alpha \to -\infty\)

\[
K_{\text{soft}}^{\text{oo}}(Y, X) \sim -\frac{2}{\alpha}K^{\text{soft}}(Y, X), \quad K_{\text{oo}}^{\text{soft}}(X, Y) \sim -\frac{\alpha}{2}K^{\text{soft}}(X, Y),
\]

\[
K_{\text{oo}}^{\text{soft}}(X, X') \sim K^{\text{soft}}(X, X').
\]
When substituted in (6.4) the factors $-2/\alpha$, $-\alpha/2$ cancel, and so agreement with (5.2) is found. Further, setting $\alpha = 0$ in (6.5) and recalling (6.6) gives

$$K_{\text{eo}}^\text{bulk}(Y, X)|_{\alpha = 0} = -\chi(X > Y) + \int_{-\infty}^{X} \frac{\sin \pi(v - Y)}{\pi(v - Y)} \, dv,$$

$$K_{\text{eo}}^\text{bulk}(X, Y)|_{\alpha = 0} = \pi \int_{0}^{1} t \sin \pi(X - Y) t \, dt, \quad K_{\text{eo}}^\text{scaled}(X, X')|_{\alpha = 0} = \frac{\sin \pi(X - X')}{\pi(X - X')}.$$

A simple calculation shows that the first of these can be rewritten as

$$K_{\text{eo}}^\text{bulk}(Y, X)|_{\alpha = 0} = -\int_{1}^{\infty} \frac{\sin \pi v(X - Y)}{\pi v} \, dv.$$

With this we obtain agreement with (5.12) in the case $|c_j - c_l| = 0, 1$, as expected.

**Acknowledgments**

We thank Eric Nordenstam for some stimulating discussions and Alexei Borodin for the content of the first paragraph in section 6. The work of PJF is supported by the Australian Research Council.

**Appendix**

Since the completion of this work, Borodin and Péché have posted a work [7] on the arXiv, subsequently published in [7], which amongst other results reclaims our proposition 5 in the Laguerre case. The strategy used is, at a technical level, quite different to that adopted here.

In this appendix we concern ourselves with another aspect of the work [7], relating to a generalization of our (3.3):

$$A_{(n+1)} = A_{(n)} + \vec{x}_{(n)} \vec{x}_{(n)}^\dagger, \quad A_{(0)} = [0]_{p \times p} \quad (A.1)$$

where $\vec{x}_{(n)}$ is a $p \times 1$ column vector of complex Gaussians with entries such that the modulus of the $ith$ component has distribution $\Gamma[1, 1/(\pi_i + \pi_n)]$. As in section 3.2, the point of interest is in the joint eigenvalue PDF for $\{A_{1}, \ldots, A_{p}\}$. This is not computed directly, but as in the discussion around (2.20), it is noted that the directed percolation in the $p \times p$ square which each lattice site $(i, j)$ containing an exponential random variable of density $(\pi_i + \pi_j)e^{-\pi_i + \pi_j}x$ has the distribution of the stochastic variable $l(p, p)$ equal to that of the distribution of the largest eigenvalue of $A_{(p)}$. For the percolation problem, the joint distribution of $\{l(j, p)\}_{j=1, \ldots, p}$ can be calculated, leading to a joint PDF for the $p$ species of variables $\{x_j^{(s)}\}, (s = 1, \ldots, p)$ with $j = 1, \ldots, s$ proportional to

$$\det[e^{-\pi x_j^{(p)}}]_{i,j=1, \ldots, p} \prod_{k=1}^{p-1} \det[e^{-\pi_{k+1}(x_j^{(k+1)} - x_i^{(k)})}\chi(x_j^{(k+1)} > x_i^{(k)})]_{i,j=1, \ldots, k+1} e^{-\pi_1 x_1^{(1)}}. \quad (A.2)$$

The question is asked as to whether this joint PDF is, in fact, the joint PDF for the eigenvalues of the matrices $A_{(s)}, s = 1, \ldots, p$.

doi:10.1088/1742-5468/2011/08/P08011
In fact, the working from [17, section 5] allows this question to be answered in the affirmative in the limit \( \pi_i \to c, (i = 1, \ldots, p) \). Thus it follows from [17, Corollary 3] that the condition PDF for the eigenvalues \( \{ a_j \}_{j=1, \ldots, n} \) of \( A(s) \) is proportional to

\[
\prod_{i=1}^{n+1} \lambda_i^{p-(n+1)} \prod_{j=1}^{n} \frac{1}{\alpha_j^{p-n}} e^{-(c+\tilde{s})n} \prod_{i<j} \lambda_i - \lambda_j \prod_{p<n} a_j - a_i \chi(\lambda < a). \tag{A.3}
\]

Let us now write \( \lambda_j \mapsto \lambda_j^{(n+1)}, a_j \mapsto \lambda^{(n)}_j \). The sort joint PDF is the product from \( n = 1, \ldots, p-1 \) of the conditional PDFs (A.3), multiplied by the PDF in the case \( n = 1 \), which is proportional to \((x_1^{(1)})^{p-1} e^{-(c+\tilde{s})x_1^{(1)}}\). This gives (A.2) with the first determinant therein replaced by \( \prod_{i=1}^{p} e^{-cx_i^{(p)}} \prod_{i<j} (x_j^{(p)} - x_i^{(p)}) \), thus verifying (A.2) in the case that \( \pi_i \to c, (i = 1, \ldots, p) \).

