Journal of Approximation Theory **118**, 290–304 (2002)
doi:10.1006/jath.2002.3725

Strong Approximation of Eigenvalues of Large Dimensional Wishart Matrices by Roots of Generalized Laguerre Polynomials

Holger Dette

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
E-mail: holger.dette@ruhr-uni-bochum.de

Communicated by Arno B. J. Kuijlaars

Received January 23, 2002; accepted in revised form June 26, 2002

The purpose of this note is to establish a link between recent results on asymptotics for classical orthogonal polynomials and random matrix theory. Roughly speaking it is demonstrated that the ith eigenvalue of a Wishart matrix $W(I_n, s)$ is close to the ith zero of an appropriately scaled Laguerre polynomial, when

$$\lim_{n,s\to\infty} n/s = y \in [0, \infty).$$

As a by-product we obtain an elemantary proof of the Marčenko–Pastur and the semicircle law without relying on combinatorical arguments.

Key Words: random matrix theory; Marčenko–Pastur law; semicircle law; Laguerre polynomials; roots of orthogonal polynomials; strong approximation.

1. INTRODUCTION

The study of sample covariance matrices is important in multivariate statistics and since the pioneering work of Marčenko and Pastur [14] much effort has been devoted to this subject (see, e.g. [2, 3, 4, 12, 15, 17] among many others). In this note we present a new approach for the derivation of the asymptotic spectral distribution of a Wishart matrix $W(I_n, s)$, when the parameters n and s both converge to infinity at appropriate rates. This method relies on a close connection between the eigenvalues of the Wishart matrix and the zeros of classical orthogonal polynomials. To be precise, let $V_s \in \mathbb{R}^{n\times s}$ denote a random matrix with i.i.d. standard normally distributed entries, define

$$M_s = \frac{1}{s} V_s V_s^T \in \mathbb{R}^{n\times n} \tag{1.1}$$

0021-9045/02 $35.00

as the sample covariance matrix and let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ denote the ordered eigenvalues of the matrix M_s, where the double index has been omitted for the sake of simplicity; i.e. $\lambda_i = \lambda_i^{(n)}$. It is well known that the joint density of the eigenvalues is proportional to the function

$$\prod_{1 \leqslant i < j \leqslant n}^{n} |\lambda_i - \lambda_j| \prod_{i=1}^{n} \lambda_i^{(s-n-1)/2} e^{-\lambda_i/2}$$

and a *typical* vector of ordered eigenvalues should be close to the mode of this density. By classical results of Stieltjes (see, e.g. [16]) the above density becomes maximal for the zeros of the Laguerre polynomial. The asymptotic properties of these polynomials have been recently investigated independently from the random matrix literature in the context of approximation theory. We refer to [5, 10] for some results on strong asymptotics for Laguerre polynomials with varying coefficients and to [7, 9, 13] for recent results on the asymptotic zero distribution of these polynomials.

It is the purpose of the present paper to provide a link between the results in random matrix theory and the theory of orthogonal polynomials. To this end we derive an almost sure approximation of the eigenvalues of the Wishart matrix M_s defined in (1.1) by the zeros of appropriately scaled generalized Laguerre polynomials, when

$$\lim_{n,s \to \infty} n/s = y \in [0, \infty]. \tag{1.2}$$

This generalizes recent work of Silverstein [15], who established almost sure convergence of the smallest eigenvalue of the Wishart matrix $W(I_n, s)$, when

$$\lim_{n,s \to \infty} n/s = y \in (0, 1).$$

As a by-product we obtain a simple proof of the Marčenko–Pastur law for the empirical spectral distribution function

$$F_{M_s}(x) = \frac{1}{n} \#\{i \mid \lambda_i \leqslant x\} \tag{1.3}$$

(note that this function has already been appropriately standardized) by using recent results on weak asymptotics for classical orthogonal polynomials. Additionally, we provide a new proof of the classical semicircle law when $n/s \to 0$.

