On the Universality of the Non-singularity of General
Ginibre and Wigner Random Matrices

Paulo Manrique*  Victor Pérez-Abreu †  Rahul Roy ‡

April 23, 2014

Abstract

We prove the universal asymptotically almost sure non-singularity of general Ginibre and
Wigner ensembles of random matrices when the distribution of the entries are independent
but not necessarily identically distributed and may depend on the size of the matrix. These
models include adjacency matrices of random graphs and also sparse, generalized, universal and
banded random matrices. We find universal rates of convergence and precise estimates for the
probability of singularity which depend only on the size of the biggest jump of the distribution
functions governing the entries of the matrix and not on the range of values of the random
entries. Moreover, no moment assumptions are made about the distributions governing the
entries. Our proofs are based on a concentration function inequality due to Kesten and allows
us to improve universal rates of convergence for the Wigner case when the distribution of the
entries do not depend on the size of the matrix.

Key terms: Adjacency matrix of random graphs, banded random matrix, decoupling, con-
centration function, generalized Wigner ensemble, Littlewood-Offord inequality, Kolmogorov-
Rogozin inequality, nondegenerate distribution, sparse random matrix.

1 Introduction and main results

Let $A_n = (\xi_{ij}^{(n)})$ be an $n \times n$ random matrix where each entry $\xi_{ij}^{(n)}$ follows a distribution $F_{ij}^{(n)}$, $1 \leq i, j \leq n$. The study of the non-singularity of such matrices has mainly been considered when $F_{ij}^{(n)} \equiv F$ and for two ensembles of random matrices, the Ginibre and Wigner. We will use the
following terminology: An $n \times n$ random matrix $G_n = (\xi_{ij})_{1 \leq i, j \leq n}$ is called a Ginibre matrix if $\xi_{ij}$, $i, j = 1, \ldots, n$ are independent random variables, and an $n \times n$ random symmetric matrix $W_n = (\xi_{ij})_{1 \leq i, j \leq n}$ is called Wigner matrix if $\xi_{ij} = \xi_{ji}$, $i, j = 1, \ldots, n$ and $\xi_{ij}$, $1 \leq i \leq j \leq n$ are independent random variables. We will not assume that the distributions of the entries have moments.

The singularity of these matrices is trivial if the distributions of $\xi_{ij}$ are degenerate. The non-singularity is also straightforward if the entries have continuous distributions. The interesting situation occurs when some of the entries have distributions with jumps. Singularity of such matrices being a highly non trivial problem.

The study of the non-singularity of Ginibre matrices goes back to the pioneering work by J. Komlós. In [15] he considers a Ginibre random matrices $GB(n, 1/2)$, entries of which are i.i.d. Bernoulli random variables, taking values 0 or 1 with probability 1/2 each. Using a very clever ‘growing rank analysis’ together with the Littlewood-Offord inequality Komlós proved that $\mathbb{P}\{\text{rank}(GB(n, 1/2)) < n\} = o(1)$ as $n \to \infty$. Bollobás [3] presents the concept of ‘strong rank’ and together with the Littlewood-Offord inequality obtains an unpublished result due to Komlós, viz. $\mathbb{P}\{\text{rank}(GB(n, 1/2)) < n\} = O(n^{-1/2})$ as $n \to \infty$. Komlós [16] was also the first in considering the singularity of Ginibre matrices whose entries are i.i.d. random variables with a common arbitrary non-degenerate distribution, proving that the probability that such an $n \times n$ matrix is singular has order $o(1)$ as $n \to \infty$. This result was improved by Kahn, Komlós and Szemerédi [12] in the case of Ginibre matrices whose entries are i.i.d. taking values $-1$ or $1$ with probability 1/2 each, showing that the probability of singularity is bounded above by $\theta^n$ for $\theta = .999$. The value of $\theta$ has been improved by Tao and Vu [24, 25] to $\theta = 3/4 + o(1)$ and by Bourgain, Vu and Wood [5] to $\theta = 1/\sqrt{2} + o(1)$. Slinko [21] considered Ginibre random matrices whose entries have the same distribution taking values in a finite set, proving also that the probability of singularity is $O(n^{-1/2})$ as $n \to \infty$.

The aim of this paper is to understand the asymptotic non-singularity of more general Ginibre and Wigner ensembles. We are interested in finding universality results with respect to general distributions of the entries and also when these distributions depend of the size of the matrix.

