The Entries of Circular Orthogonal Ensembles

Tiefeng Jiang [1]

Abstract. Let $\mathbf{V} = (v_{ij})_{n\times n}$ be a circular orthogonal ensemble. In this paper, for $1 \leq m \leq o(\sqrt{n}/\log n)$, we give a bound for the tail probability of $\max_{1\leq i,j\leq m} |v_{ij} - (1/n)\mathbf{y}_i'\mathbf{y}_j|$, where $\mathbf{Y} = (\mathbf{y}_1, \cdots, \mathbf{y}_n)$ is a certain $n \times n$ matrix whose entries are independent and identically distributed random variables with the standard complex normal distribution $\mathbb{C}N(0,1)$. In particular, this implies that, for a sequence of such matrices $\{\mathbf{V}_n = (v_{ij}^{(n)})_{n\times n},\ n \geq 1\}$, as $n \to \infty$, $\sqrt{n}v_{ij}^{(n)}$ converges in distribution to $\mathbb{C}N(0,1)$ for any $i \geq 1,\ j \geq 1$ with $i \neq j$, and $\sqrt{n}v_{ii}^{(n)}$ converges in distribution to $\sqrt{2}\cdot\mathbb{C}N(0,1)$ for any $i \geq 1$.

1 Introduction

The circular ensembles were first introduced by physicist Dyson [6, 7, 8] for the study of nuclear scattering data. For the definition of the ensembles, the density of eigenvalues, cluster functions, eigenvalue correlation functions, and eigenvalue nearest neighborhood spacing distributions as well as their connection to thermodynamics, one can see, for example, [17, 19].

There are three circular ensembles: circular orthogonal ensembles (COE), circular unitary ensembles (CUE) and circular symplectic ensembles (CSE). According to Theorem 9.1.1 from [17], a *circular orthogonal ensemble* is an $n \times n$ symmetric unitary random matrix $\mathbf{V}$, whose probability distribution is the same as that of $\mathbf{O}'\mathbf{V}\mathbf{O}$ for any unitary matrix $\mathbf{O}$. This can be realized by taking $\mathbf{V} = \mathbf{U}'\mathbf{U}$, where $\mathbf{U}$ is an Haar-invariant unitary matrix. For the other two ensembles, we will make some remarks at the end of this section.

The objective of this paper is to investigate the entries of the circular orthogonal ensembles. We will study the joint distributions of a big block of the matrices. In fact, we obtain a probability inequality on the difference between these entries and some simple functions of independent Gaussian random variables.

The main theme of random matrix theory is the investigation of eigenvalues. Another venue of this theory is the study of the dependency of entries of certain type of random matrices. For instance, D'Aristotle, Diaconis, and Newman[4], Diaconis, Eaton and Lauritzen[5], and Jiang[11, 12, 13], studied statistical testing problems and the image analysis based on the understanding of certain types of random matrices; some subsequent work are given in Li and Rosalsky [14], Li, Liu and

[1]Supported in part by NSF#DMS-0449365, School of Statistics, University of Minnesota, 224 Church Street, MN55455, tjiang@stat.umn.edu.

Key Words: random matrix theory, circular ensemble, Haar measure, weak convergence, Gaussian distribution.
AMS (2000) subject classifications: 28C10, 60B15, 60E99, 60F05.

Rosalsky [15], Liu, Lin and Shao [16] and Zhou [20]. Through investigating the entries of Haar-invariant unitary matrices, Jiang[10] recently proved that the limiting distribution of the largest eigenvalues of the Jacobi ensembles is the Tracy-Widom law.

The Hermite ensembles, the Laguerre ensembles and the Jacobi ensembles are three of major matrix models studied popularly in random matrix theory. Compared to them, the circular orthogonal ensembles are less known. In this paper, we investigate the entries of these ensembles. Before stating the main results, let us review some terminologies.

(a) If $X = (\xi + i\eta)/\sqrt{2}$, where ξ and η are independent and $N(0,1)$-distributed random variables, then we say X is a standard *complex Gaussian* random variable, and is denoted by $X \sim \mathbb{C}N(0,1)$, see, e.g., [1].

(b) For a sequence of random vectors $\{X_n;\ n = 0, 1, 2, \cdots\}$ defined on $\mathbb{C}$ or $\mathbb{R}^k$, we say X_n *converges to* X_0 *in probability* if $P(\|X_n - X_0\| \geq \epsilon) \to 0$ as $n \to \infty$ for any $\epsilon > 0$, where $\|\cdot\|$ is the Euclidean norm.

(c) Let $\{X_n;\ n = 0, 1, 2, \cdots\}$ be as in (b), we say X_n *converges weakly* or *converges in distribution to* X_0 if $\lim_{n\to\infty} Ef(X_n) \to Ef(X_0)$ for any bounded, continuous and real function $f(x)$ defined on $\mathbb{C}$ or $\mathbb{R}^k$.

Of course, (b) implies (c). Our main result is as follows.

Theorem 1 *For each $n \geq 1$, there exist an $n \times n$ circular orthogonal ensemble $\mathbf{V} = (v_{ij})$ and an $n \times n$ matrix $\mathbf{Y} = (\mathbf{y_1}, \cdots, \mathbf{y}_n)$ such that*

(i) the entries of $\mathbf{Y}$ and $\mathbf{V}$ are defined on the same probability space;

(ii) the n^2 entries of $\mathbf{Y}$ are i.i.d. random variables with the distribution of $\mathbb{C}N(0,1)$;

(iii) for any positive integers m and n satisfying that $n \geq 12m^2$, we have

$$P(\max_{1\leq i,j\leq m} |nv_{ij} - \mathbf{y}_i'\mathbf{y}_j| \geq t) \leq (K^{-1}n^2)e^{-Kt/m} \tag{1.1}$$

for any $0 < t \leq 4n/m$, where K is a constant not depending on n, m or t.

Through this probability inequality, we approximate some of the entries of a circular ensemble by $\mathbf{y}_i'\mathbf{y}_j/n$ for some (i,j)'s simultaneously. The following corollary tells us how large $m = m_n$ can be to make the uniform approximation valid.

Corollary 1.1 *For each $n \geq 1$, let the $n \times n$ matrices $\mathbf{V}_n = \mathbf{V} = (v_{ij}^{(n)})$ and $\mathbf{Y}_n = \mathbf{Y} = (\mathbf{y}_{n,1}, \cdots, \mathbf{y}_{n,n})$ be constructed as in Theorem 1. If $m_n = o(\sqrt{n}/\log n)$ as $n \to \infty$, then*

$$\max_{1\leq i,j\leq m_n} \left| \sqrt{n}v_{ij}^{(n)} - \frac{\mathbf{y}_{n,i}'\mathbf{y}_{n,j}}{\sqrt{n}} \right| \to 0$$

in probability as $n \to \infty$.

This corollary indicates that "$n^{-1/2}$" is the correct order of v_{ij}'s. In fact, we have a more refined result as follows.

Corollary 1.2 *Let* $\mathbf{V}_n = (v_{ij}^{(n)})_{n\times n}$*,* $n \geq 1$ *be a sequence of circular orthogonal ensembles. Then, as* $n \to \infty$*,* $\sqrt{n}v_{ij}^{(n)}$ *converges weakly to* $\mathbb{C}N(0,1)$ *for any* $i \geq 1$ *and* $j \geq 1$ *with* $i \neq j$*, and* $\sqrt{n}v_{ii}^{(n)}$ *converges weakly to* $\sqrt{2}\cdot\mathbb{C}N(0,1)$ *for any* $i \geq 1$*.*

Corollary 1.1 reveals that, in the asymptotic sense, the $m_n \times m_n$ upper-left block of an $n\times n$ circular orthogonal ensemble can be thought as $(1/n)\tilde{\mathbf{Y}}'\tilde{\mathbf{Y}}$, where $\tilde{\mathbf{Y}} = (y_{ij})$ is a $n\times m_n$ matrix whose entries are independent, $\mathbb{C}N(0,1)$-distributed random variables. Note that $\tilde{\mathbf{Y}}'\tilde{\mathbf{Y}}$ is not a Wishart matrix. A Wishart matrix, by using the current notation, is $\tilde{\mathbf{Y}}^*\tilde{\mathbf{Y}} = (z_{ij})_{m_n\times m_n}$, where the (i,j)-entry of $\tilde{\mathbf{Y}}^*$ is $\bar{y}_{ji}$. There is a big difference between the two matrices. In fact, Lemma 2.11 tells us that the off-diagonal entries of the Wishart matrix and those of $\tilde{\mathbf{Y}}'\tilde{\mathbf{Y}} = (w_{ij})_{m_n\times m_n}$ have the same asymptotic behavior. However, for any $i \geq 1$, the same lemma shows that $n^{-1/2}(z_{ii} - n)$ converges weakly to $N(0,1)$, whereas for the case of $\tilde{\mathbf{Y}}'\tilde{\mathbf{Y}}$, $n^{-1/2}w_{ii}$ converges weakly to $\sqrt{2}\cdot\mathbb{C}N(0,1)$ for any $i \geq 1$.

