Random Matrix Theory and its Innovative Applications*

Alan Edelman and Yuyang Wang

Abstract Recently more and more disciplines of science and engineering have found Random Matrix Theory valuable. Some disciplines use the limiting densities to indicate the cutoff between "noise" and "signal." Other disciplines are finding eigenvalue repulsions a compelling model of reality. This survey introduces both the theory behind these applications and MATLAB experiments allowing a reader immediate access to the ideas.

1 Random Matrix Theory in the Press

Since the beginning of the 20th century, Random matrix theory (RMT) has been finding applications in number theory, quantum mechanics, condensed matter physics, wireless communications, etc., see [16, 15, 12, 7]. Recently more and more disciplines of science and engineering have found RMT valuable. New applications in RMT are being found every day, some of them surprising and innovative when compared with the older applications.

For newcomers to the field, it may be reassuring to know that very little specialized knowledge of random matrix theory is required for applications, and therefore the "learning curve" to become a user is not at all steep. Two methodologies are worth highlighting.

Alan Edelman
Department of Mathematics, MIT, e-mail: edelman@math.mit.edu

Yuyang Wang
Department of Computer Science, Tufts University e-mail: ywang02@cs.tufts.edu

* The first author was supported in part by DMS 1035400 and DMS 1016125. Note to our readers: This survey paper is in large part a precursor to a book on Random Matrix Theory that will be forthcoming. We reserve the right to reuse materials in the book.

Fig. 1 New Scientist cover story entitled *Entering the matrix: the simple idea at the root of reality*. Quoting Raj Rao Nadakuditi: "It really does feel like the ideas of random matrix theory are somehow buried deep in the heart of nature."

1. **Distinguishing "signal" from "noise":** Generate a matrix of data specific to your application (e.g. a correlation matrix or a sample covariance matrix) and perhaps normalize the data to have mean 0 and variance 1. Compare the answer to the known example of singular values or eigenvalues of random matrices. (Usually it is enough to know the quarter circle law, the semi-circle law, and the Marčenko-Pastur Laws, see Section 2 for details.)

   1a) **"No correlation"**: If the answer obtained is similar enough to one of the known laws in RMT, one might declare the data to be all noise. In one example (Figure 7), Popoff et al. [20] use the fact that the distribution of the singular values of the transmission matrix follows a quarter circle law (see Section 2) to show that the matrix elements are not significantly correlated, thereby justifying the fact that their experimental procedure does not introduce spurious correlations.

   1b) **"Correlations"**: If there are singular values/eigenvalues larger than those predicted by RMT, one considers those as indicative of "signal" or correlations in the data. Most recently, Arup Chakraborty, a Chemistry and Chemical Engineering professor at MIT and his colleagues used RMT (Figure 3) to find sectors of HIV that rarely undergo multiple mutations [8]. The mathematical procedure is relatively easy to explain and apply. It is known that if we take a matrix `A=randn(m,n)` (in MATLAB notation), which is an $m \times n$ random matrix with

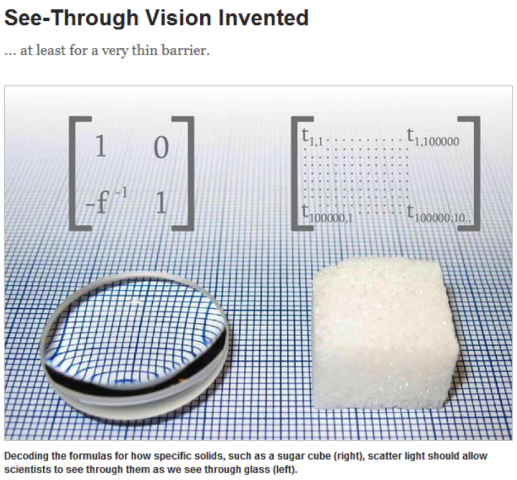

Fig. 2 Comparing the singular values of a transmission matrix to that of a random matrix suggests that there are no spurious correlations.

independent and identically distributed (i.i.d.) standard normal, then the eigenvalues of the Wishart matrix A^TA/m in the limit as $m/n = r$ and $m,n \to \infty$ are almost surely in the interval $[(1-\sqrt{r})^2,(1+\sqrt{r})^2]$. Thus if we have a correlation matrix with eigenvalues larger than $(1+\sqrt{r})^2$ we consider these as signal rather than noise. Theoretical understanding of the "signal" eigenvalues may be found in [19].

Fig. 3 Random matrix techniques in a recent study featured in the Wall Street Journal

2. **Spacing Distributions**: Generate a statistic that one considers likely to be represented by the spacings, or the largest eigenvalue, or the smallest eigenvalue of a random matrix. These might be recognized respectively by a statistic that is feeling repulsions from two sides, that is feeling pushed outward from behind, or pushed inward towards a "hard edge."

2a) **Repulsion from two sides**: The repulsion from two edges is given by what is known as the bulk distribution spacings and is described in [21].

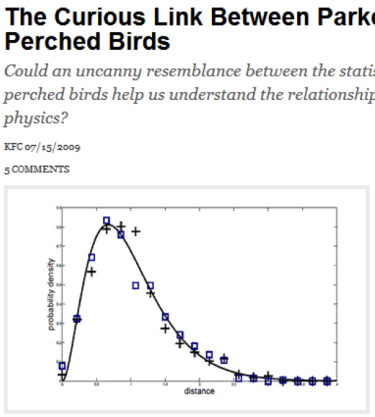

Fig. 4 Random matrix theory spacings shows itself in the gaps between parked cars, the distances between perched birds (illustrated above), the distribution peaks that neutron scatter off heavy nuclei, etc.

2b) **Repulsion from "behind" with no barrier**: The repulsion from behind is given by the Tracy-Widom distributions ($\beta = 2$ is the most common application, but $\beta = 1$ also shows up). In a remarkable first of a kind experiment [23], Kazumasa Takeuchi and Masaki Sano have measured the interface in turbulent liquid crystal growth. Two kinds of experiments are reported, flat and curved interfaces. Careful measurements of the first four moments of the random radii match those of the largest eigenvalues of real and complex matrices respectively. These are the Tracy-Widom laws. (See Section 2 for the complex Tracy-Widom law). The use of the third and fourth moments, in the form of skewness and kurtosis, indicate that the Tracy-Widom law really appears to be applying in practice. In other results where this same law is conjectured, statistics sometimes seem indistinguishable from the normal distribution.

2c) **Repulsion with barrier**: The repulsion towards the hard edge has been less commonly applied, but corresponds to a "Bessel kernel."

The rest of this paper is organized as follows. In Section 2, we introduce two classical random matrix ensembles, namely, Hermite ensembles and Laguerre ensembles. Further, we describe the limiting eigenvalue densities. Section 3 starts with a numerical question of how to calculate the eigenvalues of a random matrix efficiently. Then, we discuss theoretical implications of the computational trick.

2 Famous Laws in RMT with MATLAB experiments

We introduce the classic random matrix ensembles and then we will provide four famous laws in RMT with corresponding MATLAB experiments. Notice that although measure-theoretical probability is not required to understand and appreciate the beauty of RMT in this paper, the extension of probabilistic measure-theoretical tools to matrices is nontrivial, we refer interested readers to [6, 2].