As a first step in this direction, the results in [3, 21] was generalized by Bruneau and Germinet [4] to Ginibre random matrices whose entries follow different independent non-degenerate distributions $F_{ij}$ which do not change with the size of the matrix. Their result gives a universal rate of convergence of $n^{-1/2}$ as follows:
**Theorem 1.** Let $G_n$ be an $n \times n$ Ginibre matrix with independent entries $\xi_{ij}$ satisfying the following property $H$: there exists $\rho \in (0, 1/2)$ such that for any $i, j = 1, \ldots, n$, $\mathbb{P}\{\xi_{ij} > x_{ij}^+\} > \rho$ and $\mathbb{P}\{\xi_{ij} < x_{ij}^-\} > \rho$ for some real numbers $x_{ij}^- < x_{ij}^+$, then

$$\mathbb{P}\{\text{rank}(G_n) < n\} \leq C/\sqrt{\beta_\rho(1 - \rho)n},$$

where the constant $C$ is universal (coming from the Littlewood-Offord inequality) and $\beta_\rho$ is an implicit constant $0 < \beta_\rho < 1$ which goes to zero as $\rho \to 1$.

**Remark 1.**

a) The above theorem is proved in [4] using ideas of strong rank of [3], together with a Bernoulli representation theorem for the distribution of a random variable and the Littlewood-Offord inequality. The value of $\rho$ is not unique.

b) We point out that it is possible to express (1) in terms of the size of the biggest jump of the distribution functions governing the entries. Indeed this follows using a strong rank analysis and a sharper version of the Kolmogorov-Rogozin concentration inequality due to Kesten [14]. This inequality, stated in Section 2, will be used repeatedly in this work. Returning to (1), taking $\kappa$ to be the size of the biggest jump of $F_{ij}$, $i, j = 1, \ldots, n$, we have

$$\mathbb{P}\{\text{rank}(G_n) < n\} = C_1/\sqrt{\beta'_\kappa \Delta(1 - \kappa \Delta)n}$$

where $\kappa \Delta = \kappa + \Delta < 1$ for any $0 < \Delta < 1$ and the constant $C_1$ is universal (coming from Kesten’s concentration inequality) and $\beta'_\kappa \Delta$ is an implicit constant $0 < \beta'_\kappa \Delta < 1$ which goes to zero as $\kappa \to 1$.

c) We observe that the constants $\beta_\rho$ and $\beta'_\kappa \Delta$ are not universal—they depend on the distributions $F_{ij}$.

d) These two results highlight the fact that the non-singularity of Ginibre matrix depends only on $\rho$ or, equivalently, the size of the biggest jump $\kappa$. In other words, the universal property of a random matrix being non-singular depends neither on the range of values taken by the entries nor on other properties of their distribution except the size of the biggest jump.

As for Wigner random matrices, the study of their singularity was initiated by Costello, Tao and Vu [8] inspired by the work of Komlós [15].

**Theorem 2.** Let $W_n = (\xi_{ij})$ be an $n \times n$ Wigner matrix whose upper diagonal entries $\xi_{ij}$ are independent random variables with common Bernoulli distribution on $\{0, 1\}$ with parameter $1/2$. 
Then
\[ \mathbb{P} \{ \text{rank}(W_n) < n \} = O(n^{-1/8+\alpha}), \]

for any positive constant \( \alpha \), the implicit constant in \( O(\cdot) \) depending on \( \alpha \).

**Remark 2.**

a) The proof of the above theorem in [8] required developing a quadratic Littlewood-Offord inequality. A possible generalization to distributions other than Bernoulli was also indicated in [8].

b) The Corollary 2 to Theorem 5 below gives a better universal rate of convergence \( n^{-1/4+\alpha} \), for any Wigner random matrix \( W_n = (\xi_{ij}) \) with independent entries which need not be identical. While the off-diagonal entries need to be non-degenerate, the diagonal entries could be degenerate.

More recently, the Wigner matrix has been studied when the entries satisfy some restrictions. Nguyen [17] considered a Wigner matrix \( W_n \) with entries taking values \(-1\) or \(1\) with probability \(1/2\) each, subject to the condition that each row has exactly \(n/2\) entries which are zero. He showed that the probability of \( W_n \) being singular is \( O(n^{-C}) \), for any positive constant \( C \), the implicit constant in \( O(\cdot) \) depending on \( C \). Also recently, Vershynin [26] has considered the case of a Wigner matrix \( W_n \) whose entries satisfy the following property: the above-diagonal entries are independent and identically distributed with zero mean, unit variance and subgaussian, while the diagonal entries satisfy \( \xi_{ii} \leq K \sqrt{n} \) for some \( K \). He showed that the probability of \( W_n \) being singular is bounded above by \( 2 \exp(-n^c) \), where \( c \) depends only on the subgaussian distribution and on \( K \).