2. EIGENVALUES OF WISHART MATRICES AND ZEROS OF LAGUERRE POLYNOMIALS

Throughout this paper let for $k = 0, 1, \ldots, n\, L_k^{(\alpha_n)}(x)$ denote the kth generalized Laguerre polynomial orthogonal with respect to the weight function $x^{\alpha_n} \exp(-x) I_{(0,\infty)}(x)$. We note that the orthogonalizing measure is varying with the degree n and that we are interested in a comparison of the roots $x_1 < \cdots < x_n$ of appropriately scaled versions of the polynomial $L_n^{(\alpha_n)}(x)$ with the ordered eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ of the matrix M_s (or an appropriately scaled version) defined in (1.1). The scaling of the polynomial and the Wishart matrix depends on the limit y in (1.2) and we use the roots of the polynomial

$$L_n^{(s-n+1)}(sx)$$

in the case $y \in (0, \infty)$, the zeros of the polynomial

$$L_n^{(s)}(2\sqrt{ns}x + s + n)$$

in the case $y = 0$ and the roots of the polynomial

$$L_n^{(s-n)}(2\sqrt{ns}x + n)$$

in the case $y = \infty$ for a comparison. The scaling of the Laguerre polynomials is motivated by weak asymptotic properties of their zeros (see Theorem 2.4), which were recently obtained by Gawronski [10], Bosbach and Gawronski [5], Faldey and Gawronski [9], Dette and Studden [7], Kuijlaars and Van Assche [13]. Throughout this paper I_k denotes the $k \times k$ identity matrix. The main result of this paper is the following.

THEOREM 2.1. (a) *Let $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ denote the ordered eigenvalues of the sample covariance matrix M_s defined in* (1.1) *and $x_1 < \cdots < x_n$ denote the zeros of the Laguerre polynomial $L_n^{(s-n+1)}(sx)$. If $n, s \to \infty$, $n/s \to y \in (0, \infty)$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |\lambda_j - x_j|^2 = 0 \qquad a.s.$$

(b) *Let $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ denote the ordered eigenvalues of the sample covariance matrix*

$$N_s = \frac{1}{2\sqrt{ns}} \{ V_s V_s^T - sI_n \} \tag{2.1}$$

and $x_1 < \cdots < x_n$ denote the zeros of the Laguerre polynomial $L_n^{(s)}(2\sqrt{ns}x + s + n)$.

If $n, s \to \infty,\ n/s \to 0$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |\lambda_j - x_j|^2 = 0 \qquad a.s.$$

(c) *Let* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix*

$$P_s = \frac{1}{2\sqrt{ns}} \{V_s V_s^T - nI_n\} \tag{2.2}$$

and $-\frac{1}{2}\sqrt{n/s} = x_1 = \cdots = x_{n-s} < x_{n-s+1} < \cdots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n)}(2\sqrt{ns}x + n)$. *If* $n, s \to \infty,\ n/s \to \infty$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |\lambda_j - x_j|^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{j=n-s+1}^{n} |\lambda_j - x_j|^2 = 0 \qquad a.s.$$

Proof. (a) For a proof of part (a) we assume at first that $y \in (0,1)$, that is $s \geqslant n$ for sufficiently large n. According to [15, p. 1366] the matrix M_s is orthogonally similar to a triangular matrix $\tilde{A} = (\tilde{a}_{i,j})_{i,j=1}^n$ with entries

$$\tilde{a}_{i,i} = \frac{1}{s}(Y_{n-i+1}^2 + X_{s-i+1}^2), \qquad i = 1, \ldots, n,$$

$$\tilde{a}_{i,i+1} = \frac{1}{s} X_{s-i+1} Y_{n-i}, \qquad i = 1, \ldots, n-1,$$

$$\tilde{a}_{i+1,i} = \frac{1}{s} X_{s-i+1} Y_{n-i}, \qquad i = 1, \ldots, n-1,$$

where $Y_n^2 = 0,\ X_i^2 \sim \chi_i^2,\ Y_i^2 \sim \chi_i^2$ are independent chi-square distributed random variables ($X_i \geqslant 0$, $Y_i \geqslant 0$). Therefore it is easy to see that the matrix M_s has the same eigenvalues as the matrix $A = (a_{i,j})_{i,j=1}^n$ defined by

$$a_{i,i} = \tilde{a}_{n-i+1,n-i+1} = \frac{1}{s}(Y_i^2 + X_{s-n+i}^2), \qquad i = 1, \ldots, n,$$

$$a_{i,i+1} = \tilde{a}_{n-i,n-i+1} = \frac{1}{s} X_{s-n+i+1} Y_i, \qquad i = 1, \ldots, n-1,$$

$$a_{i+1,i} = \tilde{a}_{n-i+1,n-i} = \frac{1}{s} X_{s-n+i+1} Y_i, \qquad i = 1, \ldots, n-1.$$

