Convergence of the eigenvalue density for β-Laguerre ensembles on short scales

Philippe Sosoe[*]   Percy Wong[†]

Abstract

In this note, we prove that the normalized trace of the resolvent of the β-Laguerre ensemble eigenvalues is close to the Stieltjes transform of the Marchenko-Pastur (MP) distribution with very high probability, for values of the imaginary part greater than $m^{-1+\epsilon}$. As an immediate corollary, we obtain convergence of the one-point density to the MP law on short scales. The proof serves to illustrate some simplifications of the method introduced in our previous work to prove a local semi-circle law for Gaussian β-ensembles.

1 Introduction

Consider an $m \times n$ matrix X, whose entries are i.i.d. complex Gaussian random variables with mean 0 and variance $\mathbb{E}|X_{ij}|^2 = 1$. The $m \times m$ matrix $H = XX^*$, the star $*$ denoting the conjugate transpose, is a *Wishart matrix* [23]. It is a classical result of random matrix theory that the distribution of the eigenvalues $\lambda_1, \dots, \lambda_m$ of H is given by the density $f_2(\lambda)$ on $\mathbb{R}^m$:

$$f_2(\lambda) = Z_{2,m}^{-1} \prod_{i<j} |\lambda_i - \lambda_j|^2 \cdot \prod_{i=1}^{m} \lambda_i^{n-m} e^{-\beta \sum_{i=1}^{m} \lambda_i / 2}.$$

Here $Z_{2,m}$ is a normalization factor. Suppose the limit $m/n \to d$ exists for $0 < d \le 1$. Then the empirical distribution

$$F_{2,m}(x) = \frac{1}{m} \cdot \sharp\{1 \le i \le m : \lambda_i \le x\}$$

of the eigenvalues of the rescaled matrix H/m converges to the Marchenko-Pastur distribution, with density

$$\rho_{MP,d}(x) = \frac{1}{2d\pi x} \sqrt{(\lambda_+ - x)(x - \lambda_-)} \mathbf{1}_{[\lambda_-, \lambda_+]}(x),$$

[*]Princeton University, USA. E-mail: psosoe@math.princeton.edu
[†]D.E. Shaw & Co. E-mail: percyphwong@gmail.com

where
$$\lambda_{\pm} = (1 \pm \sqrt{d})^2.$$

In fact, V.A. Marchenko and L.A. Pastur [18] showed that weak convergence of the eigenvalue distribution also holds when the entries of X are i.i.d. but not necessarily Gaussian.

A. Edelman and I. Dumitriu [7] have introduced an infinite family of tridiagonal random matrix models, termed β-Laguerre matrices, which generalize the Wishart model, and have the explicit eigenvalue density

$$f_{\beta,a}(\lambda) = Z_{\beta,m}^{-1} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^m \lambda_i^{a - \frac{\beta}{2}(m-1)-1} e^{-\beta \sum_{i=1}^m \lambda_i/2}, \qquad (1.1)$$

for $0 < \beta < \infty$. Here $a > \beta(m-1)/2$ is a real parameter. $Z_{\beta,m}$ is another normalization factor. The (Gaussian) Wishart eigenvalue distribution corresponds to the Laguerre ensemble with $\beta = 2$ and $a = \beta n/2$. Almost-sure convergence of the eigenvalue density to the Marchenko-Pastur distribution with parameter

$$d = \lim \frac{m\beta}{2a} < 1$$

was established for the eigenvalue distributions of the β-Laguerre matrix models by the moment method in [6]; that is, for any $a < b$:

$$F_{\beta,n}(b) - F_{\beta,n}(a) \to \int_a^b \rho_{MP,d}(s)\, \mathrm{d}s,$$

almost surely as $n \to \infty$. $F_{\beta,n}(x)$ is the eigenvalue distribution function, defined as in the case $\beta = 2$ above.

The object of the present note is to extend these results on convergence of the eigenvalue distribution of general β-Laguerre ensembles to *short* intervals $[a_m, b_m]$ such that $b_m - a_m = O(m^{-1+\epsilon})$, for $\epsilon > 0$ arbitrary.

The main result is the following:

Theorem 1.1. *Let $\beta > 0$, and $0 < d \le 1$. Set $a = m\beta/2d$. Let $\delta, \kappa, \epsilon > 0$ be positive parameters, and $E \in (\lambda_- + \kappa, \lambda_+ - \kappa)$. Denote by N_I the number of eigenvalues of the β-Laguerre ensemble of size m with parameter a in the interval I. For any $k > 0$, there exists a constant $C_{\delta,\kappa,\epsilon,k}$ such that:*

$$\mathbb{P}\left(\left| \frac{1}{m} N_{[E-m^{-1+\epsilon}, E+m^{-1+\epsilon}]} - \int_{E-m^{-1+\epsilon}}^{E+m^{-1+\epsilon}} \rho_{MP,d}(x)\, \mathrm{d}x \right| \ge \delta m^{-1+\epsilon} \right) \le C_{\delta,\kappa,\epsilon,k} m^{-k}.$$

The proof of Theorem 1.1 is obtained by combining Theorem 4.1 with Corollary 4.3 in Section 4. Following recent work of L. Erdős, B. Schlein, H.T. Yau and collaborators in the case of Wigner matrices (see for example [9], [10], [11], [12], [13], [14]), our method is based on the study of the resolvent matrix of the Edelman-Dumitriu tridiagonal models.

Let $M_{\beta,d}$ denote a normalized tridiagonal β-Laguerre matrix (see Section 2 for details on notation). The imaginary part of the trace of the resolvent

$$s_{\beta,d}(z) = \frac{1}{n} \operatorname{tr}(M_{\beta,d} - z)^{-1}.$$

provides an approximation for the eigenvalue distribution on scales comparable to the distance between z and the spectrum of $M_{\beta,d}$. In [20], we showed how a resolvent

expansion, together with an iterative argument based on the Schur complement identity could be used to derive a local version of the semi-circle law for the Gaussian β-ensembles. Tridiagonal models for these eigenvalue distributions appeared in [7]. In Proposition 2.1, we use a resolvent expansion to show that $s_{\beta,d}(z)$ is close to the Stieltjes transform of $\rho_{MP,d}$ for values of z away from the spectral edges and with $\Im z > m^{-1/4+\epsilon}$. This implies that Theorem 1.1 holds for intervals I of size $|I| \geq m^{-1/4+\epsilon}$. The argument in the present work is substantially simpler than the corresponding one in [20], where we proved a local convergence result on the scale $m^{-1/2+\epsilon}$ using a resolvent expansion and asymptotics for Hermite polynomials derived by the Riemann-Hilbert method. Instead of attempting to exploit the cancellation due to oscillation of the normalized Laguerre polynomials, we use a general off-diagonal resolvent estimate (see Lemma 2.2). The computation of the limit of the normalized resolvent trace for the deterministic matrix corresponding to $\beta = \infty$ using Riemann-Hilbert asymptotics in [20] has been replaced by a less involved derivation, see Section 3.