The method of our proof of Theorem 1 is the Gram-Schmidt algorithm, which is also used in [9, 11, 12] for studying Haar-invariant orthogonal, unitary and symplectic matrices. Although the Gram-Schmidt algorithm for generating orthogonal or unitary matrices is not generally stable in practice (see, e.g., [18]), it is quite efficient in investigating the coupling results as in Theorem 1. This is because the algorithm is more explicit than others.

Remark 1. Through the Gram-Schmidt algorithm as used in this paper, some approximation results on the classical compact groups $O(n)$, $U(n)$ and $Sp(n)$ are derived in [9, 11, 12]. However, all the proofs here do not depend on any of those.

Remark 2. In Corollary 1.1, the order $\sqrt{n}$ in $m_n = o(\sqrt{n}/\log n)$ is almost the best possible one obtained from Theorem 1. It might be improved to $n/(\log n)^\alpha$ for some $\alpha > 0$. To do so, one needs to examine more precisely the behavior of the three terms in (2.5), instead of the bounds used in the proof of Lemma 2.1.

Remark 3. Theorem 1 considers the case of circular orthogonal ensembles. We know that a circular unitary ensemble is actually an Haar-invariant unitary matrix, see, e.g., Theorem 9.3.1 from [17]. A theorem in Jiang[12] shows that the entries of an $n\times n$ circular unitary ensemble can be approximated by those of $\mathbf{Y}/\sqrt{n}$ in the same fashion as in Theorem 1, where the entries of $\mathbf{Y}$ are i.i.d. standard complex normal random variables. A characterization of the circular symplectic ensembles is given in Theorem 9.2.1 from [17], see also [6, 7, 8]. It basically says that such a matrix can be constructed by $\mathbf{U}^D\mathbf{U}$, where $\mathbf{U}$ is an Haar unitary matrix from $U(2n)$, and $\mathbf{U}^D = -\mathbf{Z}_{2n}\mathbf{U}^T\mathbf{Z}_{2n}$ with

$$\mathbf{Z}_{2n} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0} \end{pmatrix}.$$

So a result similar to Theorem 1 should hold for the case of the circular symplectic ensembles: the $m\times m$ upper-left block of a circular symplectic ensemble can be approximated by $(1/n)\mathbf{Y}^D\mathbf{Y}$, where $\mathbf{Y}$ is an $n\times m$ matrix whose entries are i.i.d. standard complex normals.

The main results stated above will be proved in the next section.

2 Proofs of Main results

Let ξ_1 and ξ_2 be two independent $N(0,1)$-distributed random variables. As mentioned in the Introduction, if $Z=(\xi_1+i\xi_2)/\sqrt{2}$, we then say Z follows the complex standard normal distribution, and is denoted by $Z\sim \mathbb{C}N(0,1)$. Further, if $X_1, X_2, \cdots, X_n$ are i.i.d. random variables with the distribution of $\mathbb{C}N(0,1)$, we then write $\mathbf{X}:=(X_1,\cdots,X_n)'\sim \mathbb{C}N_n(\mathbf{0},\mathbf{I}_n)$. For two random vectors ξ and η, the notation $\xi\sim\eta$ means that ξ and η have the same probability distribution. The following facts are very useful:

$$|X_1|^2\sim \text{Exp}(1),\ \|\mathbf{X}\|^2\sim \frac{1}{2}\chi^2(2n) \text{ and } \mathbf{e}'\mathbf{X}\sim \mathbb{C}N(0,1) \tag{2.1}$$

for any unit vector $\mathbf{e}\in\mathbb{C}^n$, where Exp(1) is the exponential distribution with parameter one, and $\chi^2(k)$ is the chi-square distribution with degree of freedom k.

Throughout this section, we assume $\mathbf{y}_1,\mathbf{y}_2,\cdots,\mathbf{y}_n$ are i.i.d. $n\times 1$ random vectors, and $\mathbf{y}_1\sim \mathbb{C}N_n(\mathbf{0},\mathbf{I}_n)$. Define $\mathbf{e}_1=\mathbf{y}_1/\|\mathbf{y}_1\|$, and

$$\mathbf{e}_k=\frac{\mathbf{y}_k-\mathbf{x}_k}{\|\mathbf{y}_k-\mathbf{x}_k\|} \tag{2.2}$$

where $\mathbf{x}_1=\mathbf{0}$ and $\mathbf{x}_k=\sum_{i=1}^{k-1}(\mathbf{e}_i^*\mathbf{y}_k)\mathbf{e}_i$ for $2\le k\le n$. Then, $\mathbf{U}=(\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n)$ is an $n\times n$ Haar unitary matrix, see, e.g., Jiang [9, 11, 12] or Mezzardi [18]. Therefore, $\mathbf{V}:=\mathbf{U}'\mathbf{U}=(v_{ij})_{n\times n}$ is an $n\times n$ circular orthogonal ensemble (COE). Clearly, $\mathbf{V}$ is a symmetric matrix. It is easy to see that

$$\begin{aligned} nv_{kl}=n\mathbf{e}_k'\mathbf{e}_l &= \frac{n(\mathbf{y}_k-\mathbf{x}_k)'(\mathbf{y}_l-\mathbf{x}_l)}{\|\mathbf{y}_k-\mathbf{x}_k\|\cdot\|\mathbf{y}_l-\mathbf{x}_l\|}\\ &= \mathbf{y}_k'\mathbf{y}_l+(a_ka_l-1)\,\mathbf{y}_k'\mathbf{y}_l-a_ka_l(\mathbf{x}_k'\mathbf{y}_l+\mathbf{y}_k'\mathbf{x}_l-\mathbf{x}_k'\mathbf{x}_l)\end{aligned}$$

for $1\le k\le l\le n$, where

$$a_k=\frac{\sqrt{n}}{\|\mathbf{y}_k-\mathbf{x}_k\|}. \tag{2.3}$$

Consequently,

$$|n\mathbf{e}_k'\mathbf{e}_l-\mathbf{y}_k'\mathbf{y}_l|\le |a_ka_l-1|\cdot|\mathbf{y}_k'\mathbf{y}_l|+|a_k|\cdot|a_l|\cdot(|\mathbf{x}_k'\mathbf{y}_l|+|\mathbf{y}_k'\mathbf{x}_l|+|\mathbf{x}_k'\mathbf{x}_l|). \tag{2.4}$$

From the expression $\mathbf{x}_1=\mathbf{0}$, $\mathbf{x}_k=\sum_{i=1}^{k-1}(\mathbf{e}_i^*\mathbf{y}_k)\mathbf{e}_i$ for $2\le k\le n$. Noticing $\mathbf{y}_k'\mathbf{e}_i=\mathbf{e}_i'\mathbf{y}_k$, we see that

$$\begin{aligned} \mathbf{x}_k'\mathbf{y}_l&=\sum_{i=1}^{k-1}(\mathbf{e}_i^*\mathbf{y}_k)(\mathbf{e}_i'\mathbf{y}_l),\quad \mathbf{y}_k'\mathbf{x}_l=\sum_{i=1}^{l-1}(\mathbf{e}_i^*\mathbf{y}_l)(\mathbf{e}_i'\mathbf{y}_k), \text{ and}\\ \mathbf{x}_k'\mathbf{x}_l&=\sum_{j=1}^{l-1}\sum_{i=1}^{k-1}(\mathbf{e}_i^*\mathbf{y}_k)(\mathbf{e}_j^*\mathbf{y}_l)\mathbf{e}_i'\mathbf{e}_j \end{aligned} \tag{2.5}$$

for $1\le k\le l\le n$. When "$\sum_{i=1}^{0}\cdots$" appears here and later, the sum is understood to be zero. Recalling a_i in (2.3), given $1\le l\le n$, set

$$\begin{aligned} A_1(l)&=\sqrt{n}\cdot\max_{1\le i,j\le l}|a_ia_j-1|,\\ A_2(1)&=0,\ A_2(l)=\max_{1\le i<j\le l}\{|\mathbf{e}_i^*\mathbf{y}_j|,\ |\mathbf{e}_i'\mathbf{y}_j|\} \text{ for } 2\le l\le n \text{ and},\\ A_3(l)&=\sqrt{n}\cdot\max_{1\le i\le j\le l}|\mathbf{e}_i'\mathbf{e}_j|. \end{aligned} \tag{2.6}$$

The restriction "$i < j$" in the definition of $A_2(l)$ is important. This is because $\mathbf{e}_i$ is a function of $\mathbf{y}_1, \cdots, \mathbf{y}_i$ by (2.2), hence is independent of $\mathbf{y}_j$ for $j > i$, and we know from (2.1) that both $\mathbf{e}_i^*\mathbf{y}_j$ and $\mathbf{e}_i'\mathbf{y}_j$ follow the distribution of $\mathbb{C}N(0,1)$. This fact will be used repeatedly later in the proofs. However, there are no such restrictions in the definition of $A_1(l)$ or $A_3(l)$.