Now consider the kth Laguerre polynomial

$$\hat{L}_k^{(s-n+1)}(x)$$

orthogonal with respect to the weight function $x^{s-n+1} \exp(-x) I_{(0,\infty)}(x)$ with leading coefficient 1. According to [6] we have the recursion ($\alpha_n = s - n + 1$)

$$\hat{L}_{k+1}^{(\alpha_n)}(x) = (x - \{2k + 1 + \alpha_n\})\hat{L}_k^{(\alpha_n)}(x) - k(k + \alpha_n)\hat{L}_{k-1}^{(\alpha_n)}(x) \tag{2.3}$$

with initial conditions $\hat{L}_{-1}^{(\alpha_n)}(x) = 0$, $\hat{L}_0^{(\alpha_n)}(x) = 1$. It is now straightforward to see that the zeros of the polynomial $\hat{L}_n^{(\alpha_n)}(sx)$ are precisely the eigenvalues of the triangular matrix $B = (b_{i,j})_{i,j}^n$, where

$$b_{i,i} = \frac{1}{s}(\alpha_n + 2i - 1), \qquad i = 1, \ldots, n,$$

$$b_{i,i+1} = \frac{1}{s}\sqrt{i(i+\alpha_n)}, \qquad i = 1, \ldots, n-1,$$

$$b_{i+1,i} = \frac{1}{s}\sqrt{i(i+\alpha_n)}, \qquad i = 1, \ldots, n-1.$$

This follows by factorizing $(-1)^n$, $(\frac{1}{s})^n$ in the determinant equation

$$\det(B - \lambda I) = 0,$$

and identifying recursion (2.3) for the polynomial $\hat{L}_n^{(\alpha_n)}(s\lambda)$.

Now the discussion following Lemma 2.3 in [1] yields for the distance between the eigenvalues of the matrix M_s and the zeros of the polynomial $\hat{L}_n^{(s-n+1)}(sx)$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} |\lambda_j - x_j|^2 &\leqslant \frac{1}{n}\operatorname{tr}(A-B)^2 = \frac{1}{n}\sum_{i,j=1}^{n} (a_{i,j} - b_{i,j})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} (a_{i,i} - b_{i,i})^2 + \frac{2}{n}\sum_{i=1}^{n-1} (a_{i,i+1} - b_{i,i+1})^2, \end{aligned} \tag{2.4}$$

where the first equality follows from the symmetry of the matrices A and B.

The two terms in (2.4) are estimated separately. For the first term we have with some finite constant $c > 0$ (observing $Y_n^2 = 0$)

$$\begin{aligned} c\sum_{i=1}^{n} (a_{i,i} - b_{i,i})^2 &\leqslant \sum_{i=1}^{n-1}\left(\frac{Y_i^2 - i}{s}\right)^2 + \left(\frac{n}{s}\right)^2 + \sum_{i=1}^{n}\left(\frac{X_{s-n+i}^2 - s + n - i}{s}\right)^2 \\ &\leqslant 2nM_n^2 + \left(\frac{n}{s}\right)^2, \end{aligned}$$

where the random variable M_n is defined by

$$M_n = \max\left\{\max_{1 \leqslant i \leqslant n-1}\left|\frac{Y_i^2 - i}{s}\right|, \max_{s-(n-1) \leqslant i \leqslant s}\left|\frac{X_i^2 - i}{s}\right|\right\}. \tag{2.5}$$

From [15, p. 1367] it follows that $M_n \to 0$ a.s. and we obtain

$$\frac{1}{n}\sum_{i=1}^{n} (a_{i,i} - b_{i,i})^2 \to 0 \qquad \text{a.s.} \tag{2.6}$$

For the remaining term in (2.4) we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n-1}(a_{i,i+1}-b_{i,i+1})^2 &= \frac{1}{n}\sum_{i=1}^{n-1}\left|\frac{X_{s-n+i+1}Y_i}{s}-\frac{\sqrt{i(i+\alpha_n)}}{s}\right|^2\\ &\leqslant \frac{1}{n}\sum_{i=1}^{n-1}\left\{\left|\frac{X^2_{s-n+i+1}-(s-n+i+1)}{s}\right|^{1/2}\left|\frac{Y_i^2-i}{s}\right|^{1/2}\right.\\ &\quad+\sqrt{\frac{i+s-n+1}{s}}\left|\frac{Y_i^2-i}{s}\right|^{1/2}\\ &\quad\left.+\sqrt{\frac{i}{s}}\left|\frac{X^2_{s-n+i+1}-(s-n+i+1)}{s}\right|^{1/2}\right\}^2\\ &\leqslant (M_n+2\sqrt{M_n})^2\to 0 \qquad \text{a.s.,}\end{aligned}$$