While we expect our readers to be familiar with real and complex matrices, it is reasonable to consider quaternion matrices as well. Let us start with the Gaussian random matrices $G_1(m,n)$ (`G1 = randn(m, n)`), which is an $m \times n$ matrix with i.i.d. standard real random normals. In general, we use the parameter β to denote the number of standard real normals and thus $\beta = 1,2,4$ correspond to real, complex and quaternion respectively. $G_\beta(m,n)$ can be generated by the MATLAB command shown in Table 1. Notice that since quaternions do not exist in MATLAB they are "faked" using 2×2 complex matrices.

Table 1 Generating the Gaussian random matrix $G_\beta(m,n)$ in MATLAB

β	MATLAB command
1	`G = randn(m, n)`
2	`G = randn(m, n) + i*randn(m, n)`
4	`X = randn(m, n) + i*randn(m, n);`
	`Y = randn(m, n) + i*randn(m, n); G = [X Y; -conj(Y) conj(X)]`

If A is an $m \times n$ Gaussian random matrix $G_\beta(m,n)$ then its joint element density is given by

$$\frac{1}{(2\pi)^{\beta mn/2}} \exp\left(-\frac{1}{2}\|A\|_F^2\right), \tag{1}$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix.

The most important property of G_β, be it real, complex, or quaternion, is its *orthogonal invariance*. This makes the distribution impervious to multiplication by an orthogonal (unitary, symplectic) matrix, provided that the two are independent. This can be inferred from the joint element density in (1) since its Frobenius norm, $\|A\|_F$, is unchanged when A is multiplied by an orthogonal (unitary, symplectic) matrix. The orthogonal invariance implies that no test can be devised that would differentiate between Q_1A, A, and AQ_2, where Q_1 and Q_2 are non-random orthogonal and A is Gaussian. Readers will later see that this simple property leads to wonderful results both in practice and in theory.

The most well-studied random matrices have names such as Gaussian, Wishart, MONOVA, and circular. We prefer Hermite, Laguerre, Jacobi, and perhaps Fourier. In a sense, they are to random matrix theory as Poisson's equation is to numerical methods. Of course, we are thinking in the sense of the problems that are well-tested, well-analyzed, and well-studied because of nice fundamental analytic properties.

These matrices play a prominent role because of their deep mathematical structure. There are four channels of structure lurking underneath numeric analysis, graph theory, multivariate statistics [17] and operator algebras [18]. In this paper, we will focus on the Hermite and Laguerre ensembles, which is summarized in Table 2. The other random matrix ensembles are discussed in details in [10].

Table 2 Hermite and Laguerre ensembles.

Ensemble	Matrices	Weight function	Equilibrium measure	Numeric	MATLAB
Hermite	Wigner	$e^{-x^2/2}$	semi-circle	`eig`	`g = G(n,n); H=(g+g')/2`
Laguerre	Wishart	$x^{\nu/2-1}e^{-x/2}$	Marcenko-Pastur	`svd`	`g = G(m,n); L=(g'*g)/m;`

2.1 The most famous Semi-circle law

In [24], Wigner originally showed that the limiting eigenvalue distribution of simple random symmetric $n \times n$ matrices $X = (A + A^T)/2$ where $A = G_1(n,n)$, follow a semi-circle distribution, which is given by

$$p(x) = \frac{1}{2\pi}\sqrt{4-x^2}. \tag{2}$$

When properly normalized, the curve looks like a semi-circle of radius 2. This distribution depicts the histogram of the n eigenvalues of a symmetric random $n \times n$ matrix obtained by symmetrizing a matrix of random normals. X constructed in this way is called the β-**Hermite ensemble** or Gaussian ensemble, more specifically *Gaussian orthogonal ensemble (GOE)* ($\beta = 1$), *Gaussian unitary ensemble (GUE)* ($\beta = 2$) and *Gaussian symplectic ensemble (GSE)* ($\beta = 4$). Code 1 histograms the random eigenvalues and plots the semi-circle. The mathematical theorem requires only one matrix $t = 1$ and $n \to \infty$, though the computer is happier with much smaller values for n.

Theorem 1. *(Wigner 1955) Let X_n be a sequence of random symmetric $n \times n$ matrices $(n = 1,2,\ldots)$, satisfying*

1. *Independent elements (up to matrix symmetry): The elements x_{ij} for $i \leq j$ of each X_n are independent random variables*[2].
2. *Zero Mean: The elements x_{ij} of each X_n satisfy $\mathbb{E}(x_{ij}) = 0$.*
3. *Unit off diagonal variance (for normalization): The elements x_{ij} of each X_n satisfy $\mathbb{E}(x_{ij}^2) = 1$.*

[2] Strictly speaking, the random variables should be written $x_{ij}(n)$.

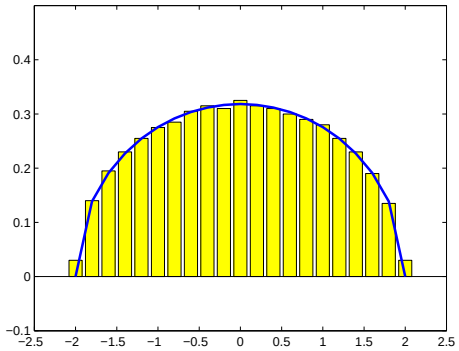

Fig. 5 Semi-circle law with one 1000×1000 matrix. Plotted is the histogram of the 1000 eigenvalues and the semi-circle on the blue line.

4. *Bounded moments: For $k = 1,2,\ldots$, There is some bound*[3] B_k, *independent of n, such that for all $m \leq k$, $\mathbb{E}(|x_{ij}|^m) \leq B_k$.*

Under these assumptions, the distribution of the eigenvalues of such $\frac{1}{\sqrt{n}}X_n$ asymptotically approaches a semi-circle distribution (Equation 2) in the following sense. As $n \to \infty$, $\mathbb{E}[\lambda^k]$ matches the moments of the semicircle distribution for a randomly chosen λ from n eigenvalues.

Wigner's original proof is combinatorial; he showed that the significant terms in $\mathbb{E}[\mathbf{Tr}(X^{2k})]$ count the ordered trees on $k+1$ vertices. This count is well-known today as the *Catalan numbers*, where the nth Catalan number is

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{3}$$

The moments of a semi-circle distribution are the very same Catalan numbers.

Furthermore, the eigenvalue density of the β-**Hermite ensemble** ($\beta = 1,2,4$) almost surely converges to the semi-circle distribution. The proof can be found in chapter 2 of [5], which also discusses the state of the art knowledge concerning the assumption of bounded moments. Roughly speaking, if the independent elements are i.i.d., the finite moment condition can be dropped. If not, there is a weaker condition by Girko that is claimed to be an if and only if.

The semi-circle law acts like a central limit theorem for (infinitely) large symmetric random matrices. If we average a large number of independent and identically distributed random matrices, the classical central limit theorem says that the elements become Gaussians.

[3] If the $x_{ij}(n)$ are identically distributed, a very common assumption, then it is sufficient to require finite moments.

Another important problem is the convergence rate of the (empirical) eigenvalue density, which was answered by Bai in [3, 4]. We refer interested readers to chapter 8 of [5].