Lemma 2.1 *Given $1 \le k \le l \le n$ and $t > 0$, set $C(l,t,n) = (13l)t^2 + 8n^{-1/2}l^2t^3$. Then*

$$\begin{aligned} P(|n\mathbf{e}_k'\mathbf{e}_l - \mathbf{y}_k'\mathbf{y}_l| \ge C(l,n,t)) \quad &\le \quad \sum_{i=1}^{3} P(A_i(l) \ge t) + P\left(\frac{\max_{1\le i\le j\le l}|\mathbf{y}_i'\mathbf{y}_j|}{\sqrt{n}} \ge t\right) \\ &\quad + P\left(\max_{1\le i\le l}\{a_i\} \ge 2\right) + P\left(\min_{1\le i\le l}\{a_i\} \le \frac{1}{2}\right). \end{aligned}$$

Proof. Remember $\mathbf{x}_1 = 0$. From (2.5), we have

$$|\mathbf{x}_k'\mathbf{y}_l| = |\sum_{i=1}^{k-1}(\mathbf{e}_i^*\mathbf{y}_k)(\mathbf{e}_i'\mathbf{y}_l)| \le \sum_{i=1}^{k-1}|\mathbf{e}_i^*\mathbf{y}_k|\cdot|\mathbf{e}_i'\mathbf{y}_l| \le kA_2(l)^2 \tag{2.7}$$

for all $1 \le k \le l \le n$ by the triangle inequality. By the same argument,

$$|\mathbf{x}_k'\mathbf{x}_l| \le \frac{l^2}{\sqrt{n}}\cdot A_2(l)^2A_3(l) \tag{2.8}$$

for all $1 \le k \le l \le n$. Now we estimate $|\mathbf{y}_k'\mathbf{x}_l|$.

First, by (2.2) and (2.3), $\mathbf{y}_k = \sum_{j=1}^{k-1}(\mathbf{e}_j^*\mathbf{y}_k)\mathbf{e}_j + (\sqrt{n}/a_k)\cdot\mathbf{e}_k$ for $1 \le k \le n$. Multiply $\mathbf{e}_i'$ from the left for both sides to have $\mathbf{e}_i'\mathbf{y}_k = \sum_{j=1}^{k-1}(\mathbf{e}_j^*\mathbf{y}_k)(\mathbf{e}_i'\mathbf{e}_j) + (\sqrt{n}/a_k)\cdot(\mathbf{e}_i'\mathbf{e}_k)$. It follows that

$$\begin{aligned} \max_{1\le i,\,k\le l}|\mathbf{e}_i'\mathbf{y}_k| \quad &\le \quad l\cdot A_2(l)\cdot\frac{A_3(l)}{\sqrt{n}} + \max_{1\le k\le l}\left\{\frac{\sqrt{n}}{a_k}\right\}\cdot\frac{A_3(l)}{\sqrt{n}} \\ &\le \quad \frac{l}{\sqrt{n}}\cdot A_2(l)A_3(l) + 2A_3(l) \end{aligned}$$

if $\min_{1\le i\le l}\{a_i\} \ge 1/2$. Second, from the middle identity in (2.5), for any $1 \le k \le l \le n$,

$$\begin{aligned} |\mathbf{y}_k'\mathbf{x}_l| \le \sum_{i=1}^{l-1}A_2(l)\cdot|\mathbf{e}_i'\mathbf{y}_k| \quad &\le \quad l\cdot A_2(l)\left(\frac{l}{\sqrt{n}}\cdot A_2(l)A_3(l) + 2A_3(l)\right) \\ &= \quad \frac{l^2}{\sqrt{n}}A_2(l)^2A_3(l) + (2l)A_2(l)A_3(l) \end{aligned} \tag{2.9}$$

provided $\min_{1\le i\le l}\{a_i\} \ge 1/2$. Combining (2.7), (2.8) and (2.9) with (2.4), for any $1 \le k \le l \le n$, we obtain

$$\begin{aligned} &|n\mathbf{e}_k'\mathbf{e}_l - \mathbf{y}_k'\mathbf{y}_l| \\ \le \quad & A_1(l)\cdot\frac{1}{\sqrt{n}}\max_{1\le i\le j\le l}|\mathbf{y}_i'\mathbf{y}_j| + \max_{1\le i\le l}\{a_i^2\}\cdot\Big\{k\cdot A_2(l)^2 + \frac{l^2}{\sqrt{n}}A_2(l)^2A_3(l) \\ & \qquad + (2l)A_2(l)A_3(l) + \frac{l^2}{\sqrt{n}}A_2(l)^2A_3(l)\Big\} \\ < \quad & t^2 + 4\left(kt^2 + \frac{l^2t^3}{\sqrt{n}} + (2l)t^2 + \frac{l^2t^3}{\sqrt{n}}\right) \le (13l)t^2 + \frac{8l^2t^3}{\sqrt{n}} \end{aligned}$$

provided $A_i(l) < t$ for $i = 1, 2, 3$, and $\max_{1\le i\le j\le l} |\mathbf{y}_i'\mathbf{y}_j|/\sqrt{n} < t$, and $1/2 < \min_{1\le i\le l}\{a_i\} \le \max_{1\le i\le l}\{a_i\} < 2$. Then the conclusion is yielded by considering the complement events. ■

With this lemma, to prove Theorem 1, it is enough to bound the six probabilities on the right hand side of the inequality in Lemma 2.1.

The following inequalities are elementary, which will be simply stated without proof. Throughout this paper, the symbol "log" stands for the natural logarithm.

Lemma 2.2 *The following inequalities hold.*

$$(i) \quad \frac{1}{(1-x)^2} \ge 1+2x \text{ and } \frac{1}{(1+x)^2} \le 1-x \text{ for } x \in (0, \frac{1}{4});$$
$$(ii) \quad \frac{x^2}{3} \le x - \log(1+x) \le \frac{x^2}{2} \text{ for } x \in (0, \frac{1}{2}];$$
$$(iii) \quad \log(1+x) - \log(1-x) > 2x \text{ for } x \in (0,1).$$

Lemma 2.3 *Suppose $n \ge 1$, $x \ge 0$ and $r \in (0, 1/4)$ satisfying that $nr \ge 2$ and $|(x/n) - 1| \ge r$. Then $|x/(n-i+1) - 1| \ge r/2$ for all $1 \le i \le nr/2$.*

Proof. From $|(x/n) - 1| \ge r$, we have that $|x - n| \ge rn$, then, by the triangle inequality, $|x - (n-i+1)| \ge rn - i + 1$, which implies $|x/(n-i+1) - 1| \ge (rn - i + 1)/(n-i+1)$. Now, $(rn-i+1)/(n-i+1) \ge r/2$ if and only if $2nr - 2i + 2 \ge nr - ir + r$, which is again equivalent to that $nr \ge 2i - (ir + 2 - r)$, which holds since $2i \le nr$ and $ir + 2 - r > 0$. ■

Lemma 2.4 *Let $\xi_1, \xi_2, \cdots, \xi_n$ be i.i.d. random variables with $\xi_1 \sim N(0,1)$. Then*

$$(i) \quad P\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 > a\right) \le 2e^{-nI(a)/2} \text{ and } P\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 < b\right) \le 2e^{-nI(b)/2} \text{ and}$$
$$(ii) \quad P\left(\frac{|\sum_{i=1}^n(\xi_i^2 - 1)|}{\sqrt{n}} \ge c\right) \le 2e^{-c^2/6}$$

for any $n \ge 1$, $a > 1$, $b \in (0,1)$ and $c \in (0, \sqrt{n}/2)$ where $I(x) = x - 1 - \log x$ for $x > 0$.