where the random variable M_n is defined in (2.5) and we have used the inequality

$$\begin{aligned}|\underline{a}\,\underline{b}-ab| \leqslant\ & |\underline{a}^2-a^2|^{1/2}|\underline{b}^2-b^2|^{1/2}\\ &+|b||\underline{a}^2-a^2|^{1/2}+|a||\underline{b}^2-b^2|^{1/2}\end{aligned}$$

for nonnegative $\underline{a}$, $\underline{b}, a, b$ (see [15]). Observing (2.4) assertion (a) of Theorem 2.1 follows in the case $y \in (0,1)$.

In the case $y > 1$ (which means $n > s$ for sufficiently large n) the result is established by interchanging the roles of s and n and from a representation for generalized Laguerre polynomials with negative parameter. To be precise, we note that in this case the matrix M_s is orthogonally similar to an $n \times n$ matrix A with principal $s \times s$ block containing the (s-dimensional) rows

$$\frac{1}{s}(X_n^2+Y_{s-1}^2, Y_{s-1}X_{n-1}, 0, \ldots, 0),$$

$$\frac{1}{s}(Y_{s-i+1}X_{n-i+1}, X^2_{n-i+1}+Y^2_{s-i}, Y_{s-i}X_{n-i}, 0, \ldots, 0) \qquad (i=2,\ldots,s-1)$$

and

$$\frac{1}{s}(0,\ldots,0, Y_1X_{n-s+1}, X^2_{n-s+1}),$$

where all other entries in the matrix A are 0 and the meaning of the random variables X_i, Y_i, X_i^2, Y_i^2 is the same as in the previous paragraph. Observing the identity

$$L_n^{(-k)}(x) = (-x)^k \frac{(n-k)!}{n!} L_{n-k}^{(k)}(x) \tag{2.7}$$

(see [16, Sect. 5.2)) the assertion now follows by similar arguments as given for the case $y \in (0, 1)$.

The remaining case $y = 1$ is proved by considering two subsequences corresponding to the cases $s \geqslant n$ and $s < n$, respectively.

(b) By the same argument as given in the proof of part (a) the eigenvalues of the matrix N_s defined in (2.1) are obtained as the eigenvalues of the tridiagonal matrix A defined by

$$a_{i,i} = \frac{1}{2\sqrt{sn}} (Y_i^2 + X_{s-n+i}^2 - s),$$

$$a_{i,i+1} = a_{i+1,i} = \frac{1}{2\sqrt{sn}} X_{s-n+i+1} Y_i.$$

Now consider the Laguerre polynomials with leading coefficient 1 and parameter $\alpha_n = s$ and define polynomials

$$p_k(x) = \hat{L}_k^{(s)}(2\sqrt{ns}x + s + n).$$

The zeros of the polynomial $p_n(x)$ are given by the eigenvalues of the matrix B defined by

$$b_{i,i} = \frac{1}{2\sqrt{ns}} (\alpha_n + 2i - 1 - n - s) = \frac{1}{2\sqrt{ns}} (2i - 1 - n), \qquad i = 1, \ldots, n,$$

$$b_{i,i+1} = b_{i+1,i} = \frac{1}{2\sqrt{ns}} \sqrt{i(i + \alpha_n)} = \frac{1}{2\sqrt{ns}} \sqrt{i(i + s)}, \qquad i = 1, \ldots, n-1,$$

where $\alpha_n = s$. The assertion now follows by similar arguments as given in the proof of part (a).

(c) The asymptotic properties in the case $y = \infty$ follow from a combination of the arguments given in the proof of part (a) for the case $y > 1$ and the proof of part (b). The matrix P_s defined in (2.2) is orthogonally similar to an $n \times n$ matrix A with principal $s \times s$ block containing the (s-dimensional) rows

$$\frac{1}{2\sqrt{ns}}\,(X_n^2 + Y_{s-1}^2 - n,\; Y_{s-1}X_{n-1}, 0, \ldots, 0),$$

$$\frac{1}{2\sqrt{ns}}\,(Y_{s-i+1}X_{n-i+1}, X_{n-i+1}^2 + Y_{s-i}^2 - n,\; Y_{s-i}X_{n-i}, 0, \ldots, 0), \qquad (i = 2, \ldots, s-1)$$