Code 1 Semicircle Law (Random symmetric matrix eigenvalues)

```
%Experiment:   Gaussian Random Symmetric Eigenvalues
%Plot:         Histogram of the eigenvalues
%Theory:       Semicircle as n->infinity
%% Parameters
n=1000;        %matrix size
t=1;           %trials
v=[];          %eigenvalue samples
dx=.2;         %binsize
%% Experiment
for i=1:t,
    a=randn(n);     % random nxn matrix
    s=(a+a')/2;     % symmetrized matrix
    v=[v;eig(s)];   % eigenvalues
end
v=v/sqrt(n/2);      % normalized eigenvalues
%% Plot
[count,x]=hist(v,-2:dx:2);
cla reset
bar(x, count/(t*n*dx),'y');
hold on;
%% Theory
plot(x,sqrt(4-x.^2)/(2*pi),'LineWidth',2)
axis([-2.5 2.5 -.1 .5]);
```

On the other hand, in the finite case where n is given, one may wonder how are the eigenvalues of an $n \times n$ symmetric random matrix A distributed? Fortunately, for the Hermite ensemble, the answer is known explicitly and the density is called the *level density* in the Physics literature [15]. It is worth mentioning that the joint element density of an $n \times n$ matrix A_β from the Hermite ensemble is [9]

$$\frac{1}{2^{n/2}} \frac{1}{\pi^{n/2+n(n-1)\beta/4}} \exp\left(-\frac{1}{2}\|A\|_F^2\right), \tag{4}$$

and the joint eigenvalue probability density function is

$$f_\beta(\lambda_1, \cdots, \lambda_n) = c_H^\beta \prod_{i<j} |\lambda_i - \lambda_j|^\beta \exp\left(-\sum_{i=1}^n \frac{\lambda_i^2}{2}\right), \tag{5}$$

with

$$c_H^\beta = (2\pi)^{-n/2} \prod_{j=1}^n \frac{\Gamma(1+\frac{\beta}{2})}{\Gamma(1+\frac{\beta}{2}j)}.$$

The level density ρ_n^A for an $n \times n$ ensemble A with real eigenvalues is the distribution of a random eigenvalue chosen from the ensemble. More precisely, the level density can be written in terms of the marginalization of the joint eigenvalue density. For example, in the Hermite ensemble case,

$$\rho_{n,\beta}^A(\lambda_1) = \int_{\mathbb{R}^{n-1}} f_\beta(\lambda_1, \cdots, \lambda_n) \mathrm{d}\lambda_2 \cdots \mathrm{d}\lambda_n. \tag{6}$$

In the following part, we will show the exact semi-circle for the GUE case and give numerical approaches that calculate the level density efficiently. Notice that such formulas also exist for the finite GOE and GSE, we refer interested readers to [15].

If A is a $n \times n$ complex Gaussian and we take $(A + A^T)/2$ which is an instance of the GUE ensemble, the eigenvalue density was derived by Wigner in 1962 as $\sum_{j=0}^{n-1} \phi_j^2(x)$, where

$$\phi_j(x) = (2^j j! \sqrt{\pi})^{-\frac{1}{2}} \exp(-x^2/2) H_j(x) \tag{7}$$

and $H_j(x)$ is the jth Hermite polynomial, which is defined as

$$H_j(x) = \exp(x^2) \left(-\frac{\mathrm{d}}{\mathrm{d}x}\right)^j \exp(-x^2) = j! \sum_{i=0}^{j/2} (-1)^i \frac{(2x)^{j-2i}}{i!(j-2i)!}.$$

Figure 6 compares the normalized level density of the GUE for different values of n.

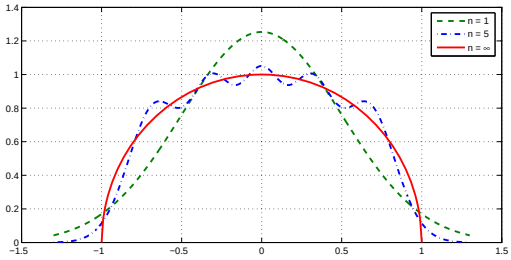

Fig. 6 Level density of the GUE ensemble ($\beta = 2$) for different values of n. The limiting result when $n \to \infty$ is Wigner's famous semi-circle law.

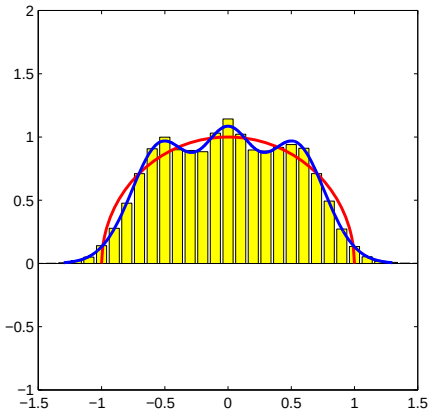

Fig. 7 The exact semicircular law with ten thousand matrices. Plotted is the histogram (30000 eigenvalues) on the real line. The unit semi-circle and the exact density function which is defined in terms of Hermite polynomials.

Code 2 Calculating the level density for the finite GUE.

```
%Experiment:   Eigenvalues of GUE matrices
%Plot:         Histogram of eigenvalues
%Theory:       Semicircle and finite semicircle
%% Parameters
n=3;           % size of matrix
s=10000;       % number of samples
d=.1;          % bin size
e=[];          % eigenvalue samples
%% Experiment
for i=1:s
    a=randn(n)+sqrt(-1)*randn(n);
    a=(a+a')/(2*sqrt(4*n));
    v=eig(a);
    e=[e v];
end
[m x]=hist(e,-1.5:d:1.5);
bar(x,m*pi/(2*d*n*s),'y');
axis('square')
axis([-1.5 1.5 -1 2])
%% Theory
hold on
t=-1:.01:1;
plot(t,sqrt(1-t.^2),'r', 'Linewidth', 2) % Plot the semicircle
levels(n)                                % Plot the finite
    semicircle
hold off
```

Most useful for computation is the three term recurrence

$$H_{j+1}(x) = 2xH_j(x) - 2jH_{j-1}(x), \tag{8}$$

starting with $H_{-1} = 0$ and $H_0 = 1$ so that $H_1(x) = 2x$. Therefore, ignoring the normalization term $(\sqrt{\pi})^{-\frac{1}{2}} \exp(-x^2/2)$ in $\phi_j(x)$, we define

$$\widetilde{\phi}_j(x) = (2^j j!)^{-\frac{1}{2}} H_j(x).$$

From Equation 8, we can get the three term recurrence for $\widetilde{\phi}_j(x)$ as follows

$$\sqrt{j} \cdot \tilde{\phi}_j(x) = \sqrt{2}x \cdot \widetilde{\phi}_{j-1}(x) - \sqrt{j-1} \cdot \widetilde{\phi}_{j-2}(x). \tag{9}$$

Based on Equation 9, one can do the direct calculation of summing $\phi_n(x)$ for each x. But there are two better ways to calculate the level density.

1. The first approach is based on the following equation

$$\sum_{j=0}^{n-1} \widetilde{\phi}_j^2(x) = n\widetilde{\phi}_n^2(x) - \sqrt{n(n+1)}\widetilde{\phi}_{n-1}(x)\widetilde{\phi}_{n+1}(x). \tag{10}$$

   This formula comes from the famous Christoffel-Darboux relationship for orthogonal polynomials. Therefore, we can combine Equation 10 with the three term recurrence for $\widetilde{\phi}$ and Code 3 realizes the idea.