Proof. By the Chernoff bound, see, e.g., Remark (c) on p.27 from [3],

$$P\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 \in A\right) \le 2e^{-nJ(A)} \tag{2.10}$$

for any Borel set $A \subset \mathbb{R}$, where $J(A) = \inf_{x\in A} J(x)$ and $J(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \log E\exp(\theta\xi_1^2)\}$. It is shown on p.35 from [3] that $J(x) = (1/2)(x - 1 - \log x)$ for $x > 0$, and $J(x) = +\infty$, otherwise. It is easy to see that $J(x)$ is increasing on $[1, +\infty)$ and decreasing on $(0, 1]$. Take $A = [a, +\infty)$ in (2.10), then the first inequality in (i) follows. Similarly, the second holds.

Now we prove (ii). By (2.10),

$$P\left(\frac{|\sum_{i=1}^n \xi_i^2 - n|}{\sqrt{n}} \ge c\right) = P\left(|\frac{1}{n}\sum_{i=1}^n \xi_i^2 - 1| \ge \frac{c}{\sqrt{n}}\right) \le 2e^{-nJ(A)}$$

where $A = \{x : |x-1| \geq cn^{-1/2}\}$. By the monotone property of $J(x)$, since $0 < cn^{-1/2} < 1$ we know that $J(A) = \min\{J(1+cn^{-1/2}), J(1-cn^{-1/2})\}$. Now

$$\begin{aligned} & 2(J(1+cn^{-1/2}) - J(1-cn^{-1/2})) \\ = \;& 2cn^{-1/2} - \log(1+cn^{-1/2}) + \log(1-cn^{-1/2}) < 0 \end{aligned}$$

by (iii) of Lemma 2.2. Now, with (ii) of Lemma 2.2,

$$2J(A) = 2J(1+cn^{-1/2}) = cn^{-1/2} - \log(1+cn^{-1/2}) \geq \frac{c^2}{3n}$$

by the given condition $cn^{-1/2} < 1/2$, that is, $c \leq \sqrt{n}/2$. So (ii) holds. ■

Lemma 2.5 *Let $\mathbf{w}_i = \mathbf{y}_i - \mathbf{x}_i$ for $1 \leq i \leq n$, where $\mathbf{x}_i$ and $\mathbf{y}_i$ are as in (2.2). Then*

$$\|\mathbf{w}_i\|^2 \sim \frac{1}{2} \cdot \chi^2(2n-2i+2) \tag{2.11}$$

for all $1 \leq i \leq n$.

Proof. First, if $i = 1$, then $\mathbf{w}_1 = \mathbf{y}_1$. The assertion (2.11) follows from (2.1). Now assume $2 \leq i \leq n$. Recall $\mathbf{x}_1 = \mathbf{0}$ and $\mathbf{x}_i = \sum_{j=1}^{i-1}(\mathbf{e}_j^*\mathbf{y}_i)\mathbf{e}_j$ for $2 \leq i \leq n$. By the orthogonality, it is easy to verify that $\|\mathbf{w}_i\|^2 = \|\mathbf{y}_i\|^2 - \sum_{j=1}^{i-1}|(\mathbf{e}_j^*\mathbf{y}_i)|^2$. Write $|(\mathbf{e}_j^*\mathbf{y}_i)|^2 = \mathbf{y}_i^*\mathbf{e}_j\mathbf{e}_j^*\mathbf{y}_i$. Then

$$\|\mathbf{w}_i\|^2 = \mathbf{y}_i^*(\mathbf{I} - \sum_{j=1}^{i-1}\mathbf{e}_j\mathbf{e}_j^*)\mathbf{y}_i = \mathbf{y}_i^*\mathbf{\Sigma}\mathbf{y}_i,$$

where $\mathbf{\Sigma} := \mathbf{I}_n - (\mathbf{e}_1, \cdots, \mathbf{e}_{i-1}) \cdot (\mathbf{e}_1, \cdots, \mathbf{e}_{i-1})^*$. Now, since $\mathbf{e}_j$ is a function of $\mathbf{y}_1, \cdots, \mathbf{y}_j$ for any $1 \leq j \leq n$, we know that $\mathbf{y}_i$ and $\mathbf{e}_1, \cdots, \mathbf{e}_{i-1}$ are independent, so $\mathbf{\Sigma}$ and $\mathbf{y}_i$ are independent. By the orthogonality again, it is easy to check that $\mathbf{\Sigma}^* = \mathbf{\Sigma}$ and $\mathbf{\Sigma}^2 = \mathbf{\Sigma}$. Thus $\mathbf{\Sigma} = \mathbf{U}^* \operatorname{diag}(\mathbf{I}_r, \mathbf{0})\mathbf{U}$ for some unitary matrix $\mathbf{U}$, which is independent of $\mathbf{y}_i$, and

$$r := \operatorname{rank}(\mathbf{\Sigma}) = tr(\mathbf{\Sigma}) = n - tr((\mathbf{e}_1, \cdots, \mathbf{e}_{i-1})(\mathbf{e}_1, \cdots, \mathbf{e}_{i-1})^*) = n - i + 1$$

by the orthogonality of $\mathbf{e}_i$'s. Finally, noting that $\mathbf{U}\mathbf{y}_i$ and $\mathbf{y}_i$ have the same distribution by (2.1), and using the expression that $\mathbf{\Sigma} = \mathbf{U}^* \operatorname{diag}(\mathbf{I}_{n-i+1}, \mathbf{0})\mathbf{U}$, we see that $\mathbf{y}_i^*\mathbf{\Sigma}\mathbf{y}_i$ and $\mathbf{y}_i^*(\mathbf{I}_{n-i+1}, \mathbf{0})\mathbf{y}_i \sim (1/2)\chi^2(2n-2i+2)$ have the same distribution. The desired result follows. ■

Lemma 2.6 *Recall a_i as in (2.3). Then*

$$P\left(\min_{1\leq i\leq l} a_i \leq \frac{1}{2}\right) \leq (2l) \cdot e^{-n/12} \text{ and } P\left(\max_{1\leq i\leq l} a_i \geq 2\right) \leq (2l) \cdot e^{-n/12}$$

for any $1 \leq l \leq n/2$.

Proof. By Lemma 2.5, $a_i^{-2} \sim \chi^2(2n-2i+2)/(2n)$ for $i = 1, 2, \cdots, n$. Then, by Lemma 2.4,

$$\begin{aligned} P\left(\min_{1\leq i\leq l} a_i \leq \frac{1}{2}\right) &\leq l \cdot \max_{1\leq i\leq l} P\left(\frac{\chi^2(2n-2i+2)}{2n} \geq 4\right) \\ &\leq l \cdot P\left(\frac{\chi^2(2n)}{2n} \geq 4\right) \leq (2l)e^{-nI(4)} \end{aligned}$$

where $I(4) = 4 - 1 - \log 4 > 1$. So the last term above is dominated by $(2l)e^{-n}$. This gives the first inequality in the lemma. By the same argument,

$$\begin{aligned} P\left(\max_{1\le i\le l} a_i \ge 2\right) &\le l \cdot \max_{1\le i\le l} P\left(\frac{\chi^2(2n-2i+2)}{2n} \le \frac{1}{4}\right) \\ &\le l \cdot P\left(\frac{\chi^2(2n-2l)}{2n} \le \frac{1}{4}\right) \le l \cdot P\left(\frac{\chi^2(2n-2l)}{2n-2l} \le \frac{1}{2}\right) \end{aligned}$$

for any $1 \le l \le n/2$. By Lemma 2.4 again, the last term above is bounded by $(2l)e^{-(n-l)I(1/2)} \le (2l)e^{-(1/2)I(1/2)n}$. Trivially, $(1/2)I(1/2) = (1/2)((1/2) - 1 - \log(1/2)) = 0.09657\cdots > 1/12$. The second inequality in the lemma follows. ■

LEMMA 2.7 *For any $1 \le l \le n$, we have that $P(A_2(l) \ge t) \le l^2 e^{-t^2}$ for any $t > 0$.*

Proof. By notation, $A_2(1) = 0$, so without loss of generality, we assume $l \ge 2$. Since $\mathbf{e}_i$ is a unit complex vector, $\mathbf{e}_i'$ and $\mathbf{e}_i^*$ are also unit vectors. Also, $\mathbf{y}_j$ and $\{\mathbf{e}_i,\, 1 \le i \le j-1\}$ are independent. Then, $\mathbf{e}_i^*\mathbf{y}_j$ and $\mathbf{e}_i'\mathbf{y}_j$ are standard complex Gaussian random variables for any $1 \le i \le j-1$ by (2.1). This says that both $|\mathbf{e}_i^*\mathbf{y}_j|$ and $|\mathbf{e}_i'\mathbf{y}_j|$ have the same distribution as that of $\sqrt{Exp(1)}$. Therefore

$$P(A_2(l) \ge t) \le l^2 \cdot P(\sqrt{Exp(1)} \ge t) \le l^2 e^{-t^2}$$

for any $t > 0$. ■

LEMMA 2.8 *Let a_i's be as in (2.3) and $A_1(l)$ as in (2.6). Then*

$$P(A_1(l) > t) \le (2l)e^{-t^2/216}$$

for $0 < t \le 3\sqrt{n}/5$ and $1 \le l \le \sqrt{n}\,t/6$.