and

$$\frac{1}{2\sqrt{ns}}\,(0, \ldots, 0,\; Y_1X_{n-s+1}, X_{n-s+1}^2 - n),$$

where all other entries in the matrix A are 0 and the meaning of the random variables X_i, Y_i, X_i^2, Y_i^2 is the same as in the proof of part (a). From (2.7) we have for some constant $c \neq 0$

$$L_n^{(s-n)}(2\sqrt{ns}x + n) = c(2\sqrt{ns}x + n)^{n-s} L_s^{(n-s)}(2\sqrt{ns}x + n),$$

where the s positive zeros of the polynomial on the right-hand side are obtained as the eigenvalues of the tridiagonal matrix B with elements

$$b_{i,i} = \frac{1}{2\sqrt{ns}}\,(2i - 1 - s), \qquad i = 1, \ldots, s,$$

$$b_{i,i+1} = b_{i+1,i} = \frac{1}{2\sqrt{ns}}\sqrt{i(i + n - s)}, \qquad i = 1, \ldots, s-1.$$

The assertion now follows by similar arguments as given in the proof of part (a). ∎

The following result is an immediate consequence of Theorem 2.1 and recent results on the location of the zeros of classical orthogonal polynomials.

COROLLARY 2.2. (a) *Let* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* M_s *defined in* (1.1) *and* $x_1 < \cdots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n+1)}(sx)$. *If* $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty,\ n/s \to y \in (0, \infty)$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} d_{\lfloor nt \rfloor} = 0 \qquad a.s.$$

for all $t \in (0, 1)$. *In particular we obtain for the smallest and largest eigenvalue of the matrix* M_s *and for the smallest and largest zero of the* *polynomial* $L_n^{(s-n+1)}(sx)$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,1]}} x_1 = \lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,1]}} \lambda_1 = (1-\sqrt{y})^2 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} x_n = \lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} \lambda_n = (1+\sqrt{y})^2 \qquad a.s.$$

and in the case $y \geqslant 1$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in[1,\infty)}} x_{n-s+1} = \lim_{\substack{n,s\to\infty \\ n/s\to y\in[1,\infty)}} \lambda_{n-s+1} = (1-\sqrt{y})^2 \qquad a.s.$$

(b) *Let* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* N_s *defined in* (2.1) *and* $x_1 < \cdots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s)}(2\sqrt{ns}x+s+n)$. *If* $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty,\ n/s \to 0$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} d_{\lfloor nt \rfloor} = 0 \qquad a.s.$$

for all $t \in (0,1)$. *In particular we obtain for the largest and smallest eigenvalue of the matrix* N_s *and for the smallest and largest zero of the polynomial* $L_n^{(s)}(2\sqrt{ns}x+s+n)$

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} x_1 = \lim_{\substack{n,s\to\infty \\ n/s\to 0}} \lambda_1 = 1 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} x_n = \lim_{\substack{n,s\to\infty \\ n/s\to 0}} \lambda_n = -1 \qquad a.s.$$

(c) *Let* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* P_s *defined in* (2.2) *and* $x_1 < \cdots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n)}(2\sqrt{ns}x+n)$. *If* $0 = d_1 = \cdots = d_{n-s} \leqslant d_{n-s+1} \leqslant d_2 \leqslant \cdots \leqslant d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty,\ n/s \to \infty$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to \infty}} d_{\lfloor nt \rfloor} = 0 \qquad a.s.$$

for all $t \in (0,1)$. *In particular we obtain for the largest and* $(n-s+1)$*th smallest eigenvalue of the matrix* P_s *and for the* $(n-s+1)$*th smallest and* *largest zero of the polynomial* $L_n^{(s-n)}(2\sqrt{ns}x+n)$

$$\lim_{\substack{n,s\to\infty\\ n/s\to\infty}} x_{n-s+1} = \lim_{\substack{n,s\to\infty\\ n/s\to\infty}} \lambda_{n-s+1} = -1 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty\\ n/s\to\infty}} x_n = \lim_{\substack{n,s\to\infty\\ n/s\to\infty}} \lambda_n = 1 \qquad a.s.$$

Proof. The first part is an immediate consequence of Theorem 2.1. The assertion regarding the largest and smallest eigenvalue follows similarly to the proof of Theorem 2.1 by an application of the Theorem of Geršgorin [11] (see [15]). The results for the largest and smallest zero of the Laguerre polynomial can be obtained from Theorem 4.4 in [7] and formula (2.7). ∎

The asymptotic properties of the largest and smallest eigenvalue in part (a) of Corollary 2.2 were already observed by Silverstein [15], but we did not find the result for sample covariance matrices for the case $n/s \to 0$ or $n/s \to \infty$ in the literature (for a proof of the analogue for $n/s \to 0$ in the case of Wigner matrices see [3]). The following example illustrates the quality of approximation in Corollary 2.2.