Code 3 Computing the level density (GUE) using Christoffel-Darboux.

```
function z=levels(n)
%Plot exact semicircle formula for GUE
x=[-1:.001:1]*sqrt(2*n)*1.3;
pold = 0*x;                          % -1st Hermite polynomial
p= 1+0*x;                            %  0th Hermite polynomial
k=p;
for j=1:n;                           % Three term recurrence
   pnew = (sqrt(2)*x.*p-sqrt(j-1)*pold)/sqrt(j);
   pold = p; p=pnew;
end
pnew = (sqrt(2)*x.*p-sqrt(n)*pold)/sqrt(n+1);
k = n*p.^2-sqrt(n*(n+1))*pnew.*pold; % Use p.420 of Mehta
% Multiply the correct normalization
k=k.*exp(-x.^2)/sqrt(pi);
% Rescale so that "semicircle" is on [-1,1] and area is pi/2
plot(x/sqrt(2*n),k*pi/sqrt(2*n),'b', 'Linewidth', 2);
```

2. The other way comes from the following interesting equivalent expression

$$\sum_{j=0}^{n-1} \phi_j^2(x) = \|(\sqrt{\pi})^{-\frac{1}{2}} \exp(-x^2/2) \cdot v\|^2, \tag{11}$$

where

$$v = \frac{u}{u_1}, \quad u = (T - \sqrt{2}x \cdot \mathbb{I})^{-1} e_{n-1}. \tag{12}$$

where u_1 is the first element of u and e_{n-1} is the column vector where only the $n-1$st entry is 1. T is a tridiagonal matrix that is related to the three term recurrence such that

$$T = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots & 0 \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \cdots & 0 \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{n-1} & 0 \end{bmatrix}.$$

To see this, from Equation 9, we have the following relation

$$\begin{bmatrix} -\sqrt{2}x & \sqrt{1} & 0 & \cdots & 0 \\ \sqrt{1} & -\sqrt{2}x & \sqrt{2} & \cdots & 0 \\ 0 & \sqrt{2} & -\sqrt{2}x & \sqrt{3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sqrt{n-1} & -\sqrt{2}x \end{bmatrix} \begin{bmatrix} \widetilde{\phi}_0(x) \\ \widetilde{\phi}_1(x) \\ \widetilde{\phi}_2(x) \\ \vdots \\ \widetilde{\phi}_{n-1}(x) \end{bmatrix} = C \times \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

where C can be determined easily by the initial condition $\widetilde{\phi}_0(x) = 1$ which justifies Equation 12.

Though one can easily use Equation 11 to compute the density at x, we can avoid inverting $T - \sqrt{2}x \cdot \mathbb{I}$ for every x. Given the eigendecomposition of $T = H\Lambda H^T$, we have

$$\begin{aligned} u &= (T - \sqrt{2}x \cdot \mathbb{I})^{-1} e_{n-1} \\ &= H(\Lambda - \sqrt{2}x \cdot \mathbb{I})^{-1} H^T e_{n-1}. \end{aligned}$$

Thus, for each x, we only need to invert the diagonal matrix $\Lambda - \sqrt{2}x \cdot \mathbb{I}$, provided that H is stored beforehand. Code 4 gives the corresponding implementation.

Code 4 Computing the level density (GUE) using the tridiagonal matrix.

```
function z = levels2(n)
%% Plot exact semicircle formula for GUE
xfull = [-1:.001:1] * sqrt(2*n) * 1.3;
% Form the Tridiagonal matrix
T = diag(sqrt(1:n-1),1);
T = T+T';
% Do the eigendecomposition of T, T = UVU'
[U, V] = eig(T);
% extract the eigenvalues
V = diag(V);
% precompute U'*e_n
% tmp_en = U' * ((0:n-1) == n-1)';
tmp_en = U(end, :)';
for i = 1:length(xfull),
    x = xfull(i);
    % generate the v vector as in (2.5)
    v = U * (tmp_en./(V - sqrt(2)*x));
    % multiply the normalization term
    y(i) = norm((sqrt(pi))^(-1/2) * exp(-x^2/2) * v/v(1))^2;
end
% Rescale so that "semicircle" is on [-1,1] and area is pi/2
plot(xfull/sqrt(2*n), y*pi/sqrt(2*n), 'r--', 'Linewidth', 2);
```

2.2 Marčenko-Pastur Law (Special case: Quarter circle Law)

Laguerre ensemble consists of random matrices $L = A^T A/m$ where $A = G_\beta(m,n)$ and the following theorem gives the limiting eigenvalue distribution of its eigenvalues [14] when $\beta = 1$ (real case).

Theorem 2. *(Marčenko, Pastur 1967) Let X_n be a sequence of random symmetric $m \times n$ matrices $(n = 1,2,\ldots)$, with $m \geq n$, satisfying*

1. *Independence: The elements x_{ij} of each $X_n$are independent random variables.*
2. *Zero Mean: The elements x_{ij} of each X_n satisfy $\mathbb{E}(x_{ij}) = 0$.*
3. *Unit variance: The elements x_{ij} of each X_n satisfy $\mathbb{E}(x_{ij}^2) = 1$.*
4. *Bounded moments: There is some bound B, independent of n, such that $\forall n, \mathbb{E}(|x_{ij}|^k) \leq B$.*
5. *Asymptotic Aspect Ratio: m depends on n in such a way that $n/m \to r \leq 1$ as $n \to \infty$.*

Under these assumptions, the distribution of the eigenvalues of $\frac{1}{m}X_n^T X_n$ asymptotically approaches the Marčenko-Pastur law as $n \to \infty$,

$$f(x) = \frac{\sqrt{(x-a)(b-x)}}{2\pi xr},$$

where $a = (1-\sqrt{r})^2$ *and* $b = (1+\sqrt{r})^2$.

Code 5 Marčenko-Pastur Law

```
%Experiment: Gaussian Random
%Plot: Histogram of the eigenvalues of X'X/m
%Theory: Marcenko-Pastur as n->infinity
%% Parameters
t=1; %trials
r=0.1;     %aspect ratio
n=2000; %matrix column size
m=round(n/r);
v=[]; %eigenvalue samples
dx=.05; %binsize
%% Experiment
for i=1:t,
    X=randn(m,n); % random mxn matrix
    s=X'*X; %sym pos def  matrix
    v=[v;eig(s)]; % eigenvalues
end
v=v/m; % normalized eigenvalues
a=(1-sqrt(r))^2;b=(1+sqrt(r))^2;
%% Plot
[count,x]=hist(v,a:dx:b);
cla reset
bar(x, count/(t*n*dx),'y');
hold on;
%% Theory
x=linspace(a,b);
plot(x,sqrt((x-a).*(b-x))./(2*pi*x*r),'LineWidth',2)
axis([0 ceil(b) -.1 1.5]);
```

According to the Marčenko-Pastur Law, we have the density of the singular values of $X/\sqrt{m}$ as

$$f(s) = \frac{\sqrt{(s^2-a^2)(b^2-s^2)}}{\pi s r}.$$

When $r = 1$, we get the special case that

$$f(s) = \frac{1}{\pi}\sqrt{4-s^2},$$

on $[0,2]$. This is the famous **quarter circle law**. The singular values of a normally distributed square matrix lie on a quarter circle. The moments are Catalan numbers. We provide the code of the eigenvalue formulations in Code 5 and the singular value formulation in Code 6 with figures shown in Figure 8.