The iterative argument leading to Theorem 4.1 in Section 4 is similar to the one in [20]. Note that Theorem 4.1 is deduced from Propositions 4.5 and 4.6, without reference to Proposition 2.1. In contrast, the local result on the intermediate scale $m^{-1/2+\epsilon}$ in [20] was used as an input for an inductive argument to reach the scale $m^{-1+\epsilon}$ [1]. Although it can be entirely replaced by the iteration in Section 4, we have chosen to present the argument for Proposition 2.1 because it provides an elementary alternative to the Schur complement approach to proving the Marchenko-Pastur law. The proof of Proposition 2.1 does not depend on the specific properties of the β-Laguerre ensembles other than concentration of the entries around their mean. It can be applied to fairly general tridiagonal models with independent entries to prove convergence of the eigenvalue distribution, and convergence up to some intermediate scale depending on the magnitude of the entries. We give examples of such extensions in Section 5.

We end this section with some references to previous literature. Local versions of the Marchenko-Pastur law for the eigenvalue distribution of XX^* when the entries of X are independent but not necessarily Gaussian have appeared in [4], [12], [21]. The first paper deals with the hard edge of the spectrum in case $d = 1$. The use of a perturbative expansion around deterministic matrices associated to Laguerre polynomials already appears in [8], where the authors study fluctuations of the spectrum in the large β limit. In [19], I. Popescu proves convergence and Gaussian fluctuations for the moments of tridiagonal matrices under a scaling assumption for the moments of the entries. In contrast to the result in [20] at the time of publication, Theorem 1.1 appears to be new for general β.

2 Tridiagonal models and resolvent expansion

We recall the main result from the work of A. Edelman and I. Dumitriu [7], which will be our starting point. Let $m \in \mathbb{N}$, $\beta > 0$ and choose $0 < d \leq 1$. Let $a = m\beta/(2d)$. Consider the bi-diagonal matrix

$$B_\beta = \begin{pmatrix} \chi_{2a} & & & \\ \chi_{\beta(m-1)} & \chi_{2a-\beta} & & \\ & \ddots & \ddots & \\ & & \chi_\beta & \chi_{2a-\beta(m-1)} \end{pmatrix}.$$

---

[1] Modulo minor changes, the inductive argument in [20] can also be used to obtain the semi-circle law down to scale $n^{-1+\epsilon}$.

In the equation above, the symbol χ_r, $r > 0$ represents a random variable with *chi* distribution with r degrees of freedom, defined by the probability density function

$$\frac{2^{1-\frac{r}{2}}}{\Gamma(r/2)} x^{r-1} e^{-\frac{x^2}{2}}.$$

Γ denotes the Gamma function, defined for $x > 0$ by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, \mathrm{d}t.$$

The random variables appearing in the matrix B_β are all independent. By [7], the eigenvalue distribution of the real tridiagonal matrix $\tilde{M}_{\beta,d} = B_\beta B_\beta^t$ is given by the density function $f_{\beta,a}$ in (1.1). We will be concerned with the scaled β-Laguerre matrix model, defined by

$$M_{\beta,d} = \frac{d}{\beta m}\tilde{M}_{\beta,d} = \frac{d}{\beta m}B_\beta B_\beta^t.$$

We will determine the behavior of the normalized trace of the resolvent, $s_{\beta,d}(z)$ in a neighborhood of the real axis in the upper half-plane $\mathbb{C}^+ = \{\Im z > 0\}$:

$$s_{\beta,d}(z) = \frac{1}{m}\mathrm{tr}\,(M_{\beta,d} - z)^{-1}.$$

The imaginary part $\Im s_{\beta,d}(z)$ is the Poisson integral of the eigenvalue distribution. As such, it is an approximation to the empirical eigenvalue distribution of $M_{\beta,d}$ at scale $\Im z$. To prove Theorem 1.1, it will suffice to show that, with probability no less than $1 - C_{\delta,\kappa,\epsilon,k}m^{-k}$, we have

$$|s_{\beta,d}(z) - s_{MP,d}(z)| < \delta,$$

where

$$s_{MP,d}(z) = \int \frac{1}{x-z}\,\rho_{MP,d}(x)\mathrm{d}x = \frac{z+1-d}{2dz} + \frac{\sqrt{(z-\lambda_-)(z-\lambda_+)}}{2dz}$$

is the Stieltjes transform of the Marchenko-Pastur distribution with parameter d, (see [13], Lemma B.1 or [14] 7.1, as well as Corollary 4.3 below).

We begin by writing

$$M_{\beta,d} = B_\beta B_\beta^t = M_{\infty,d} + \Delta, \tag{2.1}$$

where $M_{\infty,d}$ is the matrix

$$\frac{d}{m}\begin{pmatrix} m/d & \sqrt{m/d}\sqrt{m-1} & & & \\ \sqrt{m/d}\sqrt{m-1} & m/d+m-2 & \sqrt{m/d-1}\sqrt{m-2} & & \\ & \ddots & & & \\ & \sqrt{m/d-m+2}\sqrt{2} & m/d-m+4 & \sqrt{m/d-m+1}\sqrt{1} & \\ & & \sqrt{m/d-m+1}\sqrt{1} & m/d-m+2 & \end{pmatrix}.$$

The entries of the tridiagonal matrix Δ are

$$\Delta_{jj} = \frac{d\chi^2_{2a-(j-1)\beta} - \beta m + d(j-1)\beta}{\beta m} + \frac{d\chi^2_{\beta(m-j+1)} - d\beta(m-j+1)}{\beta m},$$

$$\Delta_{jj-1} = d\frac{\chi_{2a-(j-1)}\chi_{\beta(m-j)} - \sqrt{m/d-j-1}\sqrt{m-j}}{m}.$$

For large m, all the entries of Δ are simultaneously small in magnitude, with overwhelming probability:

$$|\Delta_{jk}| \le m^{-1/2+c}, \quad 1 \le j, k \le m. \tag{2.2}$$

for any $c > 0$. Here and below, we will say an event $E = E(m)$ holds *with overwhelming probability* if, for each k, there is a constant C_k such that

$$\mathbb{P}(E^c) \leq C_k m^{-k}.$$

This follows readily from the definitions, and properties of the χ_r and χ_r^2 distributions, which are concentrated around their mean

$$\mathbb{E}\chi_r = \sqrt{2} \cdot \frac{\Gamma((r+1)/2)}{\Gamma(r/2)} = \sqrt{r} \cdot (1 + O(1/r)),$$
$$\mathbb{E}\chi_r^2 = r,$$

with exponential tails.

2.1 The matrix $M_{\infty,d}$ and generalized Laguerre polynomials

The spectral theory of the symmetric matrix $M_{\infty,d}$ can be described explicitly in terms of Laguerre polynomials. For $\alpha > -1$, the *generalized Laguerre polynomials* L_k^α, $k = 0, 1, \ldots$ are orthogonal polynomials with respect to the measure

$$w_\alpha(x)\, \mathrm{d}x = x^\alpha e^{-x}\, \mathrm{d}x$$

on the positive real axis $\mathbb{R}^+ = [0, \infty)$. See [3, Section 4.5], [2, Section 22.7] L_k^α is normalized such as to have $L^2(\mathbb{R}^+, w_\alpha \mathrm{d}x)$-norm 1:

$$\int L_k^\alpha(x) L_j^\alpha(x) w_\alpha(x)\, \mathrm{d}x = \delta_{j,k}.$$

The eigenvalues of $M_{\infty,d}$ are the normalized zeros

$$\lambda_i = d \cdot l_i / m \tag{2.3}$$

of the mth generalized Laguerre polynomial with parameter

$$\alpha = m \cdot \left(\frac{1}{d} - 1 \right).$$

There is a complete set of corresponding eigenvectors v_i of the form

$$v_i = \begin{pmatrix} L_{m-1}^{m(1/d-1)+1}(l_i) \\ L_{m-2}^{m(1/d-1)+1}(l_i) \\ \vdots \\ L_0^{m(1/d-1)+1}(l_i) \end{pmatrix}, \tag{2.4}$$

for $0 \leq i \leq m - 1$. We denote by u_i the normalized eigenvectors

$$u_i = \frac{v_i}{\|v_i\|}.$$

To derive the above facts regarding the eigenvalues and eigenvectors of $M_{\infty,d}$, we start from the three term recurrence relation for the Laguerre polynomials $\tilde{L}_k^\alpha$, normalized to have leading coefficient

$$(-1)^k \frac{1}{k!}.$$

The three term recurrence relation reads (see [3, Section 4.11, p. 143])

$$x\tilde{L}_k^\alpha(x) = -(k+1)\tilde{L}_{k+1}^\alpha(x) + (2k+\alpha+1)\tilde{L}_k^\alpha(x) - (k+\alpha)\tilde{L}_{k-1}^\alpha(x). \tag{2.5}$$