EXAMPLE 2.3. Consider the case $s = 10n$ and note that for finite samples the limits $n/s \to y \in (0, \infty)$ and $n/s \to 0$ cannot be distinguished. Therefore both approximations of parts (a) and (b) in Theorem 2.1 could be used in principle. For the sake of brevity we use only case (b). Table I shows the

TABLE I
The 10 smallest zeros of the scaled Laguerre polynomials $p_n(x)$ defined in (2.8) and the n smallest eigenvalues of the standardized Wishart matrix N_{10n} defined in (2.9) for various values of n

n	10		15		20	
	λ_j	x_j	λ_j	x_j	λ_j	x_j
	−0.74099	−0.72672	−0.76857	−0.76658	−0.78275	−0.78866
	−0.57733	−0.57065	−0.64970	−0.65404	−0.68759	−0.69867
	−0.43048	−0.42409	−0.54791	−0.55201	−0.60781	−0.61862
	−0.28494	−0.27741	−0.45075	−0.45334	−0.53318	−0.54254
	−0.13478	−0.12609	−0.35486	−0.35519	−0.46089	−0.46816
	0.02480	0.03340	−0.25757	−0.25592	−0.38926	−0.39427
	0.19766	0.20499	−0.15823	−0.15433	−0.31717	−0.32007
	0.39190	0.39425	−0.05522	−0.04934	−0.24446	−0.24499
	0.62150	0.61121	0.05261	0.06015	−0.17025	−0.16852
	0.93142	0.88111	0.16661	0.17546	−0.09418	−0.09021

zeros $x_1, \dots, x_n$ of the Laguerre polynomial

$$p_n(x) = L_n^{(10n)}(2n\sqrt{10}x + 11n) \tag{2.8}$$

and the eigenvalues $\lambda_1, \dots, \lambda_n$ of the standardized matrix

$$N_{10n} = \frac{1}{2n\sqrt{10}}(V_{10n}V_{10n}^T - 10nI_n) \tag{2.9}$$

for $n = 10, 15, 20$. These eigenvalues have been obtained by simulations based on 100.000 runs. For the sake of brevity only the ten smallest eigenvalues and zeros are displayed.

We will use Theorem 2.1 for an alternative proof of the famous Marčenko–Pastur and semicircle law in the normal case using recent results for the asymptotic zero distribution of classical orthogonal polynomials. Conversely the arguments given in this paper show that the Marčenko–Pastur and semicircle law could also be used to provide an alternative proof for the asymptotic zero distribution of the Laguerre polynomials with varying *integer-valued* parameters. For the sake of completeness we recall a result on the asymptotic zero distribution for the zeros of the Laguerre polynomials with varying (not necessarily integer valued) parameters. A proof can be found in [7] (see also [8, 9] or [13]). For a real sequence $(\alpha_n)_{n\in\mathbb{N}}$ with elements > -1 let

$$N^{(\alpha_n)}(\xi) := \#\{x | L_n^{(\alpha_n)}(x) = 0, x \leqslant \xi\} \tag{2.10}$$

denote the number of zeros of the generalized Laguerre polynomial $L_n^{(\alpha_n)}(x)$ less or equal than ξ, then we have the following result.

Theorem 2.4 (Dette and Studden [7]).
(a) *If* $\lim_{n\to\infty} \frac{\alpha_n}{n} = a \geqslant 0$, *then*

$$\lim_{n\to\infty} \frac{1}{n} N^{(\alpha_n)}(n\xi) = \frac{1}{2\pi}\int_{r_1}^{\xi} \frac{\sqrt{(r_2 - x)(x - r_1)}}{x}\,dx \qquad \textit{for all } \xi \in [r_1, r_2],$$

where $r_{1,2} = 2 + a \pm 2\sqrt{1+a}$
(b) *If* $\lim_{n\to\infty} \frac{\alpha_n}{n} = \infty$, *then*

$$\lim_{n\to\infty} \frac{1}{n} N^{(\alpha_n)}(\sqrt{n\alpha_n}\xi + \alpha_n) = \frac{1}{2\pi}\int_{-2}^{\xi} \sqrt{4 - x^2}\,dx \qquad \textit{for all } |\xi| \leqslant 2.$$

Theorem 2.5. (Marčenko–Pastur and Semicircle Law).