One of the goals of this paper is to study the non-singularity of Ginibre and Wigner matrices when the distributions of the entries \( F_{ij}^{(n)} \) depend on the matrix size. This kind of random matrices appear in the study of random graphs [6], sparse matrices [7], [9] and some other models that have recently been extensively considered like the so-called generalized, universal and banded Wigner ensembles [10], [22] among other works. See also the non i.i.d. Wigner case in, for example, [2, pp 26].

One difficulty that arises in this situation is to find adequate asymptotic estimates of the probability of the singularity being zero where the involved constants in the rate of convergence do not depend on the distributions of the entries. We overcome this difficulty using a universal concentration inequality due to Kesten [14] which we express in terms of the size of the jumps of the distribution functions.
1.1 Main results

We now consider Ginibre and Wigner matrix ensembles \( G_n = \left( \xi_{ij}^{(n)} \right)_{1 \leq i,j \leq n} \) and \( W_n = \left( \xi_{ij}^{(n)} \right)_{1 \leq i,j \leq n} \), where the distribution function \( F_{ij}^{(n)} \) governing \( \xi_{ij}^{(n)} \) is allowed to change with the size of the matrix.

One of our main conclusions is the non-singularity of the above Ginibre and Wigner ensembles. More specifically, given a collection of non-degenerate distribution functions \( \left\{ F_{ij}^{(n)} : i, j \geq 1, n \geq 1 \right\} \) and a subsequence \( \left\{ m_n : n \geq 1 \right\} \) we study the singularity of the \( m_n \times m_n \) matrix with independent entries \( \xi_{ij}^{(n)} \) being governed by the distribution function \( F_{ij}^{(n)} \) for every \( 1 \leq k, l \leq m_n \).

**Theorem 3.** *(Universality of non-singularity of Ginibre and Wigner ensembles)* With the notation as above, let \( G_r^{(n)} \) and \( W_r^{(n)} \) be the \( r \times r \) Ginibre and Wigner matrices respectively, each with entries \( \xi_{ij}^{(n)} \), \( 1 \leq i \leq j \leq r \). Given any \( \varepsilon \in (0, 1) \), there exists an increasing sequence \( \{ m_n \} \) such that

\[
P \left\{ \text{rank}(G_{mn}^{(n)}) < m_n \right\} = O \left( m_n^{-(1-\varepsilon)/2} \right),
\]

(3)

\[
P \left\{ \text{rank}(W_{mn}^{(n)}) < m_n \right\} = O \left( m_n^{-(1-\varepsilon)/4} \right),
\]

(4)

where the implicit constant in \( O(\cdot) \) depends on \( \varepsilon \) and the size \( \kappa_n \) of the biggest jump of the distribution functions \( F_{ij}^{(n)} \), \( 1 \leq i, j \leq m_n \).

We will deduce Theorem 3 from the following two main technical results of this paper, which provide useful universal constants in the approximation of probability of singularity.

**Theorem 4.** Let \( G_n = (\xi_{ij}) \) be an \( n \times n \) Ginibre matrix with independent entries \( \xi_{ij} \) \( i \neq j \), having non-degenerate distributions \( F_{ij} \) and the diagonal entries \( \xi_{ii} \) being allowed to have degenerate distributions. Let \( \kappa \in (0, 1) \) be the biggest jump of the distribution functions \( F_{ij} \), \( i, j = 1, \ldots, n \). Then, for any \( \varepsilon \in (0, 1) \)

\[
P \{ \text{rank}(G_{n+1}) < n + 1 \} = O \varepsilon \left( \kappa^{(n-n^1-\varepsilon)/2} + \left[ \kappa^2 \Delta_n^{1-\varepsilon(1-\kappa)} \right]^{1/2} \right),
\]

(5)

where \( \kappa_\Delta := \kappa + \Delta < 1 \) with \( \Delta > 0 \) fixed and as small as desired, and the implicit constant in \( O(\cdot) \) depending only on \( \varepsilon \in (0, 1) \).

**Theorem 5.** Let \( W_n = (\xi_{ij}) \) be an \( n \times n \) Wigner matrix with independent entries \( \xi_{ij} \), \( i \leq j \) having non-degenerated distribution \( F_{ij} \) and the diagonal entries \( \xi_{ii} \) being allowed to have degenerate distributions. Let \( \kappa \in (0, 1) \) be the biggest jump of the distribution functions of \( F_{ij} \), \( i, j = 1, \ldots, n \).
Then, for any $\varepsilon \in (0, 1)$

$$\mathbb{P}\{\text{rank}(W_n) < n\} = O_\varepsilon\left(\frac{\kappa_n^{3/2}n^{-1/2}n^{1-\varepsilon}}{\kappa(1 - \kappa)} + \left[\frac{\kappa^2}{n^{1-\varepsilon}(1 - \kappa)}\right]^{1/4}\right), \quad (6)$$

where $\kappa := \kappa + \Delta < 1$ with $\Delta > 0$ fixed, which is a small as desired, and the implicit constant in $O(\cdot)$ depending only on $\varepsilon \in (0, 1)$.