Proof. Let $\mathbf{w}_i = \mathbf{y}_i - \mathbf{x}_i$ and $L_i = |(\sqrt{n}/\|\mathbf{w}_i\|) - 1|$ for $1 \le i \le n$, where $\mathbf{x}_i$ and $\mathbf{y}_i$ are as in (2.2). We first claim that

$$P(L_i \ge r) \le 2e^{-nr^2/24} \tag{2.12}$$

for all $r \in (0, 1/4)$ and $1 \le i \le nr/2$. Assuming this holds, we prove the desired inequality.

Write $a_i a_j - 1 = (a_i - 1)(a_j - 1) + (a_i - 1) + (a_j - 1)$. Since $L_i = |a_i - 1|$, we have

$$A_1(l) \le \sqrt{n} \cdot \max_{1\le i\le l}\{L_i^2\} + 2\sqrt{n} \cdot \max_{1\le i\le l}\{L_i\} < 3\sqrt{n}s$$

if $\max_{1\le i\le l}\{L_i\} < s \le 1$. Take $s = t/3\sqrt{n} \le 1$, and use (2.12) to obtain

$$\begin{aligned} P(A_1(l) \ge t) \le P\left(\max_{1\le i\le l}\{L_i\} \ge \frac{t}{3\sqrt{n}}\right) &\le l \cdot \max_{1\le i\le l} P\left(L_i \ge \frac{t}{3\sqrt{n}}\right) \\ &\le (2l)e^{-t^2/216} \end{aligned}$$

since $t/3\sqrt{n} \le 1/5$ by assumption, and $1 \le l \le \sqrt{n}\,t/6 = n \cdot (t/3\sqrt{n})/2 = nr/2$ for $r := t/(3\sqrt{n})$ as required in (2.12). We get the inequality.

Now we turn to prove (2.12). If $|x-1|<r$, then by the Median-value Theorem, there exists $\xi_x \in (1-r, 1+r)$ such that $|(1/\sqrt{x})-1| = |x-1|/(2\xi_x^{3/2})$. Now $2\xi_x^{3/2} \geq 2(1-(1/4))^{3/2} > 1$, we get $|(1/\sqrt{x})-1| < r$. Therefore,

$$P(L_i \geq r) \leq P\left(\left|\frac{\|\mathbf{w}_i\|^2}{n} - 1\right| \geq r\right) \leq P\left(\left|\frac{\|\mathbf{w}_i\|^2}{(n-i+1)} - 1\right| \geq \frac{r}{2}\right) \tag{2.13}$$

for all $1 \leq i \leq nr/2$ by Lemma 2.3. From Lemma 2.5, $\|\mathbf{w}_i\|^2 \sim (1/2)\sum_{i=1}^{2(n-i+1)} \xi_i^2$ where ξ_i's are i.i.d. $N(0,1)$-distributed random variables. Thus, the last term in (2.13) is equal to

$$P\left(\frac{|\sum_{i=1}^{2(n-i+1)}(\xi_i^2-1)|}{\sqrt{2(n-i+1)}} \geq \frac{r}{2}\sqrt{2(n-i+1)}\right),$$

which, from Lemma 2.4, is bounded by $2\exp(-(n-i+1)r^2/12) \leq 2\exp(-nr^2/24)$ since $1 \leq i \leq nr/2$. ■

LEMMA 2.9 *For any $1 \leq l \leq n$, we have that*

$$P\left(\frac{1}{\sqrt{n}} \max_{1\leq i,j\leq l} |\mathbf{y}_i'\mathbf{y}_j| \geq t\right) \leq (10l^2)e^{-t^2/54}$$

for $0 < t \leq 3\sqrt{n}/2$.

Proof. Note that

$$P\left(\frac{1}{\sqrt{n}} \max_{1\leq i,j\leq l} |\mathbf{y}_i'\mathbf{y}_j| \geq t\right) \leq l \cdot P\left(\frac{|\mathbf{y}_1'\mathbf{y}_1|}{\sqrt{n}} \geq t\right) + l^2 P\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}} \geq t\right). \tag{2.14}$$

First, since $\mathbf{y}_1/\|\mathbf{y}_1\|$ is a unit vector and is independent of $\mathbf{y}_2$, $\mathbf{y}_1'\mathbf{y}_2/\|\mathbf{y}_1\| \sim \mathbb{C}N(0,1)$ by (2.1), it follows that $|\mathbf{y}_1'\mathbf{y}_2| \sim \|\mathbf{y}_1\| \cdot |\eta|$, where $\|\mathbf{y}_1\| \sim \{(1/2)\chi^2(2n)\}^{1/2}$ and $|\eta| \sim \sqrt{Exp(1)}$ by (2.1) again. Therefore,

$$\begin{aligned} P\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}} \geq t\right) &= P\left(\frac{\|\mathbf{y}_1\| \cdot |\eta|}{\sqrt{n}} \geq t\right) \\ &\leq P\left(\frac{\|\mathbf{y}_1\|}{\sqrt{n}} \geq \sqrt{2}\right) + P\left(|\eta| \geq \frac{t}{\sqrt{2}}\right). \end{aligned} \tag{2.15}$$

The last probability is equal to $P(Exp(1) \geq t^2/2) = e^{-t^2/2}$. By (i) of Lemma 2.4,

$$P\left(\frac{\|\mathbf{y}_1\|}{\sqrt{n}} \geq \sqrt{2}\right) = P\left(\frac{\chi^2(2n)}{2n} \geq 2\right) \leq 2e^{-(2n)I(2)/2} \leq 2e^{-n/5} \tag{2.16}$$

where $I(x) = (x - 1 - \log x)$ for $x > 0$, and hence $I(2) = 2 - 1 - \log 2 > 1/5$. In summary,

$$P\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}} \geq t\right) \leq 2e^{-n/5} + e^{-t^2/2} \tag{2.17}$$

for any $n \geq 1$ and $t > 0$.

Now $\mathbf{y}_1' \sim (1/\sqrt{2})(\xi_1 + i\eta_1, \cdots, \xi_n + i\eta_n)$, where $\{\xi_j, \eta_j, 1 \leq j \leq n\}$ are i.i.d. random variables with $\xi_1 \sim N(0,1)$. Thus,

$$\mathbf{y}_1'\mathbf{y}_1 = \frac{1}{2}\sum_{j=1}^{n}(\xi_j + i\eta_j)^2 = \frac{1}{2}\sum_{j=1}^{n}(\xi_j^2 - \eta_j^2) + i\sum_{j=1}^{n}\xi_j\eta_j. \tag{2.18}$$

It follows that

$$\begin{aligned}|\mathbf{y}_1'\mathbf{y}_1| &\le \frac{1}{2}|\sum_{j=1}^n(\xi_j^2-\eta_j^2)|+|\sum_{j=1}^n\xi_j\eta_j| \\ &\le |\sum_{j=1}^n(\xi_j^2-1)|+|\sum_{j=1}^n(\eta_j^2-1)|+|\sum_{j=1}^n\xi_j\eta_j|. \qquad (2.19)\end{aligned}$$

Then

$$P\left(\frac{|\mathbf{y}_1'\mathbf{y}_1|}{\sqrt{n}}\ge t\right)\le 2P\left(\frac{|\sum_{j=1}^n(\xi_j^2-1)|}{\sqrt{n}}\ge\frac{t}{3}\right)+P\left(\frac{1}{\sqrt{n}}|\sum_{j=1}^n\xi_j\eta_j|\ge\frac{t}{3}\right). \qquad (2.20)$$