(a) *If* $n \to \infty$, $n/s \to y \in (0, \infty)$ *and* F_{M_s} *denotes the empirical spectral distribution function of the matrix* M_s *defined in* (1.1), *then for all* $\xi \in \mathbb{R}$

$$F_{M_s}(\xi) \to F_M(\xi) \qquad a.s., \tag{2.11}$$

where the distribution function F_M *has density*

$$f_M(x) := \frac{1}{2\pi y} \frac{\sqrt{(b-x)(x-a)}}{x} I_{[a,b]}(x),$$

the quantities a *and* b *are given by* $a = (1-\sqrt{y})^2$, $b = (1+\sqrt{y})^2$, *respectively, and there is an additional jump of size* $1 - 1/y$ *in the case* $y > 1$.

(b) *If* $n/s \to 0$ *and* F_{N_s} *denotes the empirical spectral distribution function of the matrix* (2.1), *then we have for any* $x \in [-1, 1]$

$$F_{N_s}(x) \to F_N(x) := \frac{2}{\pi} \int_{-1}^{x} \sqrt{1-t^2}\, dt \qquad a.s. \tag{2.12}$$

$(F_{N_s}(x) \to 1$ *if* $x > 1$, $F_{N_s}(x) \to 0$ *if* $x < 1)$.

(c) *If* $n, s \to \infty$, $n/s \to \infty$ *and* F_{P_s} *denotes the empirical distribution function of the s largest eigenvalues of the matrix* P_s *defined by* (2.2), *then we have for any* $x \in [-1, 1]$

$$F_{P_s}(x) \to F_N(x) = \frac{2}{\pi} \int_{-1}^{x} \sqrt{1-t^2}\, dt \qquad a.s. \tag{2.13}$$

$(F_{P_s}(x) \to 1$ *if* $x > 1, F_{P_s}(x) \to 0$ *if* $x < 1)$.

Proof. (a) Consider at first the case (a) with $y \in (0, 1]$. From [1] and Theorem 2.1 it follows for the Levy distance L between the distribution functions F_{M_s} and F_B that

$$L^3(F_{M_s}, F_B) \leqslant \frac{1}{n} \sum_{i=1}^{n} |\lambda_j - x_j|^2 \to 0 \qquad \text{a.s.},$$

where

$$F_B(\xi) = \frac{1}{n} \#\{x \mid \hat{L}_n^{(\alpha_n)}(sx) = 0, x \leqslant \xi\}$$

denotes the empirical distribution function of the zeros of the Laguerre polynomial $L_n^{(\alpha_n)}(sx)$ with parameter $\alpha_n = s - n + 1$. From the first part of

Theorem 2.4 we therefore have for any $\xi \in [r_1, r_2]$

$$F_B(\xi) = \frac{1}{n}\#\{x \mid \hat{L}_n^{(\alpha_n)}\left(n\frac{s}{n}x\right) = 0, x \leqslant \xi\}$$
$$= \frac{1}{n}N^{(\alpha_n)}\left(n\frac{s}{n}\xi\right) \underset{\substack{n\to\infty \\ \frac{n}{s}\to y}}{\longrightarrow} \frac{1}{2\pi}\int_{r_1}^{\xi/y} \frac{\sqrt{(r_2 - x)(x - r_1)}}{x}\,dx, \qquad (2.14)$$

where $r_{1,2} = (1 \pm \frac{1}{\sqrt{y}})^2$. Substitution and differentiation yields for the density of the limiting distribution

$$\frac{1}{2\pi y}\frac{\sqrt{(b-x)(x-a)}}{x} I_{[a,b]}(x),$$

where

$$a = (1 - \sqrt{y})^2,$$

$$b = (1 + \sqrt{y})^2.$$

The argument for the case $y > 1$ follows exactly in the same way using at first identity (2.7).