Letting $k = m - j$ for $j = 1, \ldots m - 1$ and replacing α by $\alpha + 1 = m(1/d - 1) + 1$ in (2.5), we find:

$$
\begin{aligned}
x\tilde{L}_{m-j}^{\alpha+1}(x) = &-(m - j + 1)\tilde{L}_{m-j+1}^{\alpha+1}(x) + (m/d + m - 2j + 2)\tilde{L}_{m-j}^{\alpha+1}(x) \\
&- (m/d - j + 1)\tilde{L}_{m-j-1}^{\alpha+1}(x).
\end{aligned}
\tag{2.6}
$$

The polynomials $\tilde{L}_k^\alpha$ are related to the orthonormal polynomials L_k^α by

$$
L_k^\alpha(x) = (-1)^k \frac{\sqrt{k!}}{\sqrt{\Gamma(\alpha + k + 1)}} \tilde{L}_k^\alpha(x),
\tag{2.7}
$$

see [3, Section 4.5]. Multipliying (2.6) by

$$
\frac{\sqrt{(m - j)!}}{\sqrt{\Gamma(\alpha + m - j + 2)}}.
$$

and using (2.7) and the relation $\Gamma(x + 1) = x\Gamma(x)$, we find

$$
\begin{aligned}
xL_{m-j}^{\alpha+1}(x) = &\sqrt{m - j + 1}\sqrt{m/d - j + 1} \cdot L_{m-j+1}^{\alpha+1}(x) \\
&+ (m/d + m - 2j + 2) \cdot L_{m-j}^{\alpha+1}(x) \\
&+ \sqrt{m - j}\sqrt{m/d - j + 1} \cdot L_{m-j-1}^{\alpha+1}(x).
\end{aligned}
\tag{2.8}
$$

Letting $x = l_i$, where l_i is one of the m (real) zeros of $L_m^{\alpha+1}(x)$, it follows that

$$
(M_{\infty,d} v_i)_j = \left(\frac{d}{m} l_i \cdot v_i \right)_j
\tag{2.9}
$$

for $j = 2, \ldots, m - 1$. On the other hand, when $j = 1$, we use the recurrence relation [2, (22.7.30)]:

$$
\tilde{L}_k^\alpha = \tilde{L}_k^{\alpha+1} - \tilde{L}_{k-1}^{\alpha+1}
$$

in (2.6) before normalizing. Since $L_m^\alpha(l_i) = 0$, equation (2.8) then simplifies to

$$
l_i \cdot L_{m-1}^{\alpha+1}(l_i) = (m/d) \cdot L_{m-1}^{\alpha+1}(l_i) + \sqrt{m - 1}\sqrt{m/d} \cdot L_{m-2}^{\alpha+1}(l_i),
\tag{2.10}
$$

which yields

$$
(M_{\infty,d} v_i)_1 = \left(\frac{d}{m} l_i \cdot v_i \right)_1.
$$

Similar reasoning when $j = m$ shows that (2.9) holds for all $j = 1, \ldots, m$.

2.2 The resolvent expansion

In this section we show that the normalized trace of the resolvent is well approximated by the Stieltjes transform of the Marchenko-Pastur density for $\Im z > m^{-1/4+\epsilon}$, $\epsilon > 0$. Specifically, we have the following

Proposition 2.1. *Let $\epsilon, \kappa > 0$. For any z such that $\Im z > m^{-1/4+\epsilon}$ and*

$$
\lambda_- + \kappa < \Re z < \lambda_+ - \kappa,
$$

we have

$$
|s_\beta(z) - s_\infty(z)| = O_{\epsilon,\kappa}(m^{-\epsilon/8})
$$

with overwhelming probability.

The choice of the scale $m^{-1/4+\epsilon}$ in Proposition 2.1 is somewhat arbitrary, given that the inductive argument of Section 4 will gradually improve any initial convergence result to the optimal scale $m^{-1+\epsilon}$. The exponent $\frac{1}{4}-\epsilon$ was chosen because it represents a scale accessible without detailed information on the size of the entries of the resolvent. Using the estimates for Laguerre polynomials in [16], and the method of [15] to obtain a lower bound for $\|v_i\|$, one can prove sharper bounds than (2.14) for $|(M_{\infty,d}-z)^{-1}_{ij}|$ when $|i-j| \leq m^{1/4}$. The expansion (2.12) can then be summed for values of $\Im z$ smaller than $m^{-1/4}$, but still much larger than $m^{-1+\epsilon}$. We will not pursue this here.

The proof of Proposition 2.1 proceeds by a resolvent expansion comparing the trace of the resolvent of $M_{\beta,d}$ to that of the deterministic matrix $M_{\infty,d}$. The precise spectral information available for the matrix $M_{\infty,d}$ allows us to calculate the large m limit of the normalized trace $s_\infty(z)$, and to control the resolvent expansion (2.12). Starting from (2.1), write:

$$(M_{\beta,d}-z)^{-1} = (M_{\infty,d}-z)^{-1} - (M_{\infty,d}-z)^{-1}\Delta(M_{\beta,d}-z)^{-1}.$$

Upon iteration, this yields

$$
(M_{\beta,d}-z)^{-1} = (M_{\infty,d}-z)^{-1}
$$
$$
+ \sum_{k=1}^{l-1} \left(-(M_{\infty,d}-z)^{-1}\Delta\right)^k (M_{\infty,d}-z)^{-1} + \left(-(M_{\infty,d}-z)^{-1}\Delta\right)^l (M_{\beta,d}-z)^{-1}. \quad (2.11)
$$

Taking traces and normalizing:

$$
s_{\beta,d}(z) = s_\infty(z) + \frac{1}{m}\sum_{k=1}^{l-1} \mathrm{tr}\left(-(M_{\infty,d}-z)^{-1}\Delta\right)^k (M_{\infty,d}-z)^{-1}
$$
$$
+ \frac{1}{m}\mathrm{tr}(-(M_{\infty,d}-z)^{-1}\Delta)^l (M_{\beta,d}-z)^{-1}. \quad (2.12)
$$

By the estimate (3.4) for the normalized trace $s_\infty(z)$ of $(M_{d,\infty}-z)^{-1}$, we can replace the first term on the right with $s_{MP,d}(z)$, introducing an $O_\kappa(m^{-\epsilon})$ error. It suffices to show that the remaining terms also decay at least at an $O(m^{-\epsilon})$ rate as $m \to \infty$.