As a consequence of the last two theorems we have

**Corollary 1.** Under the assumptions of Theorem 4,

$$\mathbb{P}\{\text{rank}(G_n) < n\} \leq C(\kappa, \varepsilon)n^{-(1-\varepsilon)/2} + o(1),$$

where $C(\kappa, \varepsilon)$ is a constant that depends on $\varepsilon > 0$ and $\kappa \in (0, 1)$.

This corollary does not give a better rate of convergence as the universal rate $n^{-1/2}$ in Theorem 1. This is due to the use of growing rank analysis instead of the strong rank analysis. However, (5) allows us to handle the situation when the distributions of the entries of the matrix depend on the size of the matrix as we will see later on. However, for the Wigner case we have an interesting universal rate of convergence.

**Corollary 2.** Under the assumptions of Theorem 5

$$\mathbb{P}\{\text{rank}(W_n) < n\} \leq C(\kappa, \varepsilon)n^{-(1-\varepsilon)/4} + o(1), \quad (7)$$

where $C(\kappa, \varepsilon)$ is a constant that depends on $\varepsilon > 0$ and $\kappa \in (0, 1)$.

The bound $n^{-1/4+\alpha}$ in (7) improves the rate $n^{-1/8+\alpha}$ in Theorem 2 of Costello, Tao and Vu [8]. Moreover, from (7) and analogous to the Ginibre case we conjecture that for the Wigner case the universal rate of convergence is $n^{-1/4}$.

For the proof of Theorem 4 we follow the growing rank analysis of Komlós [15]. For the proof of Theorem 5 we follow ideas in [8]. However in both these proofs we use the improvement by Kesten [14] of the Kolmogorov-Rogozin concentration inequality.

We now turn to Theorem 3. A natural question is whether one always has the case $m_n = n$. An example where this is not true is the following. Let $GB(n, p)$ and $WB(n, p)$ denote the $n \times n$ Ginibre and Wigner matrices whose entries have Bernoulli distribution on $\{0, 1\}$ with parameter
Let \( ZGB_n (ZWB_n) \) be the event that the first row of \( GB(n, 1/n), (WB(n, 1/n)) \) contains only zeros, then
\[
P \{ ZGB_n \} = \left( 1 - \frac{1}{n} \right)^n, \quad P \{ ZW B_n \} = \left( 1 - \frac{1}{n} \right)^n,
\]
hence
\[
e^{-1} \leq \lim_{n \to \infty} P \{ \text{rank} (GB (n, 1/n)) < n \},
\]
\[
e^{-1} \leq \lim_{n \to \infty} P \{ \text{rank} (WB (n, 1/n)) < n \},
\]
but
\[
P \{ \text{rank} (GB (n^2, 1/n)) < n^2 \} \leq C \left( n^{-1/4} \right) + o(1), \tag{8}
\]
\[
P \{ \text{rank} (WB (n^2, 1/n)) < n^2 \} \leq C \left( n^{-1/5} \right) + o(1). \tag{9}
\]
However, there are examples where the maximum jump converges to one (and therefore the corresponding distributions converge to a degenerate distribution) and \( m_n = n \): if \( \alpha \in (0, 1) \)
\[
P \{ \text{rank} (GB (n, n^\alpha /n)) < n \} \leq C \left( n^{-\alpha/4} \right) + o(1), \tag{10}
\]
\[
P \{ \text{rank} (WB (n, n^\alpha /n)) < n \} \leq C \left( n^{-\alpha/5} \right) + o(1), \tag{11}
\]
where \( C \) is a universal constant in the Kolmogorov-Rogozin concentration inequality.