The middle term above is bounded by $4e^{-t^2/54}$ for $0<t\le 3\sqrt{n}/2$ by (ii) of Lemma 2.4. Apply the same argument between (2.14) and (2.15) to the real case, we get that $\sum_{j=1}^n\xi_j\eta_j/\sqrt{n}\sim\sqrt{\chi^2(n)/n}\cdot N(0,1)$. This leads to

$$P\left(\frac{1}{\sqrt{n}}|\sum_{j=1}^n\xi_j\eta_j|\ge\frac{t}{3}\right)\le P\left(\frac{\chi^2(n)}{n}\ge 2\right)+P\left(|N(0,1)|\ge\frac{t}{3\sqrt{2}}\right). \qquad (2.21)$$

Similar to (2.16), the middle probability above is bounded by $2e^{-n/10}$. Using the inequality $P(|N(0,1)|\ge x)\le 2(\sqrt{2\pi}x)^{-1}e^{-x^2/2}$ for $x>0$, we have that

$$P\left(|N(0,1)|\ge\frac{t}{3\sqrt{2}}\right)\le\frac{6\sqrt{2}}{\sqrt{2\pi}}\cdot\frac{1}{t}\cdot e^{-t^2/36}\le e^{-t^2/36}$$

as $t\ge 4$ since $6\sqrt{2}/(4\sqrt{2\pi})\le 1$. Combining (2.20) with (2.21), and noting that $e^{-n/10}\le e^{-2t^2/45}$ for $0<t<3\sqrt{n}/2$, we have that

$$P\left(\frac{|\mathbf{y}_1'\mathbf{y}_1|}{\sqrt{n}}\ge t\right)\le 4e^{-t^2/54}+2e^{-n/10}+e^{-t^2/36}\le 7\cdot e^{-t^2/54} \qquad (2.22)$$

as $4\le t\le 3\sqrt{n}/2$. Now, if $t\in(0,4)$, then $7\cdot e^{-t^2/54}\ge 7\cdot e^{-16/54}\ge 7/e>1$. Thus, (2.22) holds for any $0<t<3\sqrt{n}/2$. Reviewing (2.17),

$$P\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}}\ge t\right)\le 2e^{-n/5}+e^{-t^2/2}\le 3e^{-t^2/12}$$

provided $0<t\le 3\sqrt{n}/2$. This together with (2.14) and (2.22) yields

$$P\left(\frac{1}{\sqrt{n}}\max_{1\le i,j\le l}|\mathbf{y}_i'\mathbf{y}_j|\ge t\right)\le(7l)e^{-t^2/54}+(3l^2)e^{-t^2/12}\le(10l^2)e^{-t^2/54}$$

for any $n\ge l\ge 1$ and $0<t\le 3\sqrt{n}/2$. ■

Lemma 2.10 *Let $n\ge 3$, $1\le l\le n$, and $A_3(l)$ be as in (2.6). Then*

$$P(A_3(l)\ge t)\le(47l^2)e^{-t^2/4050}$$

for $0<t\le 15\sqrt{n}/2$.

Proof. Recalling the statement immediately below (2.2), $(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ is an Haar-invariant unitary matrix, hence $\mathbf{e}_i$ and $\mathbf{e}_1$ have the same distribution, and $(\mathbf{e}_i, \mathbf{e}_j)$ and $(\mathbf{e}_1, \mathbf{e}_2)$ have the same distribution for any $1 \le i \ne j \le n$. It follows that

$$P(A_3(l) \ge s) \le l^2 \cdot \max_{1\le i\le j\le 2} P(\sqrt{n}\cdot|\mathbf{e}_i'\mathbf{e}_j| \ge s) \tag{2.23}$$

for any $s > 0$. Recall (2.2) to have that

$$\begin{aligned} \mathbf{e}_1 &= \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}; \\ \mathbf{e}_2 &= \frac{\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1}{\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|} = \frac{\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}}{\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|}. \end{aligned} \tag{2.24}$$

Hence, $\sqrt{n}|\mathbf{e}_1'\mathbf{e}_1| = \sqrt{n}\,|\mathbf{y}_1'\mathbf{y}_1|/\|\mathbf{y}_1\|^2 < 2s$, provided $\|\mathbf{y}_1\|^2/n > 1/2$ and $|\mathbf{y}_1'\mathbf{y}_1|/\sqrt{n} < s$. Therefore

$$P\left(\sqrt{n}|\mathbf{e}_1'\mathbf{e}_1| \ge 2s\right) \le P\left(\frac{\chi^2(2n)}{2n} \le \frac{1}{2}\right) + P\left(\frac{|\mathbf{y}_1'\mathbf{y}_1|}{\sqrt{n}} \ge s\right), \tag{2.25}$$

where the fact that $\|\mathbf{y}_1\|^2 \sim (1/2)\cdot\chi^2(2n)$ in (2.1) is used. Also, noting that $|\mathbf{y}_1'\mathbf{y}_1| \le \|\mathbf{y}_1\|^2$ by the Cauchy-Schwartz inequality, from (2.24),

$$\begin{aligned} \sqrt{n}|\mathbf{e}_1'\mathbf{e}_2| &= \frac{n}{\|\mathbf{y}_1\|\cdot\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|}\cdot\left|\frac{\mathbf{y}_1'\mathbf{y}_2}{\sqrt{n}} - \frac{\mathbf{y}_1'\mathbf{y}_1}{\sqrt{n}\,\|\mathbf{y}_1\|}(\mathbf{e}_1^*\mathbf{y}_2)\right| \\ &\le \frac{\sqrt{n}}{\|\mathbf{y}_1\|}\cdot\frac{\sqrt{n}}{\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|}\cdot\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}} + \frac{\|\mathbf{y}_1\|}{\sqrt{n}}|\mathbf{e}_1^*\mathbf{y}_2|\right) \\ &< \sqrt{2}\cdot\sqrt{2}\cdot(s+\sqrt{2}\,s) < 5s \end{aligned}$$

if $1/\sqrt{2} < \|\mathbf{y}_1\|/\sqrt{n} < \sqrt{2}$, $\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|/\sqrt{n} > 1/\sqrt{2}$, $|\mathbf{y}_1'\mathbf{y}_2|/\sqrt{n} < s$, and $|\mathbf{e}_1^*\mathbf{y}_2| < s$. Thus

$$\begin{aligned} P(\sqrt{n}|\mathbf{e}_1'\mathbf{e}_2| \ge 5s) \le\ & P\left(\frac{\|\mathbf{y}_1\|}{\sqrt{n}} \ge \sqrt{2}\right) + P\left(\frac{\|\mathbf{y}_1\|}{\sqrt{n}} \le \frac{1}{\sqrt{2}}\right) + P\left(\frac{\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|}{\sqrt{n}} \le \frac{1}{\sqrt{2}}\right) \\ & + P\left(\frac{|\mathbf{y}_1'\mathbf{y}_2|}{\sqrt{n}} \ge s\right) + P(|\mathbf{e}_1^*\mathbf{y}_2| \ge s). \end{aligned} \tag{2.26}$$

Now, $\|\mathbf{y}_1\|^2 \sim (1/2)\cdot\chi^2(2n)$ by (2.1), and $\|\mathbf{y}_2 - (\mathbf{e}_1^*\mathbf{y}_2)\mathbf{e}_1\|^2 \sim (1/2)\cdot\chi^2(2n-2)$ by Lemma 2.5 (taking $i=2$), and $|\mathbf{e}_1^*\mathbf{y}_2|^2 \sim Exp(1)$ by independence and (2.1). These together with (2.25) and (2.26) conclude that

$$\begin{aligned} & \max_{1\le i\le j\le 2} P(\sqrt{n}|\mathbf{e}_i'\mathbf{e}_j| \ge 5s) \\ \le\ & P\left(\frac{\chi^2(2n)}{2n} \ge 2\right) + 2P\left(\frac{\chi^2(2n-2)}{2n} \le \frac{1}{2}\right) + \max_{1\le i\le j\le 2} P\left(\frac{|\mathbf{y}_i'\mathbf{y}_j|}{\sqrt{n}} \ge s\right) + e^{-s^2} \end{aligned} \tag{2.27}$$

for all $n \ge 2$ and $s > 0$. By (2.16) and Lemma 2.9, we obtain

$$P\left(\frac{\chi^2(2n)}{2n} \ge 2\right) \le 2e^{-n/5} \le 2e^{-4s^2/45} \text{ and} \tag{2.28}$$

$$\max_{1\le i\le j\le 2} P\left(\frac{|\mathbf{y}_i'\mathbf{y}_j|}{\sqrt{n}} \ge s\right) \le 40\cdot e^{-s^2/54} \tag{2.29}$$