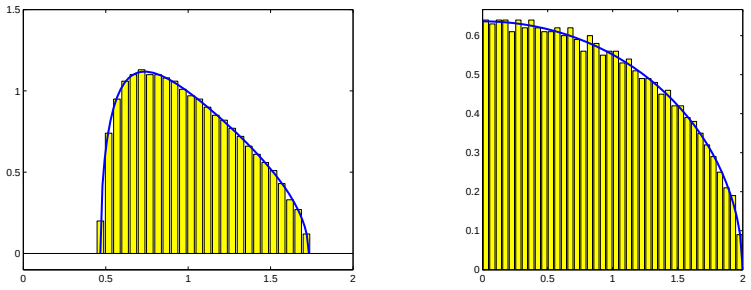

Fig. 8 Left: Marčenko-Pastur Law ($r = 0.1$) with a 20000×2000 matrix X. Plotted is the histogram of the 2000 eigenvalues of $X^T X/20000$; Right: Quarter Circle Law ($r = 1$) with a 2000×2000 matrix. Plotted are histograms of its singular values.

Code 6 Quarter circle law

```
%Experiment: Gaussian Random
%Plot: Histogram singular values
%Theory: Quatercircle Law
%% Parameters
t=1; %trials
r=1;     %aspect ratio
n=2000; %matrix column size
m = n;
v=[]; %eigenvalue samples
dx=.05; %binsize
a = 0; b = 2;
%% Experiment
for i=1:t,
  v=[v;svd(randn(n))]; % singular values
end
v=v/sqrt(m); % normalized
close all;
[count, x]=hist(v,(a-dx/2):dx:b); cla reset
bar(x, count/(t*n*dx),'y'); hold on;
%% Theory
x=linspace(a,b);
plot(x,sqrt(4 - x.^2)/pi,'LineWidth',2)
axis square
axis([0 2 0 2/3]);
```

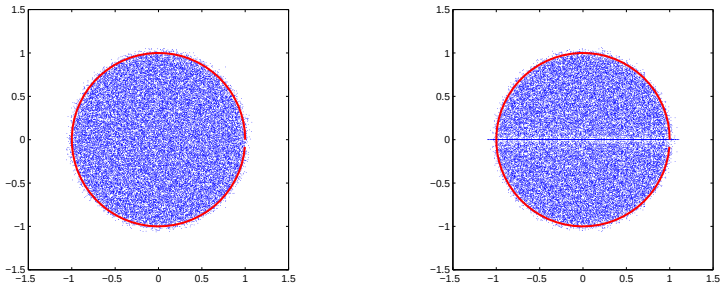

Fig. 9 Circular Law for the real and the complex case with 200 200×200 random matrices. Plotted is eigenvalues in the complex plain (40000 eigenvalues).

2.3 Circular Law

The eigenvalues of Hermite and Laguerre ensembles are distributed on the real line. In general, an interesting question is that when properly normalized, how are the eigenvalues `randn(n)` distributed on the complex plain. The following theorem provides the answer [13].

Theorem 3. *(Girko, 1984) The complex eigenvalues divided by $\sqrt{n}$ of an $n \times n$ random matrix with independent elements of mean 0 and variance 1 converge (under reasonable conditions) to the uniform distribution on the unit disk in the complex plane.*

The "Saturn effect": Notice the concentration of eigenvalues on the real line and the gap near the real line. The real line serves as an attractor to some of the eigenvalues. Two things are worth mentioning:

1. The Saturn effect is consistent with the circular law. As $n \to \infty$, the $O(\sqrt{n})$ eigenvalues on the real line do not matter. Also the diminishing repulsion is consistent with the circular law.
2. There are $O(\sqrt{n})$ real eigenvalues sometimes clashes with our intuition. After all, the real line is a set of measure 0. Why should there be so many eigenvalues on the real line?

2.4 Tracy-Widom Distribution (Law)

Tracy-Widom law provides the limiting density for the largest eigenvalue of the Hermite ensemble. The probability density for the Tracy-Widom distribution is given by the formula

$$p(x) = \frac{\mathrm{d}}{\mathrm{d}x} \exp\left(- \int_x^\infty (t-x) q(t)^2 \mathrm{d}t \right),$$

where $q(t)$, is defined as the solution of a so-called Painlevé II differential equation:

$$\ddot{q}(t) = tq(t) + 2q(t)^3,$$

with the boundary condition that as $t \to \infty$, $q(t)$ is asymptotic to the Airy function $\mathrm{Ai}(t)$.

While this may seem more formidable than the normal and semi-circle distributions, there are codes that may be used as black boxes for accurately calculating the Tracy-Widom distribution. This distribution depicts the histogram of the largest eigenvalue of a complex version of the random symmetric matrices. The distribution has also been showing up in many other applications. We show in Code 7 that even the formidable is but a few lines of MATLAB. It is based on solving the following differential equation

$$\frac{d}{dt}\begin{Bmatrix} q \\ q' \\ I \\ I' \end{Bmatrix} = \begin{Bmatrix} q' \\ tq + 2q^3 \\ I' \\ q^2 \end{Bmatrix},$$

where $I(x) = \int_x^\infty (t-x)q(t)^2 dt$. This has the advantage of evolving the needed quantity $I(x)$, as we go rather than post-processing. In Code 8, we calculate the largest eigenvalue of an instance from the 2-Hermite ensemble (GUE). We normalize the eigenvalues by subtracting them by $2\sqrt{n}$ and multiplying them by $n^{1/6}$.

Code 7 Calculating the Tracy-Widom Distribution.

```
%Theory: Compute and Plot the Tracy-Widom Distribution
%%Parameters
t0=5;     %right endpoint
tn=-8;    %left endpoint
dx=.005; %discretization
%%Theory: The differential equation solver
deq=@(t,y) [y(2); t*y(1)+2*y(1)^3; y(4); y(1)^2];
opts=odeset('reltol',1e-12,'abstol',1e-15);
y0=[airy(t0);airy(1,t0);0;airy(t0)^2]; % boundary conditions
[t,y]=ode45(deq,t0:-dx:tn,y0,opts); %solve
F2=exp(-y(:,3));     % the distribution
f2=gradient(F2,t);  % the density
%% Plot
%cla reset
plot(t,f2,'LineWidth',2)
axis([-5 2 0 .5]);
```

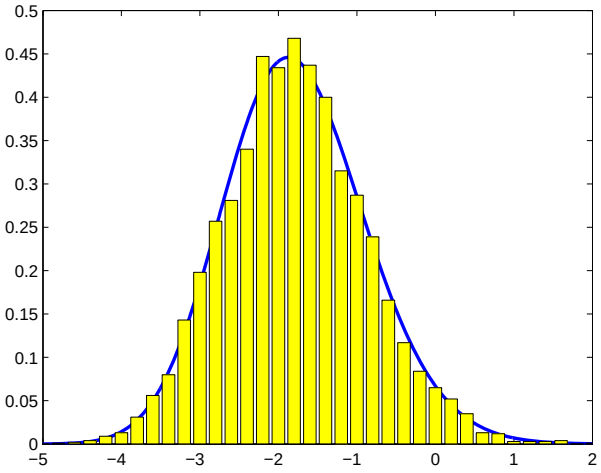

Fig. 10 Tracy-Widom law with 5000 matrices. Plotted is the histogram of the 500000 eigenvalues and the Tracy-Widom distribution on the blue line.