Furthermore, as an application of the Wigner case, Theorem 5 we obtain an estimation of the probability that the adjacency matrix of a sparse random graph (not necessarily an Erdös-Rényi graph) is non-singular. Costello and Vu \[6\] have analyzed the adjacency matrices of sparse Erdös-Rényi graphs, where each entry is equal to 1 with the same probability \( p(n) \) which tends to 0 as \( n \) goes to infinity (see also Costello and Vu \[7\] where a generalization of \[6\] is considered in which each entry takes the value \( c \in \mathbb{C} \) with probability \( p \) and zero with probability \( 1 - p \), and the diagonal entries are possibly non-zero). It is proved in \[6\] that when \( c \ln(n)/n \leq p(n) \leq 1/2, c > 1/2, \) then with probability \( 1-O((\ln \ln(n))^{-1/4}) \), the rank of the adjacency matrix equals the number of non-isolated vertices. In particular, if \( p(n) = c \ln(n)/n, c < 1/2, \) then the adjacency matrix is singular. From Theorem 5 we observe that if the size of the adjacency matrix is \( n^2 \times n^2 \), then it is non-singular with high probability. Moreover, we have that the adjacency matrix is singular for more models of random graphs. In particular, we consider the following model extension of Erdös-Rényi graphs, where vertices \( i \) and \( j \) are linked with a probability that depends on \( i \) and \( j \).
and the number of vertices. Furthermore, the rate of convergence is an improvement of the one given in [9].

**Proposition 1.** Let \( \{p_{ij} \in (0, 1), 1, 2, \ldots\} \) be a double sequence of positive numbers, then there is a random graph with \( n \) vertices such that the vertex \( i \) is linked with the vertex \( j \) with probability \( p_{ij} \), \( 1 \leq i < j \leq n \), and if \( A_n \) is the adjacency matrix, then

\[
P \{ \text{rank}(A_n) < n \} \leq C(p_n^*) n^{-(1-\varepsilon)/4} + o(1),
\]

where \( C(p_n^*) \) is a constant depending on \( p_n^* = \max_{1 \leq i \leq j \leq n} \{p_{ij}\} \) and \( \varepsilon \in (0, 1) \).

The above examples correspond to the case when the distribution of each entry converges to a degenerate distribution as the matrix size goes to infinity. For (10), (11) the corresponding ensemble of random matrices is asymptotically almost surely non-singular when \( m_n = n \), but in (8), (9) and possibly for (12) this is not the case.

On the other hand, convergence of the size of the maximum jumps in \((0, 1)\) also ensures that \( m_n = n \) and therefore non-singularity. Let \( G^{(n)}_n \) and \( W^{(n)}_n \) be as in Theorem 3 and let \( \kappa_n = \max_{1 \leq i, j \leq n} \{\sup \{F_{ij}^{(n)}(x^+) - F_{ij}^{(n)}(x) : x \in \mathbb{R}\}\} \). Suppose \( \kappa_n \to \kappa \) as \( n \to \infty \) for some \( 0 < \kappa < 1 \), then, as the following result shows, both \( G^{(n)}_n \) and \( W^{(n)}_n \) are asymptotically almost surely non-singular.

**Proposition 2.** With the notation as above, we have

\( i) \)

\[
P \{ \text{rank}(G^{(n)}_n) < n \} = O_{\varepsilon} \left( n^{-\frac{\varepsilon}{2}} + \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)} \right]^{1/2} \right),
\]

where \( \kappa_\Delta := \kappa + \Delta < 1, \Delta \in (0, 1) \) and the implicit constant in \( O(\cdot) \) depending only on \( \varepsilon \in (0, 1) \).

\( ii) \)

\[
P \{ \text{rank}(W^{(n)}_n) < n \} = O_{\varepsilon} \left( \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)} + \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)} \right]^{1/4} \right),
\]

where \( \kappa_\Delta := \kappa + \Delta < 1, \Delta \in (0, 1) \), and the implicit constant in \( O(\cdot) \) depending only on \( \varepsilon \in (0, 1) \).

**Remark 3.** The second summands in the expression (13) and (14) determine the rate of convergence to non-singularity, while the first summands are necessary to handle the situation when the distribution of entries changes with matrix size. This produces constants in \( O_{\varepsilon} \) independent of the distribution.
Remark 4. a) In many applications of random matrices one considers ensembles of the form $G_n^{(n)} = a_n^{-1} G_n$ and $W_n^{(n)} = a_n^{-1} W_n$ where $a_n \to \infty$ as $n \to \infty$ and the non-degenerate distributions of the entries of $G_n$ and $W_n$ do not depend on matrix size $n$. In this case $\kappa_n = \kappa$ for all $n \geq 1$ and the distribution of each entry converges to zero almost surely. However the ensembles $G_n^{(n)}$ and $W_n^{(n)}$ are asymptotically almost surely non-singular. In fact, this holds for any sequence $a_n \to \infty$ and the rate of convergence to zero of the probability of singularity is not affected by the rate of convergence of $a_n$ if the distributions of the entries have discrete support without any accumulation point.

b) The case $a_n = 1/\sqrt{n}$ is the set-up of problems of random matrices appearing in the study of the asymptotic spectral distributions [1], [2], geometric functional analysis [23], [18] and restricted isometries [20], among others.

c) Finally, the results in the Ginibre case have a straightforward extension to non-square $n \times m$ random matrices whose entries are independent random variables and have distributions with jumps.