For $1 \leq i,j \leq m$, we let

$$R_{ij} = R_{ij}(z) := (M_{\infty,d}-z)^{-1}_{ij},$$

and

$$R^{\beta,d}_{ij} = R^{\beta,d}_{ij}(z) := (M_{\beta,d}-z)^{-1}_{ij}.$$

Expand the kth term in of the sum in (2.12) to find:

$$
(-1)^k \sum_{1 \leq i_1,\ldots i_{k+1} \leq m} R_{i_1 i_2}\Delta_{i_2 i'_2} R_{i'_2,i_3} \cdots R_{i'_k i_{k+1}}\Delta_{i_{k+1} i'_{k+1}} R_{i'_{k+1} i_1},
$$

where $i'_l = i_l$, $i_l + 1$ or $i_l - 1$. Taking absolute values and using (2.2), this is bounded by

$$
m^{-(k+2)(1/2-c)} \sum_{1 \leq i_1,\ldots i_{k+1} \leq m} |R_{i_1 i_2}||R_{i'_2,i_3}| \cdots |R_{i'_{k+1} i_1}|, \quad (2.13)
$$

with overwhelming probability.

The next lemma gives an estimate for the entries R_{ij} of the resolvent which is efficient when i and j are widely separated.

Lemma 2.2. *For any $1 \le i, j \le m$, we have*

$$|R_{ij}(z)| \le C(\Im z)^{-1} \cdot \exp(-c_d(\Im z) \cdot |i - j|). \tag{2.14}$$

Several results on exponential decay of the resolvent entries could be used to obtain Lemma 2.2. We will use a slight variant of the estimate of Combes-Thomas type developed by M. Aizenman in the context of localization for discrete random Schrödinger operators:

Lemma 2.3 (M. Aizenman, [1], Lemma II.1). *Let Γ be a countable set with a metric $d : \Gamma \times \Gamma \to [0, \infty)$, and let H be a self-adjoint operator on $\ell^2(\Gamma)$ whose off-diagonal elements are exponentially summable:*

$$S_\alpha = \sup_x \sum_{y \ne x} |H(x,y)| e^{\alpha d(x,y)} < \infty. \tag{2.15}$$

Then, for energies not in the spectrum of H, with

$$\Delta = \mathrm{dist}(z, \sigma(H)) > 0,$$

we have

$$\left| (H - z)^{-1}_{xy} \right| \le \frac{2}{\Delta} \exp\left(-\frac{\alpha\Delta}{\Delta + 2S_\alpha} \cdot d(x,y) \right).$$

Note that the hypothesis (2.15) differs from the one in Lemma II.1 in [1], but the same proof works also for Lemma 2.3 as stated.

To bound (2.13), we can assume that $i_j = i'_j$. Indeed, the bound (2.14) only changes by a constant factor when the index i is changed by a unit. Consider the sum over i_1 in (2.13):

$$\sum_{i_1=1}^{m} |R_{i_1 i_2}||R_{i_{k+1} i_1}|.$$

We split this sum into three regions:

$$\left(\sum_{|i_1 - i_2| \ge m^{1/4}} + \sum_{|i_1 - i_2| \le m^{1/4}, |i_1 - i_{k+1}| \ge m^{1/4}} + \sum_{|i_1 - i_2| \le m^{1/4}, |i_1 - i_{k+1}| \le m^{1/4}} \right) |R_{i_1 i_2}||R_{i_{k+1} i_1}|.$$

By (2.14), the first two sums are bounded by

$$Cm^{1/2} \exp(-m^\epsilon).$$

We are reduced to considering the sum

$$\sum_{|i_1 - i_2| \le m^{1/4}} |R_{i_1 i_2}||R_{i_1 i_{k+1}}|. \tag{2.16}$$

The second factor in each summand of (2.16) is estimated as follows:

$$|R_{i_1 i_{k+1}}| = \left| \sum_{i=1}^{m} \frac{u_i(i_1) u_i(i_{k+1})}{\lambda_i - z} \right| \le \left(\sum_{i=1}^{m} \frac{|u_i(i_1)|^2}{|\lambda_i - z|} \right)^{1/2} \left(\sum_{i=1}^{m} \frac{|u_i(i_{k+1})|^2}{|\lambda_i - z|} \right)^{1/2}. \tag{2.17}$$

The first factor on the right side of the inequality (2.17) is no greater than $m^{1/8 - \epsilon/2}$, since $\|u_i\| = 1$ and

$$\frac{1}{|\lambda_i - z|} \le (\Im z)^{-1}.$$

In summary, (2.16) is bounded by

$$m^{1/8-\epsilon/2} \left(\sum_{i=1} \frac{|u_i(i_{k+1})|^2}{|\lambda_i - z|} \right)^{1/2} \cdot \sum_{|i_1-i_2|\leq m^{1/4}} |R_{i_1 i_2}| \leq m^{1/8+1/2-(3/2)\epsilon} \cdot \left(\sum_{i=1} \frac{|u_i(i_{k+1})|^2}{|\lambda_i - z|} \right)^{1/2}.$$

We proceed to sum over i_2 in (2.13):

$$\sum_{i_2} |R_{i_2 i_3}| \leq \sum_{|i_2-i_3|\leq m^{1/4}} |R_{i_2 i_3}| + C \leq Cm^{1/4} \cdot m^{1/4-\epsilon}.$$

We can now repeatedly sum over $i_3, \ldots, i_{k-1}$, using (2.14) at every step. Each summation results in an additional factor of $m^{1/2-\epsilon}$. Thus (2.13) is bounded by

$$C^{k-2} m^{1/8-\epsilon/2} \cdot m^{(k-1)/2-(k-1)\epsilon} \sum_{i_k} \left(\sum_{i=1}^{m} \frac{|u_i(i_k)|^2}{|\lambda_i - z|} \right)^{1/2} \cdot \sum_{|i_k-i_{k+1}|\leq m^{1/4}} |R_{i_k i_{k+1}}|. \quad (2.18)$$

Using the Cauchy-Schwarz inequality as in (2.17), we have

$$|R_{i_k i_{k+1}}| \leq m^{1/8-\epsilon/2} \cdot \left(\sum_{i=1} \frac{|u_i(i_{k+1})|^2}{|\lambda_i - z|} \right)^{1/2}.$$

By (3.4) and Corollary 4.3,

$$\sum_{i_{k+1}} \sum_{i=1} \frac{|u_i(i_{k+1})|^2}{|\lambda_i - z|} = \sum_{i} \frac{1}{|\lambda_i - z|} \leq Cm \log m.$$

Inserting the previous two inequalities into (2.18), we find that the sum (2.13) is no greater than

$$C^k \log m \cdot m^{(k+2)/2-k\epsilon}.$$

Choosing c smaller than $\epsilon/2$ in (2.2), this last quantity is $O(m^{-k\epsilon/2})$ with overwhelming probability. As for the final term in (2.12), it is bounded by

$$m^{-(k+2)(1/2-c)} \sum_{1\leq i_1,\ldots i_k \leq m} |R_{i_1 i_2}||R_{i'_2,i_3}| \cdots |R_{i'_{k-1} i_k}| \cdot m^{1/4-\epsilon},$$

since

$$|R_{ij}^{\beta,d}(z)| \leq m^{1/4-\epsilon}$$

for each i and j, provided $\Im z > m^{1/4-\epsilon}$. Performing each of the k sums using (2.14) as previously, we find that the sum is bounded by

$$C^k m^{-k\epsilon/2} \cdot m^{1/4},$$

with overwhelming probability. Letting k in (2.12) be larger than $8/\epsilon$, we obtain

$$s_\beta(z) = s_\infty(z) + O(m^{-\epsilon/2}). \quad (2.19)$$

with overwhelming probability. Combined with the approximation (3.4), the relation (2.19) implies Proposition 2.1.