Code 8 Largest Eigenvalues of a random Hermitian matrix.

```
%Experiment:  Largest Eigenvalue of Random Hermitian Matrices
%Plot:        Histogram of the normalized largest eigenvalues
%Theory:      Tracy-Widom as n->infinity
%% Parameters
n=100;        %matrix size
t=5000;       %trials
v=[];         %eigenvalue samples
dx=.2;        %binsize
%% Experiment
for i=1:t,
    a=randn(n)+sqrt(-1)*randn(n);    % random nxn complex matrix
    s=(a+a')/2;          % Hermitian matrix
    v=[v;max(eig(s))];   % Largest Eigenvalue
end
v=n^(1/6)*(v-2*sqrt(n));    % normalized eigenvalues
%% Plot
hold on
[count,x]=hist(v,-5:dx:2);
bar(x, count/(t*dx),'y');
```

3 Random Matrix Factorization

A computational trick can also be a theoretical trick. Therefore do not dismiss an efficient computation as a mere "implementation detail", it may be where the next theory comes from.

Direct random matrix experiments usually involve `randn(n)`. Since many linear algebra computations require $\mathcal{O}(n^3)$ operations, it seems more feasible to take n relatively small, and take a large number of Monte Carlo instances. This has been our strategy in the example codes so far.

In fact, matrix computations involve a series of reductions. With normally distributed matrices, the most expensive reduction steps can be avoided on the computer as they can be done with mathematics! All of a sudden $\mathcal{O}(n^3)$ computations become $\mathcal{O}(n^2)$ or even better.

3.1 The Chi-distribution and orthogonal invariance

There are two key facts to know about a vector of independent standard normals. Let v_n denote such a vector. In MATLAB this would be `randn(n,1)`. Mathematically, we say that the n elements are independent and i.i.d. standard normals (mean 0, variance 1).

- **Chi distribution**: the Euclidean length $\|v_n\|$, which is the square root of the sum of the n squares of Gaussians, has what is known as the χ_n distribution.
- **Orthogonal Invariance**: for any fixed orthogonal matrix Q, or if Q is random and independent of v_n, the distribution of Qv_n is identical to that of v_n. In other words, it is impossible to tell the difference between a computer-generated v_n or Qv_n upon inspecting only the output.

We shall see that these two facts allow us to very powerfully transform matrices involving standard normals to simpler forms. For reference, we mention that the χ_n distribution has the probability density

$$f(x) = \frac{x^{n-1}e^{-x^2/2}}{2^{n/2-1}\Gamma(n/2)}. \tag{13}$$

There is no specific requirement that n be an integer, despite our original motivation as the length of a Gaussian vector. The square of χ_n is the distribution that underlies the well known *Chi-squared test*. It can be seen that the mean of χ_n^2 is n. (For integers, it is the sum of the n standard normal variables). We have that v_n is the product of the random scalar χ_n, which serves as the length, and an independent vector that is uniform on the sphere, which serves as the direction.

3.2 The QR decomposition of `randn(n)`

Given a vector v_n, we can readily construct an orthogonal reflection or rotation H_n such that $H_n v_n = \pm\|v_n\|e_1$, where e_1 denotes the first column of the identity. In matrix computations, there is a standard technique known as constructing a *Householder transformation* which is a reflection across the external angle bisector of these two vectors.

Therefore, if v_n follows a multivariate standard normal distribution, $H_n v_n$ yields a Chi distribution for the first element and 0 otherwise. Furthermore, let `randn(n)` be an $n \times n$ matrix of iid standard normals. It is easy to see now that through successive Householder reflections of size $n, n-1, \ldots$ we can orthogonally transform `randn(n)` into the upper triangular matrix

$$H_1 H_2 \cdots H_{n-1} H_n \times \texttt{randn(n)} = R_n = \begin{pmatrix} \chi_n & G & G & \ldots & G & G & G \\ & \chi_{n-1} & G & \ldots & G & G & G \\ & & \chi_{n-2} & \ldots & G & G & G \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \chi_3 & G & G \\ & & & & & \chi_2 & G \\ & & & & & & \chi_1 \end{pmatrix}.$$

Here all elements are independent and represent a distribution and the "G"' are all i.i.d. standard normals. It is helpful to watch a 3×3 real Gaussian matrix ($\beta = 1$) matrix turn into R:

$$\begin{pmatrix} G & G & G \\ G & G & G \\ G & G & G \end{pmatrix} \to \begin{pmatrix} \chi_3 & G & G \\ 0 & G & G \\ 0 & G & G \end{pmatrix} \to \begin{pmatrix} \chi_3 & G & G \\ 0 & \chi_2 & G \\ 0 & 0 & G \end{pmatrix} \to \begin{pmatrix} \chi_3 & G & G \\ 0 & \chi_2 & G \\ 0 & 0 & \chi_1 \end{pmatrix}.$$

The "G"'s as the computation progresses are not the same numbers, merely indicating the distribution. One immediate consequence is the following interesting fact

$$\mathbb{E}[\mathbf{det}(\texttt{randn(n)}^2)] = n! \tag{14}$$

This could also be obtained for any $n \times n$ matrix with independent entries with mean 0 and variance 1, by squaring the "big formula" for the determinant, noting that cross terms have expectation 0, and the $n!$ squared terms each have expectation 1.

3.2.1 Haar measure on Orthogonal matrices

Let Q be a random orthogonal matrix, such that one can not tell the difference between the distribution of AQ and Q for any fixed orthogonal matrix A. We say that Q has the uniform or *Haar distribution* on orthogonal matrices.

From the previous part of this section, with a bit of care we can say that `randn(n)`=(orthogonal uniform with Haar measure)(R_n) is the QR decomposition of `randn(n)`. Therefore, code for generating Q can be as simple as `[Q,~]=qr(randn(n))`.

Similarly, `[Q,~]=qr(randn(n)+sqrt(-1)*randn(n))` gives a random unitary matrix Q. For unitary matrix Q, its eigenvalues will be complex with a magnitude of 1, i.e. they will be distributed on the unit circle in the complex plane. Code 9 generates a random unitary matrix and histograms the angles of its eigenvalues.

Code 9 Sample a random unitary matrix.

```
%Experiment:   Generate random orthogonal/unitary matrices
%Plot:         Histogram eigenvalues
%Theory:       Eigenvalues are on unit circle
%% Parameters
t=5000;          %trials
dx=.05;          %binsize
n=10;            %matrix size
v=[];            %eigenvalue samples
%% Experiment
for i=1:t
    % Sample random unitary matrix
    [X ~]=qr(randn(n)+sqrt(-1)*randn(n));
    % If you have non-uniformly sampled eigenvalues, you may
        need this fix
    X=X*diag(sign(randn(n,1)+sqrt(-1)*randn(n,1)));
    v=[v;eig(X)];
end
%% Plot
x=(-(1+dx/2):dx:(1+dx/2))*pi;
h1=rose(angle(v),x);
set(h1,'Color','black')
%% Theory
hold on
h2=polar(x, t*n*dx/2*x.^0);
set(h2,'LineWidth',2)
hold off
```

3.2.2 Longest increasing subsequence

There is an interesting link between the moments of the eigenvalues of Q and the number of permutations of length with longest increasing subsequence k. For example, the permutation (3 1 8 4 5 7 2 6 9 10) has (1 4 5 7 9 10) or (1 4 5 6 9 10) as the longest increasing subsequence of length 6. For $n = 4$, there are 24 possible permutations listed in Table 3. We underline the fourteen permutations with longest increasing subsequence of length 2. Of these, one permutation (4 3 2 1) has length 1 and the other thirteen have length 2.