2 Preliminaries on Concentration Inequalities

In this section we present the Kolmogorov-Rogozin concentration inequalities that we use for the proofs of our main results on non-singularity. We express these inequalities in terms of the size of biggest jump of the non-degenerate distribution functions.

The Lévy concentration function $Q(\xi; \lambda)$ of a random variable $\xi$ is defined by

$$Q(\xi; \lambda) = \sup_{x \in \mathbb{R}} \mathbb{P}\{\xi \in [x, x + \lambda]\}, \quad \lambda > 0.$$ 

Let $\xi_1, \xi_2, \ldots.$ be independent random variables and $S_n = \sum_{i=1}^{n} \xi_i$. An expression that relates the concentration function of $S_n$ to the concentration functions of the summands $\xi_i$ was given by Kolmogorov-Rogozin; see [13].

**Lemma 1** (Kolmogorov-Rogozin Inequality). There exists a universal constant $C$ such that for any independent random variables $\xi_1, \ldots, \xi_n$ and any real numbers $0 < \lambda_1, \ldots, \lambda_n \leq L$, one has

$$Q(S_n; L) \leq CL \left\{ \sum_{i=1}^{n} \lambda_i^2 [1 - Q(X_i; \lambda_i)] \right\}^{-1/2}.$$ 

Kesten [14] obtained the following refinement of the above inequality.
Lemma 2. For the constant $C$ of the Kolmogorov-Rogozin inequality and any independent random variables $\xi_1, \ldots, \xi_n$, and real numbers $0 < \lambda_1, \ldots, \lambda_n \leq 2L$, one has

\[
Q(S_n; L) \leq 4 \cdot 2^{1/2}(1 + 9C)L \sum_{i=1}^{n} \lambda_i^2 \left[1 - Q(\xi_i; \lambda_i) \right] \frac{Q(\xi_i; L)}{\{\sum_{i=1}^{n} \lambda_i^2 \left[1 - Q(\xi_i; \lambda_i) \right] \}^{3/2}}.
\]

For the study of non-singularity of random matrices, one has to find an estimate of the probability that a polynomial of independent random variables equals a real number. In the case of Ginibre and Wigner matrices the polynomial are of degree one and two respectively. Our first goal is to write Kesten inequality in terms of the size of the biggest jump and then obtain the corresponding linear and quadratic concentration inequalities.

We first discuss the relation between the size of the biggest jump of a non-degenerate distribution $F$ and its corresponding Lévy concentration function. Let $D_F$ be the set of discontinuities of $F$ and $\kappa$ its biggest jump, i.e., $\kappa = \sup_{x \in \mathbb{R}} \{ \xi = x \}$ where $\xi$ has distribution function $F$.

We note the following:

1. There exists $x_\kappa \in \mathbb{R}$ such that $\mathbb{P} \{ \xi = x_\kappa \} = \kappa$.

2. Let $p_i = \mathbb{P} \{ \xi = x_i \}$, $i \in \mathbb{N}$, then $\sum_{i \geq 1} p_i \leq 1$, i.e., for all $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that $\sum_{i \geq n} p_i \leq \varepsilon$ for all $n \geq N(\varepsilon)$.

3. If $F$ is a discrete distribution (\sum_{i \in \mathbb{N}} p_i = 1) and $x_\kappa$ is not an accumulation point of $D_F$, there exists $\delta_1 > 0$ with

\[
\sup_{x \in \mathbb{R}} \mathbb{P} \{ \xi \in [x, x + \delta_1] \} = \kappa.
\]

Otherwise, if $F$ is not discrete or $x_\kappa$ is an accumulation point of $D_F$, there exists some $\Delta > 0$, which may be taken as small as desired, such that, for $\Delta$ fixed, there is $\delta_2 > 0$ with

\[
\sup_{x \in \mathbb{R}} \mathbb{P} \{ \xi \in [x, x + \delta_2] \} = \kappa + \Delta < 1.
\]