as $0 < s \leq 3\sqrt{n}/2$. Now, $n/(n-1) \leq 3/2$ for all $n \geq 3$, then

$$P\left(\frac{\chi^2(2n-2)}{2n} \leq \frac{1}{2}\right) \leq P\left(\frac{\chi^2(2n-2)}{2n-2} \leq \frac{3}{4}\right) \leq 2 \cdot e^{-(n-1)I(3/4)}$$

by (i) of Lemma 2.4. Now $I(3/4) = \log(1+(1/3)) - 1/4 \geq (1/3) - (1/(2 \cdot 3^2)) - (1/4) = 1/36$ by (ii) of Lemma 2.2. Moreover, $n - 1 \geq n/2$ for all $n \geq 2$. We then have

$$P\left(\frac{\chi^2(2n-2)}{2n} \leq \frac{1}{2}\right) \leq 2 \cdot e^{-n/72} \leq 2 \cdot e^{-s^2/162} \tag{2.30}$$

as $n \geq 3$ and $0 < s \leq 3\sqrt{n}/2$. Now, combine (2.23), and (2.27) to (2.30), we obtain that

$$P(A_3(l) \geq 5s) \leq (47l^2)e^{-s^2/162}$$

for all $n \geq 3$ and $0 < s \leq 3\sqrt{n}/2$. The desired conclusion follows by setting $t = 5s$. ■

Proof of Theorem 1. First, the condition that $n \geq 12m^2$ implies that $n \geq 12$. Recall $C(l,t,n) = (13l)t^2 + 8n^{-1/2}l^2t^3$. By Lemma 2.1 and Lemmas 2.6 to 2.10, for any $1 \leq k \leq l \leq n$,

$$\begin{aligned} & P(|n\mathbf{e}_k'\mathbf{e}_l - \mathbf{y}_k'\mathbf{y}_l| \geq C(l,n,t)) \\ \leq\ & (2l)e^{-t^2/216} + l^2e^{-t^2} + (47l^2)e^{-t^2/4050} + (10l^2)e^{-t^2/54} + (4l)e^{-n/12} \\ \leq\ & (64l^2)e^{-t^2/4050} \end{aligned}$$

for t, l and n satisfying that $0 < t \leq 3\sqrt{n}/5$ and $1 \leq k \leq l \leq \sqrt{n}\,t/6$, and $0 < t \leq 15\sqrt{n}/2$, and $0 < t \leq 3\sqrt{n}/2$, and $1 \leq l \leq n/2$, where the inequality $(4l)e^{-n/12} \leq (4l^2)e^{-25t^2/108}$ as $0 < t \leq 3\sqrt{n}/5$ is used in the last step. Optimize these conditions to have

$$P(|n\mathbf{e}_k'\mathbf{e}_l - \mathbf{y}_k'\mathbf{y}_l| \geq C(l,n,t)) \leq (64l^2)e^{-t^2/4050}$$

as $0 < t \leq 3\sqrt{n}/5$ and $1 \leq k \leq l \leq \sqrt{n}\,t/6$. For convenience, changing indices k to i and l to j respectively, the above becomes

$$P(|n\mathbf{e}_i'\mathbf{e}_j - \mathbf{y}_i'\mathbf{y}_j| \geq C(j,n,t)) \leq (64j^2)e^{-t^2/4050} \tag{2.31}$$

for $0 < t \leq 3\sqrt{n}/5$ and $1 \leq i \leq j \leq \sqrt{n}\,t/6$. We will sharpen this result next in three steps.

Step 1. If $0 < jt/\sqrt{n} \leq 1/2$, then

$$C(j,t,n) = (13j)t^2 + 8n^{-1/2}j^2t^3 \leq 17jt^2.$$

It follows from (2.31) that

$$\begin{aligned} & P(\max_{1\leq i,j\leq m} |n\mathbf{e}_i'\mathbf{e}_j - \mathbf{y}_i'\mathbf{y}_j| \geq 17mt^2) \\ \leq\ & m^2 \cdot \max_{1\leq i\leq j\leq m} P(|n\mathbf{e}_i'\mathbf{e}_j - \mathbf{y}_i'\mathbf{y}_j| \geq 17jt^2) \leq (64m^4)e^{-t^2/4050} \end{aligned} \tag{2.32}$$

as $0 < t \leq 3\sqrt{n}/5$, $1 \leq m \leq \sqrt{n}\,t/6$, and $mt/\sqrt{n} \leq 1/2$.

Step 2. Since the condition $mt/\sqrt{n} \le 1/2$ implies that $0 < t \le 3\sqrt{n}/5$, we then know that (2.32) holds only if $6m/\sqrt{n} \le t \le (1/2)\sqrt{n}/m$, which is valid by condition that $n \ge 12m^2$. Now, since $12m^2 \le n$, then $m \le \sqrt{n/12}$. Hence, if $0 < t < 6m/\sqrt{n}$, we have $0 < t \le 6/\sqrt{12} \le 3$. Notice $(64m^4)e^{-t^2/4050} \ge 64e^{-9/4050} > 64/e > 1$ for $t \in [0, 3]$, thus (2.32) holds only if $0 < t \le (1/2)\sqrt{n}/m$.

Step 3. Now, set $s = 17t^2$. Using condition that $m \le \sqrt{n/12}$, we have from (2.32)

$$P(\max_{1\le i,j\le m} |n\mathbf{e}_i'\mathbf{e}_j - \mathbf{y}_i'\mathbf{y}_j| \ge ms) \le \frac{64n^2}{144} \cdot \exp\left\{-\frac{s}{4050 \cdot 17}\right\} \tag{2.33}$$

as $0 < s = 17t^2 \le (17/4)(n/m^2)$. Now, take $v_{ij} = \mathbf{e}_i'\mathbf{e}_j$ and $K = (4050 \cdot 17)^{-1}$ to have

$$P(\max_{1\le i,j\le m} |nv_{ij} - \mathbf{y}_i'\mathbf{y}_j| \ge ms) \le K^{-1}n^2 \cdot \exp\{-Ks\} \tag{2.34}$$

for any $0 < s \le 4n/m^2$. Denote $t = ms$ and plug it into the above inequality, then

$$P(\max_{1\le i,j\le m} |nv_{ij} - \mathbf{y}_i'\mathbf{y}_j| \ge t) \le K^{-1}n^2 \cdot e^{-Kt/m}$$

for $0 < t \le 4n/m$. ■

Proof of Corollary 1.1. Theorem 1 says that

$$P\left(\max_{1\le i,j\le m_n} |\sqrt{n}v_{ij}^{(n)} - \frac{1}{\sqrt{n}}\mathbf{y}_{n,i}'\mathbf{y}_{n,j}| \ge \frac{t}{\sqrt{n}}\right) \le (K^{-1}n^2)e^{-Kt/m_n} \tag{2.35}$$

for any $0 < t \le 4n/m_n$. Let $0 < \delta_n \to 0$ be such that $m_n = (\sqrt{n}/\log n) \cdot \delta_n$ and $t_n = \sqrt{n\delta_n}$ for $n \ge e^e$. Evidently, $n \ge 12m_n^2$ as n is large. Further, one can see that

$$\begin{aligned} &\frac{t_n}{m_n} = \frac{\log n}{\sqrt{\delta_n}} \gg \log n, \quad \frac{t_n}{\sqrt{n}} = \sqrt{\delta_n} \to 0 \text{ and} \\ &0 < t_n = \sqrt{n\delta_n} < \frac{4\sqrt{n}\log n}{\delta_n} = \frac{4n}{m_n} \end{aligned} \tag{2.36}$$

as $n \to \infty$. Plugging $t = t_n$ and $m = m_n$ into (2.35), we see that the right hand side of (2.35) goes to 0 from the first assertion in (2.36). The conclusion follows. ■

Lemma 2.11 *The following hold:*

(i) As $n \to \infty$, both $\mathbf{y}_i'\mathbf{y}_j/\sqrt{n}$ and $\mathbf{y}_i^*\mathbf{y}_j/\sqrt{n}$ converge weakly to $\mathbb{C}N(0,1)$ for any $i \ne j$.

(ii) As $n \to \infty$, for fixed $i \ge 1$, $\mathbf{y}_i'\mathbf{y}_i/\sqrt{n}$ converges weakly to $\sqrt{2} \cdot \mathbb{C}N(0,1)$, and $(\mathbf{y}_i^*\mathbf{y}_i - n)/\sqrt{n}$ converges weakly to $N(0,1)$.