Table 3 Permutations for $n = 4$

1 2 3 4	2 1 3 4	3 1 2 4	4 1 2 3
1 2 4 3	2 1 4 3	3 1 4 2	4 1 3 2
1 3 2 4	2 3 1 4	3 2 1 4	4 2 1 3
1 3 4 2	2 3 4 1	3 2 4 1	4 2 3 1
1 4 2 3	2 4 1 3	3 4 1 2	4 3 1 2
1 4 3 2	2 4 3 1	3 4 2 1	4 3 2 1

Given a permutation of the numbers from 1 through n, the longest increasing subsequence may be found with the following admittedly cryptic algorithm "Patience sort".

Code 10 Patience sort.

```
function z=patiencesort(p)
% Patience sort
%
% Parameters
%  p : Permutation
% Returns
%  z : Length of longest increasing subsequence
piles=[];
for i=1:length(p)
    piles(1+sum(p(i)>piles))=p(i);
end
z=length(piles);
```

A remarkable result from random matrix theory is that the number of permutations of length n with longest increasing subsequence less than or equal to length k is given by

$$\mathbb{E}_{Q_k} |\mathbf{Tr}(Q_k)|^{2n}, \tag{15}$$

where Q_k is a $k \times k$ random unitary matrix. Who would have thought the moments of a random unitary matrix would count such a combinatorial object? Notice that the length of the permutation is n. Permutations of size 10 indicate the 20th moment of the absolute trace. The size of the matrix is the length of the longest increasing subsequence. Sometimes this may seem backwards, but it is correct.

Code 11 Random Orthogonal matrices and the Longest increasing sequence.

```
%Experiment:  Counts longest increasing subsequence statistics
t=200000;     % Number of trials
n=4;          % permutation size
k=2;          % length of longest increasing subsequence
v=zeros(t,1); % samples
for i=1:t
    [X,DC]=qr(randn(k)+sqrt(-1)*randn(k));
    X=X*diag(sign(randn(k,1)+sqrt(-1)*randn(k,1)));
    v(i)=abs(trace(X))^(2*n);
end
z = mean(v);
p = perms(1:n); c = 0;
for i=1:factorial(n)
    c = c + (patiencesort(p(i,:)) <= k);
end
[z c]
```

3.3 The tridiagonal reductions of GOE

Eigenvalues are usually defined early in one's education as the roots of the characteristic polynomial. Many people just assume that this is the definition that is used during a computation, but it is well established that this is not a good method for computing eigenvalues. Rather, a matrix factorization is used. In the case that S is symmetric, an orthogonal matrix Q is found such that $Q^T SQ = \Lambda$ is diagonal. The columns of Q are the eigenvectors and the diagonal of Λ are the eigenvalues.

Mathematically, the construction of Q is an iterative procedure, requiring infinitely many steps to converge. In practice, S is first tridiagonalized through a finite process which usually takes the bulk of the time. The tridiagonal is then iteratively diagonalized. Usually, this takes a negligible amount of time to converge in finite precision.

If $X =$ `randn(n)` and $S = (X + X^T)/\sqrt{2}$, then the eigenvalues of S follow the semi-circle law while the largest one follows the Tracy-Widom law. We can tridiagonalize S with the finite Householder procedure. The result is

$$T_n = \begin{pmatrix} G\sqrt{2} & \chi_{n-1} & & & & & \\ \chi_{n-1} & G\sqrt{2} & \chi_{n-2} & & & & \\ & \chi_{n-2} & G\sqrt{2} & \chi_{n-3} & & & \\ & & \chi_{n-3} & \ddots & \ddots & & \\ & & & \ddots & G\sqrt{2} & \chi_2 & \\ & & & & \chi_2 & G\sqrt{2} & \chi_1 \\ & & & & & \chi_1 & G\sqrt{2} \end{pmatrix}, \tag{16}$$

where $G\sqrt{2}$ refers to a Gaussian with mean 0 and variance 2. The superdiagonal and diagonal are independent, as the matrix is symmetric. The matrix T_n has the same eigenvalue distribution as S, but numerical computation of the eigenvalues is considerably faster when the right software is used.

A dense eigensolver requires $\mathscr{O}(n^3)$ operations and will spend nearly all of its time constructing T_n. Given that we know the distribution for T_n a priori, this is wasteful. The eigenvalues of T_n require $\mathscr{O}(n^2)$ time or better. In addition, dense matrix requires $\mathscr{O}(n^2)$ in storage while the tridiagonal matrix only needs $\mathscr{O}(n)$.

In a similar fashion, we can compute the singular values of a rectangular $m \times n$ matrix considerably faster by reducing it to bidiagonal form (shown here for $n > m$), as follows

$$B_n = \begin{pmatrix} \chi_m & \chi_{n-1} & & & & & & & \\ & \chi_{m-1} & \chi_{n-2} & & & & & & \\ & & \chi_{m-2} & \chi_{n-3} & & & & & \\ & & & \ddots & \ddots & & & & \\ & & & & \chi_3 & \chi_{n-m+1} & & & \\ & & & & & \chi_2 & \chi_{n-m} & & \\ & & & & & & \chi_1 & \chi_{n-m-1} \end{pmatrix}.$$

The story gets better. Random matrix experiments involving complex numbers or even over the quaternions reduce to real matrices even before they need to be stored on a computer. For general G_β, the R_n, tridiagonal and bidiagonal reduction have the following extensions

$$R_n = \begin{pmatrix} \chi_{n\beta} & G_\beta & G_\beta & \dots & G_\beta & G_\beta & G_\beta \\ & \chi_{(n-1)\beta} & G_\beta & \dots & G_\beta & G_\beta & G_\beta \\ & & \chi_{(n-2)\beta} & \dots & G_\beta & G_\beta & G_\beta \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \chi_{3\beta} & G_\beta & G_\beta \\ & & & & & \chi_{2\beta} & G_\beta \\ & & & & & & \chi_\beta \end{pmatrix}.$$

$$T_n = \begin{pmatrix} G\sqrt{2} & \chi_{(n-1)\beta} & & & & & \\ \chi_{(n-1)\beta} & G\sqrt{2} & \chi_{(n-2)\beta} & & & & \\ & \chi_{(n-2)\beta} & G\sqrt{2} & \chi_{(n-3)\beta} & & & \\ & & \chi_{(n-3)\beta} & \ddots & \ddots & & \\ & & & \ddots & G\sqrt{2} & \chi_{2\beta} & \\ & & & & \chi_{2\beta} & G\sqrt{2} & \chi_\beta \\ & & & & & \chi_\beta & G\sqrt{2} \end{pmatrix}.$$

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

Of interest is that T_n and B_n are real matrices whose eigenvalue and singular value distributions are exactly the same as the original complex and quaternion matrices. This leads to even greater computational savings because only real numbers need to be stored or computed with. Table 4 summaries how to generate instances from Hermite and Laguerre ensemble efficiently.