We defined $\kappa_\Delta$, for $\Delta \in (0, 1)$ fixed, by $\kappa_\Delta := \kappa$ if $F$ is discrete and $x_\kappa$ is not an accumulation point of $D_F$ and otherwise, $\kappa_\Delta := \kappa + \Delta$. So, we have that there is $\delta > 0$ such that

\[
\sup_{x \in \mathbb{R}} \mathbb{P} \{ \xi \in [x, x + \delta] \} = \kappa_\Delta. \tag{15}
\]
4. We fix $\Delta \in (0, 1)$ and $\delta > 0$ that satisfies \ref{15}. If $a \in \mathbb{R}$ with $|a| \geq 1$, then

$$\sup_{x \in \mathbb{R}} \mathbb{P}\{a\xi \in [x, x + \delta]\} \leq \kappa_\Delta.$$ 

Indeed, if $\sup_{x \in \mathbb{R}} \mathbb{P}\{a\xi \in [x, x + \delta]\} > \kappa_\Delta$, then there exists some $x^* \in \mathbb{R}$ such that $\mathbb{P}\{a\xi \in [x^*, x^* + \delta]\} > \kappa_\Delta$, but

$$\delta \geq |a\xi - x^*| = |a| \left| \xi - \frac{x^*}{a} \right| \geq \left| \xi - \frac{x^*}{a} \right|,$$

which is a contradiction to the definition of $\kappa_\Delta$. So, we have that

$$Q(\beta \xi, \delta) \leq \kappa_\Delta$$

for $|\beta| \geq 1$.

Now let $\xi_1, \ldots, \xi_n$ be independent random variables with distribution functions $F_1, \ldots, F_n$, respectively. Let $\Delta > 0$ be fixed, for each $\xi_i$ we consider $\kappa(i)$ and $\kappa_\Delta(i)$ defined as above. We first prove the following concentration inequality in terms of the biggest jumps of the distribution functions.

**Lemma 3** (Linear Concentration Inequality). Let $\xi_1, \ldots, \xi_n$ be independent random variables with non-degenerate distributions $F_1, \ldots, F_n$, respectively, and let $\alpha_1, \ldots, \alpha_n$ be real numbers with $\alpha_i \neq 0$, $i = 1, \ldots, n$. Then

$$\sup_{x \in \mathbb{R}} \mathbb{P}\left\{ \sum_{i=1}^n \alpha_i \xi_i = x \right\} = O\left( \frac{\sum_{i=1}^n (1 - \kappa(i)) \kappa_\Delta(i)}{\left\{ \sum_{i=1}^n [1 - \kappa_\Delta(i)] \right\}^{3/2}} \right),$$

where the implicit constant in $O(\cdot)$ does not depend on $F_i$, $i = 1, \ldots, n$.

**Proof.** Let $a = \min_{1 \leq i \leq n} \{|\alpha_i|\}$ and $\delta = \min_{1 \leq i \leq n} \{\delta_i\}$, where $\delta_i > 0$ satisfies $\kappa_\Delta(i) = Q(\xi_i, \delta_i)$, $i = 1, \ldots, n$. We have for $x \in \mathbb{R}$

$$\mathbb{P}\left\{ \sum_{i=1}^n \alpha_i \xi_i = x \right\} = \mathbb{P}\left\{ \sum_{i=1}^n \frac{\alpha_i}{a} \xi_i = \frac{x}{a} \right\} = \mathbb{P}\left\{ \sum_{i=1}^n \alpha'_i \xi_i = x' \right\},$$

where $\alpha_i/a = \alpha'_i$ and $x/a = x'$. Now,

$$\mathbb{P}\left\{ \sum_{i=1}^n \alpha'_i \xi_i = x' \right\} \leq \sup_{y \in \mathbb{R}} \mathbb{P}\left\{ \sum_{i=1}^n \alpha'_i \xi_i \in [y, y + \delta]\right\} \leq 4 \cdot 2^{1/2} (1 + 9C) \frac{\sum_{i=1}^n (1 - \kappa(i)) \kappa_\Delta(i)}{\left\{ \sum_{i=1}^n [1 - \kappa_\Delta(i)] \right\}^{3/2}},$$

11
Remark 5. a) If $\xi_1, \ldots, \xi_n$ are iid random variables,

$$\sup_{x \in \mathbb{R}} \mathbb{P} \left\{ \sum_{i=1}^{n} \alpha_i \xi_i = x \right\} = O \left( \left[ \frac{(1-\kappa)^2 \kappa^2}{n(1-\kappa)^2} \right]^{1/2} \right),$$

and if they are discrete random variables and $x_\kappa$ is not an accumulation point of $D_F$,

$$\sup_{x \in \mathbb{R}} \mathbb{P} \left\{ \sum_{i=1}^{n} \alpha_i \xi_i = x \right\} = O \left( \left[ \frac{\kappa^2}{n(1-\kappa)} \right]^{1/2} \right).$$

b) Lemma 3 holds when $r$ many of the random variables $\xi_1, \ldots, \xi_n$ are degenerate for some $1 \leq r < n$. Contribution to the bound of the concentration inequality is provided only by the non-degenerated random variables.