Using different strategies, Dieker and Warren [9] have now proved that the joint PDF (A.2) is the joint PDF for the eigenvalues of the matrices \( A(s), s = 1, \ldots, p \) in the general case.

References

[1] Baryshnikov Y, GUEs and queues, 2001 Probab. Theory Relat. Fields 119 256
[2] Baik J and Rains E M, Algebraic aspects of increasing subsequences, 2001 Duke Math. J. 109 1
[3] Borodin A, Periodic Schur process and cylindric partitions, 2007 Duke Math. J. 140 391
[4] Borodin A, Ferrari P L and Sasamoto T, Fluctuation properties of the TASEP with periodic initial configuration, 2007 J. Stat. Phys. 129 1055
[5] Borodin A, Ferrari P L and Sasamoto T, Transition between Airy_1 and Airy_2 processes and the TASEP fluctuations, 2008 Commun. Pure Appl. Math. 61 1603
[6] Borodin A and Forrester P J, Increasing subsequences and the hard-to-soft transition in matrix ensembles, 2003 J. Phys. A: Math. Gen. 36 2963
[7] Borodin A and Peché S, Airy kernel with two sets of parameters in directed percolation and random matrix theory, 2008 J. Stat. Phys. 132 275
[8] Boutillier C, The bead model and limit behaviours of dimer models, 2009 Ann. Probab. 37 107
[9] Dieker A B and Warren J, On the largest-eigenvalue process for generalized Wishart matrices, 2009 Alea 6 369
[10] Dyson F J, Correlations between eigenvalues of a random matrix, 1970 Commun. Math. Phys. 29 235
[11] Forrester P J, 2010 Log-gases and Random Matrices (Princeton, NJ: Princeton University Press)
[12] Forrester P J, The spectrum edge of random matrices, 1993 Nucl. Phys. B 403 709
[13] Forrester P J, Growth models, random matrices and Painlevé transcendent, 2003 Nonlinearity 16 R27
[14] Forrester P J, Beta random matrix ensembles, 2009 Random Matrix Theory and its Applications (Lecture Notes Series, IMS, NUS) vol 18, ed Z Bai, Y Chen and Y-C Liang (Singapore: World Scientific) p 27
[15] Forrester P J, Nagao T and Honner G, Correlations for the orthogonal-unitary and symplectic-unitary transitions at the hard and soft edges, 1999 Nucl. Phys. B 553 601
[16] Forrester P J and Rains E M, Correlations for superpositions and decimations of Laguerre and Jacobi orthogonal matrix ensembles with a parameter, 2004 Probab. Theory Relat. Fields 130 518
[17] Forrester P J and Rains E M, Interpretations of some parameter dependent generalizations of classical matrix ensembles, 2005 Probab. Theory Relat. Fields 131 1
[18] Glynn P and Whitt W, Departures from many queues in series, 1991 Ann. Appl. Probab. 1 546
[19] Gorin V, Non-intersecting paths and Hahn polynomial ensembles, 2008 Funkts. Anal. Prilozh. 42 23
[20]Greene C, An extension of Schensted’s theorem, 1974 Adv. Math. 14 254
[21] Johansson K, Shape fluctuations and random matrices, 2000 Commun. Math. Phys. 209 437
[22] Johansson K, Discrete polynuclear growth and determinantal processes, 2003 Commun. Math. Phys. 242 277
[23] Johansson K, The arctic circle boundary and the Airy process, 2005 Ann. Probab. 33 1
[24] Johansson K, Non-intersecting, simple, symmetric random walks and the extended Hahn kernel, 2005 Ann. Inst. Fourier 55 2129
[25] Johansson K and Nordenstam E, Eigenvalues of GUE minors, 2006 Elect. J. Probab. 11 1342

doi:10.1088/1742-5468/2011/08/P08011
Determinantal correlations for classical projection processes

[26] Johnstone I M, On the distribution of the largest principal component, 2001 Ann. Stat. 29 295
[27] Knuth D E, Permutations, matrices and generalized Young tableaux, 1970 Pac. J. Math. 34 709
[28] Macêdo A M S, Universal parametric correlations at the soft edge of the spectrum edge of random-matrix ensembles, 1994 Europhys. Lett. 26 641
[29] Nagaï T and Forrester P J, Multilevel dynamical correlation functions for Dyson’s Brownian motion model of random matrices, 1998 Phys. Lett. A 247 42
[30] Nordenstam E J G, Erratum to Eigenvalues of GUE minors, 2007 Elect. J. Probab. 12 1048
[31] Okounkov A and Reshetikhin N, Correlation function of Schur process with applications to local geometry of a random 3-dimensional Young diagram, 2003 J. Am. Math. Soc. 16 581
[32] Okounkov A and Reshetikhin N, The birth of a random matrix, 2006 Moscow Math J. 6 553
[33] Porter C E, 1965 Statistical Theories of Spectra: Fluctuations (New York: Academic)
[34] Sohnikov A, A note on the universality of the distribution of the largest eigenvalues in certain sample covariance matrices, 2002 J. Stat. Phys. 108 1033
[35] Szegö G, 1975 Orthogonal Polynomials 4th edn (Providence, RI: American Mathematical Society)