Table 4 Generating the β-Hermite and β-Laguerre ensembles efficiently.

Ensemble	MATLAB commands

Hermite	`% Pick n, beta`
	`d = sqrt(chi2rnd(beta * [n:-1:1]))';`
	`H = spdiags(d, 1, n, n) + spdiags(randn(n, 1), 0, n, n);`
	`H = (H + H') / sqrt(2);`

Laguerre	`% Pick m, n, beta`
	`% Pick a > beta * (n - 1)/2`
	`d = sqrt(chi2rnd(2 * a - beta * [0:1:n-1]))';`
	`s = sqrt(chi2rnd(beta * [n:-1:1]))';`
	`B = spdiags(s, -1, n, n) + spdiags(d, 0, n, n);`
	`L = B * B';`

There are many extremely important practical steps we can take at this point. We outline two interesting practical points.

Sturm sequences can be used with T_n for the computation of histograms [1]. This is particularly valuable when there is interest in a relatively small number of histogram intervals (say 20 or 30) and n is very large. This is an interesting idea, particularly because most people think that histogramming eigenvalues first requires that they compute the eigenvalues, then sort them into bins. The Sturm sequence idea gives a count without computing the eigenvalues at all. This is a fine example of not computing more than is needed: if you only need a count, why should one compute the eigenvalues at all?

For the largest eigenvalue, the best trick for very large n is to only generate the upper left $10\,n^{1/3} \times 10\,n^{1/3}$ of the matrix. Because of what is known as the "Airy" decay in the corresponding eigenvector, the largest eigenvalue – which technically depends on every element in the tridiagonal matrix – numerically depends significantly only on the upper left part. This is a huge savings in a Monte Carlo generation. Further savings can be obtained by using the Lanczos "shift and invert" strategy given an estimate for the largest eigenvalue. We refer interested reads to Section 10 of [10]. Code 12 provides an example of how we succeed to compute

the largest eigenvalue of a billion by billion matrix in the time required by naive methods for a hundred by hundred matrix.

Code 12 Compute the largest eigenvalues of a billion by billion matrix.

```
%% This code requires statistics toolbox
beta = 1; n = 1e9; opts.disp = 0; opts.issym = 1;
alpha = 10; k = round(alpha * n^(1/3)); % cutoff parameters
d = sqrt(chi2rnd(beta * n: -1: (n - k - 1)))';
H = spdiags(d, 1, k, k) + spdiags(randn(k, 1), 0, k, k);
H = (H + H')/sqrt(4 * n * beta); % Scale so largest eigenvlaue is
    near 1
eigs(H, 1, 1, opts);
```

3.4 Generalization beyond complex and quaternion

There is little reason other than history and psychology to only consider the reals, complexes, and quaternions $\beta = 1, 2, 4$. The matrices given by T_n and B_n are well defined for any β, and are deeply related to generalizations of the Schur polynomials knows as the Jack Polynomials of parameter $\alpha = 2/\beta$. Much is known, but much remains to be known. Edelman [11] proposes in his method of "Ghosts and Shadows" that even G_β exists and has a meaning upon which algebra might be doable.

Another interesting story comes from the fact that the reduced forms connect random matrices to the continuous limit, stochastic operators, which these authors believe represents a truer view of whys random matrices behave as they do [22].

4 Conclusion

In this paper, we give a brief summary of recent innovative applications in random matrix theory. We introduce the Hermite and Laguerre ensembles and give four famous laws (with MATLAB demonstration) that govern the limiting eigenvalue distributions of random matrices. Finally, we provide the details of matrix reductions that do not require a computer and give an overview of how these reductions can be used for efficient computation.

Acknowledgement

We acknowledge many of our colleagues and friends, too numerous to mention here, whose work has formed the basis of Random Matrix Theory. We particularly thank Raj Rao Nadakuditi for always bringing the latest applications to our attention.

References

1. J.T. Albrecht, C.P. Chan, and A. Edelman. Sturm sequences and random eigenvalue distributions. *Foundations of Computational Mathematics*, 9(4):461–483, 2009.
2. G.W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Cambridge Studies in Advanced Mathematics 118, 2010.
3. Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. Part II. Sample covariance matrices. *Annals of Probability*, 21:649–672, 1993.
4. Z.D. Bai. Convergence rate of expected spectral distributions of large random matrices. part i. wigner matrices. *The Annals of Probability*, pages 625–648, 1993.
5. Zhidong Bai and Jack Silverstein. *Spectral Analysis of Large Dimensional Random Matrices, 2nd edn*. Science Press, Beijing, 2010.
6. O.E. Barndorff-Nielsen and S. Thorbjørnsen. Lévy laws in free probability. *Proceedings of the National Academy of Sciences*, 99(26):16568, 2002.
7. R. Couillet and M. Debbah. *Random matrix methods for wireless communications*. Cambridge Univ Pr, 2011.
8. V. Dahirel, K. Shekhar, F. Pereyra, T. Miura, M. Artyomov, S. Talsania, T.M. Allen, M. Altfeld, M. Carrington, D.J. Irvine, et al. Coordinate linkage of hiv evolution reveals regions of immunological vulnerability. *Proceedings of the National Academy of Sciences*, 108(28):11530, 2011.
9. Ioana Dumitriu. *Eigenvalue Statistics for Beta-Ensembles*. Phd thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 2003.
10. A. Edelman and N.R. Rao. Random matrix theory. *Acta Numerica*, 14(233-297):139, 2005.
11. Alan Edelman. The random matrix technique of ghosts and shadows. *Markov Processes and Related Fields*, 16(4):783–790, 2010.
12. P.J. Forrester. *Log-gases and random matrices*. Number 34. Princeton Univ Pr, 2010.
13. V. L. Girko. Circular law. *Teor. Veroyatnost. i Primenen.*, 29:669–679, 1984.
14. VA Marčenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457, 1967.
15. M.L. Mehta. *Random matrices*, volume 142. Academic press, 2004.
16. F. Mezzadri and N.C. Snaith. *Recent perspectives in random matrix theory and number theory*. Cambridge Univ Pr, 2005.
17. R. J. Muirhead. *Aspects of multivariate statistical theory*. John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.
18. Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*. Number 335 in London Mathematical Society Lecture Note Series. Cambridge University Press, New York, 2006.
19. D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, 17(4):1617, 2007.
20. SM Popoff, G. Lerosey, R. Carminati, M. Fink, AC Boccara, and S. Gigan. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Physical review letters*, 104(10):100601, 2010.
21. P. Šeba. Parking and the visual perception of space. *Journal of Statistical Mechanics: Theory and Experiment*, 2009:L10002, 2009.

22. B.D. Sutton. *The stochastic operator approach to random matrix theory*. PhD thesis, Massachusetts Institute of Technology, 2005.
23. K.A. Takeuchi and M. Sano. Universal fluctuations of growing interfaces: evidence in turbulent liquid crystals. *Physical review letters*, 104(23):230601, 2010.
24. E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62:548–564, 1955.