3 Convergence for $\beta = \infty$

In this section, we identify the limit of the resolvent

$$s_\infty(z) = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{\lambda_j - z} = \frac{1}{m}\frac{\left(\tilde{L}_m^\alpha(mz/d)\right)'}{\tilde{L}_m^\alpha(mz/d)} \tag{3.1}$$

for z as close as $m^{-1/2+\epsilon}$ to the limiting spectrum, but away from its edges. The eigenvalues λ_i are given by (2.3). In [20], the corresponding quantity for Hermite polynomials was estimated using asymptotics for these orthogonal polynomials in the complex plane. Here, we use a differential equation for $s_\infty(z)$ derived from the ODE satisfied by Laguerre polynomials ([2, (22.6.15)], [3, Section 4.5, (4.5.1)]):

$$x(\tilde{L}_m^\alpha)''(x) + (\alpha + 1 - x)(\tilde{L}_m^\alpha)'(x) + m\tilde{L}_m^\alpha(x) = 0, \tag{3.2}$$

with $\alpha = m(1/d - 1)$. Differentiating the ratio on the right of (3.1) and using (3.2), we find an equation for $s_\infty(z)$ (see [8, Section 2.2.2, Eqn. (20)]),

$$zd(s_\infty(z))^2 + s_\infty(z)(d - 1 + z) + 1 + \frac{d}{m}s_\infty(z) + \frac{dz}{m}s_\infty(z)' = 0. \tag{3.3}$$

To solve the equation approximately, we treat the final two terms

$$\varepsilon = \frac{d}{m}s_\infty(z) + \frac{dz}{m}s_\infty(z)'$$

as error terms, and use the rough estimates

$$|s_\infty(z)| \le m^{1/2-\epsilon}, |s_\infty'(z)| \le m^{1-2\epsilon}.$$

These follow at once from $|\lambda_i - z| \ge m^{1/2-\epsilon}$, so that $\varepsilon = O(m^{-\epsilon})$. For z in the region

$$\{\Im z > m^{-1/2+\epsilon}, |z - \lambda_\pm| > \kappa, |z| \le 4\},$$

we find the solutions

$$s^\pm(z) = \frac{z + d - 1 \pm \sqrt{(z - \lambda_+)(z - \lambda_-)}}{2dz} + O_\kappa(m^{-\epsilon}).$$

The solution s^- satisfies $\Im s^- > 0$ for $\Im z > 0$, and we conclude that the right side of the equation is an approximation for $s_{\beta,d}$:

$$s_\infty(z) = s_{MP,d}(z) + O_\kappa(m^{-\epsilon}). \tag{3.4}$$

We expect that the error term can be replaced by $O(1/m)$, but (3.4) is sufficient for our purposes.

4 Inductive argument

In this section, we improve the convergence of the density on short scales from the level $m^{-1/4+\epsilon}$ to the optimal level of $m^{-1+\epsilon}$ by an inductive argument.

Theorem 4.1. *Let* $s(z) = s_{\beta,d}(z)$ *be the Stieltjes transform of the measure induced by the eigenvalues of the normalized β-Laguerre ensemble matrix $M_{\beta,d}$. Let $s_{MP,d}(z)$ be the Stieltjes transform of the Marchenko-Pastur distribution. Then, with overwhelming probability,*

$$\sup_{z \in D_{\epsilon,\kappa}} |s(z) - s_{MP,d}(z)| = o(1) \tag{4.1}$$

where the domain D is defined as

$$D_{\epsilon,\kappa} := \{z : \Im z > m^{-1+\epsilon}, \lambda_- + \kappa < \Re z < \lambda_+ - \kappa, |z| \le 4\}.$$

To prove Theorem 4.1, we need three facts about the tridiagonal models and Stieltjes transforms. The first can be found in [6]:

Proposition 4.2. *One can diagonalize*

$$M_{\beta,d} = Q\Lambda Q^*$$

such that the first row of Q is independent of Λ and consists of independent entries with χ_β distribution, normalized to unit norm.

The next corollary establishes the link between control of Stieltjes transform and control of the eigenvalue density. See for example [13]:

Corollary 4.3. *Let $z = E + i\eta$ and $m^{-1+\epsilon} \leq \eta \leq \frac{1}{2}E$, $\frac{1}{2}\lambda_- \leq E \leq 10$ and $\epsilon, \tau > 0$. Suppose that one has the bound*

$$|s_{\beta,d,m}(z) - s_{MP,d}(z)| \leq \tau$$

with overwhelming probability for all such z. Then for any interval I in $[\lambda_- + \epsilon, \lambda_+ - \epsilon]$ with $|I| \geq \eta$, one has

$$\left| N_I - n \int_I \rho_{MP,d}(y)\, \mathrm{d}y \right| \leq C_\epsilon \tau n |I|$$

with overwhelming probability, where N_I denotes the number of eigenvalues in I.

Finally, we shall need the following standard result. A proof can be found in [20]:

Lemma 4.4. *Suppose the Marchenko-Pastur distribution holds at level m^a for some $-1 < a \leq 0$, that is, for any $c > 0$,*

$$|s(z) - s_{MP,d}(z)| < c$$

for sufficiently large m, with overwhelming probability.
Then we have

$$\frac{1}{m}\sum_j \frac{1}{|\lambda_j - z|^2} \leq C(\Im z)^{-2} m^a \log m \tag{4.2}$$

for any z such that $m^{-1} < \Im z < m^a$.

The proof of Theorem 4.1 is a combination of the following propositions:

Proposition 4.5. *Suppose the Marchenko-Pastur law holds at level m^a for some $-1 < a \leq 0$, and that we have*

$$|R_{11}^{\beta,d}(z) - s_{MP,d}(z)| = O(1) \tag{4.3}$$

for $\Im z > m^a$ with overwhelming probability. Then, for any $\delta > 0$,

$$|R_{11}^{\beta,d}(z) - s_{MP,d}(z)| = o(1) \tag{4.4}$$

holds for z such that $\Im z > m^{(a-1)/2+\delta}$ with overwhelming probability.

For a given sequence X_n of random variables, we write

$$X_n = o(1)$$

with overwhelming probability if, for every $c > 0$, the event $\{|X_n| \leq c\}$ has overwhelming probability.

Proposition 4.6. *Suppose that with overwhelming probability, $|R_{11}^{\beta,d}(z) - s_{MP,d}(z)| = o(1)$ for z such that $\Im z > m^a$ for some $-1 < a \leq 0$, then we have an improved Marchenko-Pastur law, that is:*

$$\left| \frac{1}{m} \sum_j \frac{1}{\lambda_j - z} - s_{MP,d}(z) \right| = o(1) \tag{4.5}$$

for z such that $\Im z > m^{(a-1)/2+\delta}$ for any $\delta > 0$ with overwhelming probability.