(b) Again we obtain from Theorem 2.1

$$L^3(F_{N_s}, F_B) \to 0 \qquad \text{a.s.},$$

where F_B denotes the empirical distribution function of the roots of the polynomial $L_n^{(\alpha_n)}(2\sqrt{n\alpha_n}x + s + n)$ with $\alpha_n = s$, that is

$$F_B(\xi) = \frac{1}{n}\#\{x \mid L_n^{(\alpha_n)}(2\sqrt{n\alpha_n}x + s + n) = 0, x \leqslant \xi\}$$
$$= \frac{1}{n}N^{(\alpha_n)}(2\sqrt{n\alpha_n}\xi + s + n).$$

Observing $\frac{n}{\sqrt{n\alpha_n}} = \sqrt{\frac{n}{s}} = o(1)$ and Example 2.7 in [7] the second part of Theorem 2.4 now gives

$$\lim_{n\to\infty} F_B(\xi) = \lim_{n\to\infty} \frac{1}{n}N^{(\alpha_n)}(2\sqrt{n\alpha_n}\xi + \alpha_n)$$
$$= \frac{1}{2\pi}\int_{-2}^{2\xi}\sqrt{4 - x^2}\,dx = \frac{2}{\pi}\int_{-1}^{\xi}\sqrt{1 - t^2}\,dt$$

whenever $|\xi| \leqslant 1$, which proves the assertion of Theorem 2.5(b).

(c) This is proved in the same way using identity (2.7). ∎

Remark 2.6. We finally note that most of the results on Wishart matrices hold under more general assumptions (existing fourth moments and not necessarily normally distributed random variables). A direct extension of the presented equivalence seems to be difficult, because our proofs rely heavily on Silverstein's [15] work, which definitively requires the assumption of a normal distribution. Research in this direction is planned for the future.

ACKNOWLEDGMENTS

The author would like to thank I. Gottschlich, who typed most parts of the paper with considerable technical expertise and I. Johnstone for some help with the references. He is also grateful to an unknown referee for some comments on an earlier version of this paper. The work of the author was supported by the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen).

REFERENCES

1. Z. D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a review, *Statist. Sinica* **9** (1999), 611–677.
2. Z. D. Bai and Y. Q. Yin, Convergence to the semicircle law, *Ann. Probab.* **16** (1988), 863–875.
3. Z. D. Bai and Y. Q. Yin, Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix, *Ann. Probab.* **16** (1988), 1729–1741.
4. Z. D. Bai and Y. Q. Yin, Limit of the smallest eigenvalue of a large dimensional covariance matrix, *Ann. Probab.* **21** (1993), 1275–1294.
5. C. Bosbach and W. Gawronski, Strong asymptotics for Laguerre polynomials with varying weights, *J. Comput. Appl. Math.* **99** (1998), 77–89.
6. T. Chihara, "An Introduction to Orthogonal Polynomials," Gordon and Breach, London, 1978.
7. H. Dette and W. J. Studden, Some new asymptotic properties for the zeros of Jacobi, Laguerre and Hermite polynomials, *Constr. Approx.* **11** (1995), 227–238.
8. H. Dette and W. K. Wong, Recurrence moment formulas for *D*-optimal designs, *Scand. J. Statist.* **22** (1995), 505–512.
9. J. Faldey and W. Gawronski, On the limit distribution of the zeros of Jonquiere polynomials and generalized classical orthogonal polynomials, *J. Approx. Theory* **81** (1995), 231–249.
10. W. Gawronski, Strong asymptotics and the asymptotic zero distributions of Laguerre polynomials $L_n^{(an+\alpha)}$ and Hermite polynomials $H_n^{(an+\alpha)}$, *Analysis* **13** (1993), 29–67.
11. S. A. Geršgorin, Über die Abgrenzung der Eigenwerte einer Matrix, *Ivz. Akad. Nauk. SSSR Ser. Fiz.-Mat.* **6** (1931), 749–754.
12. I. M. Johnstone, On the distribution of the largest principal component, *Ann. Statist.*, **29** (2001), 295–327.
13. A. B. J. Kuijlaars and W. Van Assche, The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients, *J. Approx. Theory* **99** (1999), 167–197.

14. V. A. Marčenko and L. A. Pastur, Distributions of eigenvalues of some sets of random matrices, *Math. USSR-Sb.* **1** (1967), 507–536.
15. J. Silverstein, The smallest eigenvalue of a large dimensional Wishart matrix, *Ann. Probab.* **13** (1985), 1364–1368.
16. G. Szegö, ‘‘Orthogonal Polynomials,’’ Amer. Math. Soc., Providence, RI, 1959.
17. E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. Math.* **62** (1955), 548–564.