Proof. For each $n \ge 1$, by assumption, the random vectors $\mathbf{y}_1, \cdots, \mathbf{y}_n$ are i.i.d., thus it suffices to prove (i) and (ii) with $i = 1$ and $j = 2$.

(i) From the argument between (2.14) and (2.15), one see that $\mathbf{y}_1'\mathbf{y}_2/\|\mathbf{y}_1\| \sim \mathbb{C}N(0,1)$. Write $\|\mathbf{y}_1\|^2 = (1/2)\sum_{i=1}^{2n} \xi_i^2$ for some i.i.d. random variables $\{\xi_i;\, i \ge 1\}$ with $\xi_1 \sim N(0,1)$. By the law of large numbers, $\|\mathbf{y}_1\|/\sqrt{n} = \sqrt{\sum_{i=1}^{2n} \xi_i^2/2n}$ converges to one in probability. Then $\mathbf{y}_1'\mathbf{y}_2/\sqrt{n} = (\|\mathbf{y}_1\|/\sqrt{n}) \cdot (\mathbf{y}_1'\mathbf{y}_2/\|\mathbf{y}_1\|)$ goes to $\mathbb{C}N(0,1)$ weakly as $n \to \infty$ by using the Slusky lemma. Observe that the probability distributions of $\sqrt{n}\mathbf{y}_1^*\mathbf{y}_2$ and $\sqrt{n}\mathbf{y}_1'\mathbf{y}_2$ are the same, thus the second statement in (i) also follows.

(ii) By (2.18),

$$\frac{\mathbf{y}_1^*\mathbf{y}_1 - n}{\sqrt{n}} = \frac{\sum_{i=1}^n(\xi_i^2+\eta_i^2)-2n}{\sqrt{4n}} \quad \text{and} \tag{2.37}$$

$$\frac{\mathbf{y}_1'\mathbf{y}_1}{\sqrt{n}} = \sqrt{2}\cdot\left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right)\cdot\frac{1}{\sqrt{n}}\sum_{i=1}^n\begin{pmatrix}(\xi_i^2-\eta_i^2)/2\\ \xi_i\eta_i\end{pmatrix} \tag{2.38}$$

where $\{\xi_i;\, i\geq 1\}$ and $\{\eta_i;\, i\geq 1\}$ are two independent sequences of i.i.d. random variables with $\xi_1\sim N(0,1)$ and $\eta_1\sim N(0,1)$.

Note that $E(\xi_1^2+\eta_1^2)=2$, $\mathrm{Var}(\xi_1^2+\eta_1^2)=2\cdot\mathrm{Var}(\xi_1^2)=2(E(\xi_1^4)-1)=2(3-1)=4$. By the standard central limit theorems for i.i.d. random variables, we have from (2.37) that $(\mathbf{y}_1^*\mathbf{y}_1-n)/\sqrt{n}$ converges weakly to $N(0,1)$ as $n\to\infty$. So the second conclusion of (ii) follows.

Easily, $E(\xi_1\eta_1)=E(\xi_1^2-\eta_1^2)/2=0$. Now, by independence, $\mathrm{Var}(\xi_1\eta_1)=E(\xi_1\eta_1)^2=1$, and $\mathrm{Var}((\xi_1^2-\eta_1^2)/2)=\mathrm{Var}(\xi_1^2)/2=1$ since $\mathrm{Var}(\xi_1^2)=2$. Lastly, $\mathrm{Cov}(\xi_1\eta_1,(\xi_1^2-\eta_1^2)/2)=E\{\xi_1\eta_1(\xi_1^2-\eta_1^2)\}/2=0$. By the standard central limit theorem again,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\begin{pmatrix}(\xi_i^2-\eta_i^2)/2\\ \xi_i\eta_i\end{pmatrix}\text{ converges weakly to } N_2\left(\mathbf{0},\mathbf{I}_2\right).$$

By (2.38) and the Slusky lemma, we obtain the first conclusion of (ii). ■

Proof of Corollary 1.2. Obviously, for each $n\geq 1$, $\mathbf{V}_n=(v_{ij}^{(n)})$ and $\mathbf{V}$ in Theorem 1 have the same distribution. Thus, to prove the corollary, without loss of generality, we assume $\mathbf{V}_n=(v_{ij}^{(n)})=\mathbf{V}$. By Corollary 1.1, there exist i.i.d. random vectors $\mathbf{y}_{n,1},\cdots,\mathbf{y}_{n,n}$ with $\mathbf{y}_{n,1}\sim\mathbb{C}N_n(\mathbf{0},\mathbf{I}_n)$ such that

$$\max_{1\leq i,j\leq[\log n]}\left|\sqrt{n}v_{ij}^{(n)}-\frac{\mathbf{y}_{n,i}'\mathbf{y}_{n,j}}{\sqrt{n}}\right|\to 0$$

in probability as $n\to\infty$, where the notation $[\log n]$ is the integer part of $\log n$. The desired conclusions follow by using Lemma 2.11 and the Slusky lemma. ■

Acknowledgement. The author thanks Professor Peter Forrester for his question about the structure of the entries of the circular orthogonal ensembles; thanks professor Greg Anderson for the discussions about the circular ensembles; thanks Xue Ding and Danning Li for discussions on the proofs. Part of the research was done in the Institute for Mathematical Sciences, National University of Singapore. The author thanks the institute and director Louis Chen for their hospitality for providing excellent academic and living environments.

References

[1] Andersen, H. H., Hojbjerre, M., Sorensen, D., Eriksen, P. S. (1995). Linear and Graphical Models: for the Multivariate Complex Normal Distribution (Lecture Notes in Statistics). Springer-Verlag.

[2] Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis. John Wiley & Sons, Second edition.

[3] Dembo, A. and Zeitouni, O. (1998). Large Deviations Techniques and Applications. Springer, 2nd edition.

[4] D'Aristotle, A., Diaconis, P. and Newman, C. (2003). Brownian Motion and the Classical Groups, Probability, Statisitcs and their applications: Papers in Honor of Rabii Bhattacharaya. Edited by K. Athreya, et al. 97-116. Beechwood, OH: Institute of Mathematical Statistics (2003).

[5] Diaconis, P., Eaton, M. and Lauritzen, L. (1992). Finite deFinetti theorem in linear models and multivariate analysis. Scand. J. Statist. 19(4) 289-315.

[6] Dyson, F. J. (1962a). Statistical theory of energy levels of complex systems, I. J. Math. Phys. 3, 140-156.

[7] Dyson, F. J. (1962b). Statistical theory of energy levels of complex systems, II. J. Math. Phys. 3, 166-175.

[8] Dyson, F. J. (1962c). Statistical theory of energy levels of complex systems, III. J. Math. Phys. 3, 1191-1198.

[9] Jiang, T. (2008). The Entries of Haar-invariant Matrices from the Classical Compact Groups. Technical Report of School of Statistics, University of Minnesota.

[10] Jiang, T. (2009). Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles. Probability Theory and Related Fields, 144(1), 221-246.

[11] Jiang, T. (2006). How many entries of a typical orthogonal matrix can be approximated by independent normals? Ann. Probab. 34(4) 1497-1529.

[12] Jiang, T. (2005). Maxima of Entries of Haar Distributed Matrices, Probability Theory and Related Fields, 131 121-144.

[13] Jiang, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices, Ann. of Applied Probab. 14(2) 865-880.

[14] Li, D. and Rosalsky, A. (2006). Some strong limit theorems for the largest entries of sample correlation matrices. Ann. Appl. Probab. 16, 423-447.

[15] Li, D., Liu, W. and Rosalsky, A. (2008). Necessary and sufficient conditions for the asymptotic distributions of the largest entry of a sample correlation matrix. To appear in *Probability Theory and Related Fields.*

[16] Liu, W. D., Lin, Z. and Shao, Q. M. (2008). The asymptotic distribution and the Berry-Esseen bound of a new test for independence in high dimension with application to stochastic optimazation. To appear in *Ann. Appl. Probab.*; available at http://wwww.imstat.org/aap/future_papers.html.

[17] Mehta, M. L. (1991). Random Matrices, 2nd edition, Academic Press, Boston.

[18] Mezzadri, F. (2007). How to generate random matrices from the classical compact groups. Notices to the AMS, Vol. 54(5) 592-604.

[19] Reichl, L. E. (2004). The Transition to Chaos: Conservative Classical Systems and Quantum Manifestations, 2nd edition, Springer.

[20] Zhou, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. Trans. Amer. Math. Soc. 359 5345-5363.