Proof of Theorem 1.1. The condition (4.3) holds trivially if $\Im z \geq 1$. The result follows by repeatedly applying Propositions 4.5 and 4.6. $\qquad\square$

What remains is the proof of the two propositions above:

Proof of Proposition 4.5. By Schur's complement, we have the following relation

$$R_{11}^{\beta,d}(z) = \frac{1}{a_{11}^2 - z - a_{11}^2 a_{21}^2 \frac{\widehat{R}_{11}^{\beta,d}}{1 + a_{21}^2 \widehat{R}_{11}^{\beta,d}}}. \tag{4.6}$$

Here a_{11} denotes the normalized $(1,1)$-entry of the bidiagonal matrix B_β, $a_{2,1}$ the normalized $(2,1)$-entry of B_β, $\widehat{R}^{\beta,d}$ is the resolvent of $\widehat{M_\beta} := \frac{d}{\beta m} \widehat{B}_\beta \widehat{B}_\beta^t$ where $\widehat{B}_\beta$ is formed by removing the first row and column of B_β. We remark that a_{11}^2 is distributed as $\frac{d}{\beta m} \chi_{2a}$ and a_{21}^2 distributed like $\frac{d}{\beta m} \chi_{\beta(m-1)}$. Lastly, a_{11}, a_{21} and $\widehat{R}$ are independent.

By the argument in Section 3, we need only show that $R_{11}^{\beta,d}$ satisfies the approximate functional equality

$$R_{11}^{\beta,d}(z) + \frac{1}{d - 1 + z + zd \cdot R_{11}^{\beta,d}(z)} = o(1), \tag{4.7}$$

with overwhelming probability for z such that $\Re z \in [\lambda_- + \kappa, \lambda_+ - \kappa]$ and $\Im z \geq m^{(a-1)/2+\delta}$, with $\delta > 0$ arbitrary. We first note that, due to the restriction on $\Re z$ (in particular, $\Re z > \kappa$ since $\lambda_- > 0$), equation (4.7) is equivalent to:

$$zd \cdot (R_{11}^{\beta,d}(z))^2 + (d - 1 + z) \cdot R_{11}^{\beta,d}(z) + 1 = o(1) \tag{4.8}$$

where from now on we suppress the superscript β, d.

Rewriting equation (4.6), we have

$$za_{21}^2 \cdot \widehat{R_{11}}^{\beta,d} R_{11}^{\beta,d} + zR_{11}^{\beta,d} + a_{21}^2 \widehat{R_{11}}^{\beta,d} - a_{11}^2 R_{11}^{\beta,d} + 1 = 0. \tag{4.9}$$

We can further rewrite the above as

$$zd \cdot \frac{m-1}{m} \cdot \widehat{R_{11}}^{\beta,d} R_{11}^{\beta,d} + zR_{11}^{\beta,d} + d\frac{m-1}{m} \cdot \widehat{R_{11}}^{\beta,d} - R_{11}^{\beta,d} + 1$$
$$+ \left(z \cdot \widehat{R_{11}}^{\beta,d} R_{11}^{\beta,d} + \widehat{R_{11}}^{\beta,d} \right) \left(a_{21}^2 - d\frac{m-1}{m} \right) - (a_{11}^2 - 1) \cdot R_{11}^{\beta,d} = 0. \tag{4.10}$$

It suffices, therefore, to show that

$$\frac{m-1}{m} \cdot \widehat{R_{11}}^{\beta,d} = R_{11}^{\beta,d} + o(1) \tag{4.11}$$

and

$$\left(z \cdot \widehat{R_{11}}^{\beta,d} R_{11}^{\beta,d} + \widehat{R_{11}}^{\beta,d} \right) \left(a_{21}^2 - d\frac{m-1}{m} \right) - (a_{11}^2 - 1) \cdot R_{11}^{\beta,d} = o(1) \tag{4.12}$$

for the set of z we are interested in.

We first note that a_{21}^2 has distribution $\frac{d}{\beta m}\chi_{\beta(m-1)}^2$ and so

$$\left|a_{21}^2 - d\frac{m-1}{m}\right| \leq Cm^{-1/2+\delta}$$

with overwhelming probability for all $\delta > 0$. Similarly, a_{11}^2 has the distribution of $\frac{d}{\beta m}\chi_{\beta m/d}^2$ and so

$$|a_{11}^2 - 1| \leq Cm^{-1/2+\delta}$$

with overwhelming probability for all $\delta > 0$. The assumption of the proposition implies that for all w with $\Im w > m^a$, we have

$$|R_{11}^{\beta,d}(w)| \leq C,$$

for some constant C. By (4.11), which will be proved independently below, we also have

$$|\widehat{R_{11}}^{\beta,d}(w)| \leq C.$$

The inequality

$$|R_{11}^{\beta,d}(z)| \leq \frac{w}{z}|R_{11}^{\beta,d}(w)|,$$

for $\Im z \leq \Im w$, now implies that approximation (4.12) holds with overwhelming probability whenever $\Im z > m^{(a-1)/2+\delta}$.

We turn our attention to (4.11). Firstly, we write

$$\widehat{R}_{11}^{\beta,d} = \sum_j \frac{\widehat{q}_j^2}{\widehat{\lambda}_j - z},$$

and similarly

$$R_{11}^{\beta,d} = \sum_j \frac{q_j^2}{\lambda_j - z},$$

where $(\widehat{q}_1, \ldots \widehat{q}_{m-1})$ is the first row of the eigenvectors for $\widehat{M_\beta}$, $\widehat{\lambda}_j$ are the eigenvalues and $(q_1, \ldots, q_m)$ is the first row of eigenvectors for M_β, and λ_j the eigenvalues.

We write

$$\begin{aligned}
R_{11}^{\beta,d} - \frac{m-1}{m}\cdot\widehat{R}_{11} = &\sum_j \frac{q_j^2}{\lambda_j - z} - \mathbb{E}_q\sum_j \frac{q_j^2}{\lambda_j - z} \\
&+ \mathbb{E}_q\sum_j \frac{q_j^2}{\lambda_j - z} - \frac{m-1}{m}\cdot\mathbb{E}_{\widehat{q}}\sum_j \frac{\widehat{q}_j^2}{\widehat{\lambda}_j - z} \\
&+ \frac{m-1}{m}\cdot\left(\mathbb{E}_{\widehat{q}}\sum_j \frac{\widehat{q}_j^2}{\widehat{\lambda}_j - z} - \sum_j \frac{\widehat{q}_j^2}{\widehat{\lambda}_j - z}\right),
\end{aligned} \quad (4.13)$$

where $\mathbb{E}_X$ denotes the expectation with respect to the random variables X.

By Proposition 4.2, q and λ_j are independent and so are $\hat{q}$ and $\hat{\lambda}_j$. We write the first term in (4.13) as

$$\sum_j \frac{q_j^2}{\lambda_j - z} - \mathbb{E}_q\sum_j \frac{q_j^2}{\lambda_j - z} = \sum_j \frac{q_j^2 - \frac{1}{m}}{\lambda_j - z}.$$

Since q and λ_j are independent, we can condition on λ_j and apply Proposition 4.7 below to the sum to conclude that the right side of the previous equation is bounded with overwhelming probability by

$$\frac{1}{m^{1-c}}\left(\sum_j \frac{1}{|\lambda_j - z|^2}\right)^{1/2}.$$

By Lemma 4.4, the latter is bounded by $C(\Im z)^{-1}m^{-1/2+a/2+c}\log m$ for some constant C and any $c > 0$. The fifth and sixth terms of equation (4.13) are similarly bounded. To deal with the middle two terms, we use the interlacing property of the eigenvalues of a matrix and its minor, and then the interlacing property of the eigenvalues of a matrix and a rank 1 perturbation, to obtain that these terms are bounded by

$$\frac{1}{m\Im z}.$$

We end the proof of Proposition 4.5 with the following simple variant of McDiarmid's inequality, a proof of which can be found in the appendix of [20]:

Proposition 4.7. *Let $X_1, \ldots, X_m$ be independent subgaussian random variables. Let $\Omega \subset \mathbb{R}^n$ be such that*

$$(X_1, \ldots, X_m) \in \Omega$$

with overwhelming probability. Let F be a real function of m real variables such that if $x_1, \ldots, x_m, \tilde{x}_i \in \Omega$, then

$$|F(x_1, \ldots, x_n) - F(x_1, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots, x_n)| \leq c_i$$

for all $1 \leq i \leq m$, and outside of Ω, F is bounded by a polynomial in m. Then for any $\lambda > 0$, one has

$$\mathbb{P}(|F(X) - \mathbb{E}(F(X))| \geq \lambda\sigma) \leq C\exp(-c\lambda^2)$$

for some absolute constants $C,c > 0$, and $\sigma = \sum_{i=1}^{m} c_i^2$.

$\square$

Proof of Proposition 4.6. The proof is similar to that of Proposition 4.5. By the assumption, it suffices to establish that

$$\left| \frac{1}{m} \sum_j \frac{1}{\lambda_j - z} - R_{11}^{\beta,d}(z) \right| = o(1)$$

with overwhelming probability. The difference inside the absolute value sign is $\sum_j \frac{\frac{1}{m} - q_j^2}{\lambda_j - z}$. The statement now follows by another concentration argument and Lemma 4.4. $\square$

5 Some extensions

The resolvent expansion in Section 2 allows for a quick proof of convergence of the eigenvalue distribution for tridiagonal matrices with independent entries "close to" a Jacobi matrix whose limiting spectral density is known. Consider a sequence of random tridiagonal matrices given by

$$A_n = \begin{pmatrix} a_n & b_{n-1} & & & & \\ b_{n-1} & a_{n-1} & b_{n-2} & & & \\ & & \ddots & & & \\ & & & b_1 & a_1 & b_0 \\ & & & & b_0 & a_0 \end{pmatrix}, \tag{5.1}$$

where a_j, b_j, $1 \leq j \leq n$. Suppose that

$$\frac{b_j - \bar{b}_j}{n^\alpha} = o(1), \tag{5.2}$$

$$\frac{a_j - \bar{a}_j}{n^\alpha} = o(1) \tag{5.3}$$

uniformly in $1 \leq j \leq n$ with overwhelming probabilty, for some $\alpha > 0$, where

$$\bar{b}_j = \mathbb{E}b_j,$$
$$\bar{a}_j = \mathbb{E}a_j.$$

If the deterministic tridiagonal matrix $\frac{\mathbb{E}A_n}{n^\alpha}$ has a limiting spectral density μ, that is, if the spectral measures μ_n of $n^{-\alpha} \cdot \mathbb{E}A_n$ converge weakly to some measure μ, then

$$\frac{1}{n} \operatorname{tr} \left(n^{-\alpha} \mathbb{E}A_n - z \right)^{-1} = \int \frac{1}{z - \lambda} \mu(\mathrm{d}\lambda) + o(1)$$

for z in any compact subset D of $\{\Im z > 0\}$. Using Lemma 2.2 as in the proof of Proposition 2.1 to perform a one-step resolvent rexpansion, we have, for $z \in D$:

$$\frac{1}{n} \operatorname{tr} \left(n^{-\alpha} A_n - z \right)^{-1} - \int \frac{1}{z - \lambda} \mu(\mathrm{d}\lambda = o(1) \tag{5.4}$$

with overwhelming probability. Thus the eigenvalue distribution of $n^{-\alpha}A_n$ almost surely converges weakly to μ.

As an application, consider the following example from [19]. Consider the Jacobi matrix

$$J_n = \begin{pmatrix} 0 & (n-1)^\alpha & & & & \\ (n-1)^\alpha & 0 & (n-2)^\alpha & & & \\ & (n-2)^\alpha & \ddots & & & \\ & & & & & 2^\alpha \\ & & & & 2^\alpha & 0 & 1 \\ & & & & & 1 & 0 \end{pmatrix}.$$

The moments of the rescaled matrix $n^{-\alpha}J_n$ converge to those of the *Nevai-Ullman distribution* [22], [17, Section 5]:

$$\frac{1}{n} \operatorname{tr}(n^{-\alpha}J_n)^k \to \mu_k,$$

$$\mu_k = \int x^k \, \nu_\alpha(\mathrm{d}x),$$

$$\nu_\alpha(\mathrm{d}x) = \mathbf{1}_{[-2,2]}(x) \frac{1}{\alpha\pi} \int_{|x|/2}^{2} \frac{t^{-1+1/\alpha}}{\sqrt{4-t^2}} \, \mathrm{d}t.$$

By a standard density argument,

$$\frac{1}{n} \operatorname{tr}(n^{-\alpha}J_n - z)^{-1} = \int \frac{1}{\lambda - z} \, \nu_\alpha(\mathrm{d}\lambda) + o(1).$$

If we let

$$b_j = (J_n)_{jj-1} + n^\beta X_j,$$
$$a_j = (J_n)_{jj} + n^\beta Y_j,$$

where X_j, $1 \leq j \leq n-1$ and Y_j, $1 \leq j \leq n$, are mean zero random variables whose kth moment is bounded uniformly in j for all k, then for any $\beta < \alpha$ (5.4) holds for the matrix $n^{-\alpha}A_n$. If we instead require that

$$\mathbb{E}\left(\frac{b_n}{n^\alpha}\right)^k \to 1 \tag{5.5}$$

$$\sup_n \mathbb{E}|a_n|^k < \infty.$$

for $k = 1, 2$, then for any $c > 0$, we have

$$\mathbb{P}\left(\left|\frac{b_j - 1}{n^\alpha}\right| \geq c\right) \leq \frac{1}{c^2}\mathbb{E}\left|\frac{b_j - 1}{n^\alpha}\right|^2 = o(1),$$

uniformly in j, and a similar bound for a_j. The convergence (5.4) with $\mu = \nu_\alpha$ holds in probability. Popescu [19] assumes (5.5) holds for all k and obtains convergence of the moments of $n^{-\alpha}A_n$ as well as almost sure convergence to the limiting distribution. Note that [19] contains more general results that apply also to cases not easily accessible by our method.

Given more precise information on the size of the entries of A_n or the rate of convergence of the Stieltjes transform for the deterministic model, one can improve on the above. To give a simple example, we introduce a "positive temperature" version of the Jacobi matrices associated to the orthogonal polynomials $p_n(x)$, $n \geq 0$, for the measures

$$e^{-x^{2m}}\,\mathrm{d}x. \tag{5.6}$$

Define the density

$$g_{n,m}(x) = \mathbf{1}_{[0,\infty)}\frac{x^{n-1}e^{-x^{2m}}}{\frac{1}{2m}\Gamma(\frac{n}{2m})}.$$

Let the random variable X_j have distribution given by $g_{j,m}$, with all the X_j independent. We have

$$\mathbb{E}X_j = \frac{\Gamma\left(\frac{j+1}{2m}\right)}{\Gamma\left(\frac{j}{2m}\right)} = \frac{2^{-1/2m}}{e^{1/2m}}\left(\frac{j+1}{m}\right)^{1/2m}\cdot\left(1+\frac{1}{j}\right)^{j/2m} + O(1/j),$$

and, by a straightforward computation:

$$\mathbb{P}\left(|X_j - \mathbb{E}X_j| \geq n^\epsilon\right) \leq Ce^{-n^\epsilon},$$

uniformly for $1 \leq j \leq n$. Consider the matrix (5.1), with entries

$$b_j = \frac{e^{1/2m}}{2^{1+1/2m}}\cdot\left(1+\frac{1}{j}\right)^{j/2m}X_j,$$

$$a_j = N(0,1).$$

Let $J_n = (m\alpha_m/n)^{1/(2m)}\bar{A}_n$, with $\bar{A}_n = \mathbb{E}A_n$, and

$$\alpha_m = \prod_{j=1}^m \frac{2j-1}{2j}.$$

The matrix J_n approximates a Jacobi matrix associated with the orthogonal polynomials for the measure (5.6). Indeed, by [5, Eq. (2.33)], we have

$$\tilde{b}_n = \frac{1}{2}\left(\frac{n+1}{\alpha_m m}\right)^{1/2m} + O(1/n^{2-1/2m}).$$

Here $\tilde{b}_n$ is the coefficient in the three-term recurrence relation:

$$xp_n(x) = \tilde{b}_{n+1}p_{n+1}(x) + \tilde{a}_n p_n(x) + \tilde{b}_{n-1}p_{n-1}(x).$$

Note that $\tilde{a}_n = 0$. The eigenvalues of the matrix

$$\tilde{J}_n = \left(\frac{m\alpha_m}{n}\right)^{1/2m}\cdot\begin{pmatrix} 0 & \tilde{b}_{n-1} & & & & \\ \tilde{b}_{n-1} & 0 & \tilde{b}_{n-2} & & & \\ & \tilde{b}_{n-2} & \ddots & & & \\ & & & & \tilde{b}_1 & \\ & & & \tilde{b}_1 & 0 & \tilde{b}_0 \\ & & & & \tilde{b}_0 & 0 \end{pmatrix},$$

are the (rescaled) zeros of the nth orthogonal polynomial with respect to (5.6). Their limiting density can be explicitly computed; see [5, Eq. (2.4)]. It has the form

$$\mu(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{1-x^2}h_m(x)\mathbf{1}_{[-1,1]}(x)\,\mathrm{d}x,$$

where $h_m(x)$ is a polynomial of degree $2m-2$. A calculation as in [20] using the Riemann-Hilbert asymptotics in [5] shows that the normalized trace of $(\tilde{J}_n - z)^{-1}$ approximates the Stieltjes transform of the limiting density $\mu(\mathrm{d}x)$ with precision $O(1/n)$ for $\Im z > n^{-1+\epsilon}$ and z away from ± 1. By the resolvent expansion argument of Section 2 and Corollary 4.3, convergence of the empirical eigenvalue distribution for $(\alpha_m n)^{-1/2m}A_n$ holds for intervals of size $|I| \geq n^{-1/4m+\epsilon}$ strictly inside the bulk of the limiting density.

References

[1] Aizenman, M., Localization at weak disorder: some elementary bounds, *Rev. Math. Phys.* (1994), 6. MR-1301371

[2] Abraomowitz, M., Segun, I., eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1964.

[3] Beals, R., Wong, R., *Special Functions. A Graduate Text*, Cambridge studies in advanced mathematics **126**, Cambridge University Press, 2010. MR-2683157

[4] Cacciapuoti, C., Maltsev, A., Schlein, B., *Local Marchenko-Pastur Law at the Hard Edge of Sample Covariance Matrices*, preprint, arXiv:1206.1730. MR-3088802

[5] Deift, P., et al. *Strong asymptotics of orthogonal polynomials with respect to exponential weights.* Communications on Pure and Applied Mathematics **52**, 12, 1999. MR-1711036

[6] Dumitriu, I.: *Eigenvalue statistics for beta ensembles.* Ph.D. thesis, MIT, 2003. MR-2717094

[7] Dumitriu, I. and Edelman, A. *Matrix models for beta ensembles.*, J. Math. Phys. **47**, 2006. J. Math. Phys., **43**, 11, 5830-5847, 2002. MR-1936554

[8] Dumitriu, I. and Edelman, A. *Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models.* J. Math. Phys. **47**, 6, 2006. MR-2239975

[9] Erdös, L., Schlein, B. and Yau, H.-T., *Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices*, Ann. Prob. **37**, 815–852, 2009. MR-2537522

[10] Erdös, L., Schlein, B. and Yau, H.-T., *Local semicircle law and complete delocalization for Wigner random matrices*, Comm. Math. Phys. **287**, 641–655, 2009. MR-2481753

[11] Erdös, L., Schlein, B. and Yau, H.-T., *Wegner estimate and level repulsion for Wigner random matrices*, IMRN **2010**, 436–479, 2009. MR-2587574

[12] Erdös, L., Schlein, B., Yau, H.-T. and Yin, J. : *The local relaxation flow approach to universality of the local statistics for random matrices.* Ann. Inst. H. Poincaré B: Probab. Statist. **48**, 1–46, 2012. MR-2919197

[13] Erdös, L., Ramirez, J.A., Schlein, B., Yau, H.-T.. : *Universality of Sine-Kernel for Wigner matrices with a small Gaussian perturbation.* Elec. J. Probab., **15**, 18, 2010. MR-2639734

[14] Erdös, L., Knowles, A., Yau, H.-T. : *The Local Semi-circle Law for a General Class of Random Matrices*, preprint, arXiv:1212.0164.

[15] Krasikov, I., *Nonnegative quadratic forms and bounds on orthogonal polynomials*, J. Approx. Th. **111**, 2001. MR-1840019

[16] Krasikov, I., *Inequalities for orthonormal Laguerre polynomials*, J. Approx. Th. **144**, 2007. MR-2287374

[17] Krasovsky, I. V. Asymptotic distribution of zeros of polynomials satisfying difference equations. J. Comput. Appl. Math. 150 (2003), no. 1, 56–70. MR-1946882

[18] Marcenko, V.A., Pastur, L., *Distribution of eigenvalues in certain sets of random matrices.* Mat. Sb. (N.S.) **72** (**114**), 507–536, 1967. MR-0208649

[19] Popescu, I., *General Tridiagonal Random Matrix Models, Limiting Distributions and Fluctuations*, Prob. Theory Related Fields, **144**, 1-2, 2009. MR-2480789

[20] Sosoe, P., Wong, P., *Local semicircle law in the bulk for Gaussian β-ensembles*, J. Stat. Phys., **148**, 2, 204–232, 2012. MR-2966359

[21] Tao, T., Vu, V., *Random covariance matrices: Universality of local statistics of eigenvalues*, Ann. Prob. **40**, 3, 2012. MR-2962092

[22] Saff, E. B.; Ullman, J. L.; Varga, R. S. Incomplete polynomials: an electrostatics approach. Approximation theory, III (Proc. Conf., Univ. Texas, Austin, Tex., 1980), pp. 769–782, Academic Press, New York-London, 1980 MR-0602801

[23] Wishart, J., The generalised product moment distribution in samples from a normal multivariate population, *Biometrika* **20A**, 1928.