In order to prove the so-called Quadratic Concentration Inequality, we recall the decoupling argument.

Lemma 4 (Decoupling). Let $X \in \mathbb{R}^{m_1}$ and $Y \in \mathbb{R}^{m_2}$ be independent random variables, with $m_1 + m_2 = n$, and let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a Borel function. Let $X'$ be a variable independent of $X$ and $Y$, but having the same distribution as $X$. For any interval $I$ of $\mathbb{R}$, we have

$$\mathbb{P}^2 \{ \varphi(X,Y) \in I \} \leq \mathbb{P} \{ \varphi(X,Y) \in I, \varphi(X',Y) \in I \}.$$

A quadratic Littlewood-Offord inequality for independent $\{0,1\}$-Bernoulli random variables with probability $1/2$ was proved in [8]. The result below is for independent random variables not necessarily identically distributed and without any assumption on their moments.

Lemma 5 (Quadratic Concentration Inequality). Let $\xi_1, \ldots, \xi_n$ be independent random variables with non-degenerate distributions $F_1, \ldots, F_n$, respectively. Let $(c_{ij})_{1 \leq i,j \leq n}$ be a symmetric $n \times n$ array of constants. Suppose that there is a partition of $\{1,2,\ldots,n\} = S_1 \sqcup S_2$ such that for each $j \in S_2$ there are at least $r$ different $i \in S_1$ for which $c_{ij} \neq 0$. Let

$$\varphi = \varphi \{ \xi_1, \ldots, \xi_n \} = \sum_{1 \leq i,j \leq n} c_{ij} \xi_i \xi_j$$
be the quadratic form whose coefficients are the $c_{ij}$. Then

$$
\sup_{x \in \mathbb{R}} \mathbb{P} \{ \varphi = x \} = O \left( \left[ \frac{\sum_{i=1}^{r}(1 - \pi(l_i))\kappa_{\Delta}(L_i)}{\sum_{i=1}^{s} |\kappa_{\Delta}(L_i)|} \right]^{3/2} + \frac{\sum_{i=1}^{s} |\kappa_{\Delta}(L_i)|}{\sum_{i=1}^{s} |\kappa_{\Delta}(L_i)|} \right)^{1/2}
$$

where $\pi(l_i)$ and $\kappa_{\Delta}(L_i)$ are the jumps associated with $\xi_i - \xi'_i$, $l_i \in \{1, \ldots, |S_1|\}$, $\kappa(l_i)$ and $\kappa_{\Delta}(L_i)$ are the jumps associated with $\xi_i$, $i \in \{|S_1| + 1, \ldots, n\}$. The implicit constant in $O(\cdot)$ does not depend on $F_i$, $i = 1, \ldots, n$.

**Proof.** Let $\delta = \min_{1 \leq i \leq n} \{\delta_i\}$ where $\delta_i > 0$ satisfies $\kappa_{\Delta}(i) = Q(\xi_i, \delta_i)$, $i = 1, \ldots, n$. If $x \in \mathbb{R}$, we have

$$
\mathbb{P} \{ \varphi = x \} \leq \mathbb{P} \{ \varphi \in [x, x + \delta/2] \}.
$$

Write $I = [x, x + \delta/2]$, $X = (\xi_1, \ldots, \xi_{|S_1|})$, $Y = (\xi_{|S_1|+1}, \ldots, \xi_n)$ and $X' = (\xi'_1, \ldots, \xi'_{|S_1|})$, with $X'$ independent of $X$ and $Y$, but having the same distribution as $X$. By Lemma 4

$$
\mathbb{P}^2 \{ \varphi(X,Y) \in I \} = \mathbb{P} \{ \varphi(X,Y) \in I, \varphi(X',Y) \in I \} \\
\leq \mathbb{P} \{ \varphi(X,Y) - \varphi(X',Y) \in [-\delta/2, \delta/2] \}.
$$

We can rewrite $\varphi(X,Y) - \varphi(X',Y)$ as

$$
\varphi(X,Y) - \varphi(X',Y) = g(X, X') + 2 \sum_{j=|S_1|+1}^{n} \xi_j \left( \sum_{i=1}^{|S_1|} c_{ij} (\xi_i - \xi'_i) \right) \\
= g(X, X') + 2 \sum_{j=|S_1|+1}^{n} \xi_j \eta_j,
$$

where $g(X, X')$ is some function of the variables in $X$ and $X'$, and

$$
\eta_j = \sum_{i=1}^{|S_1|} c_{ij} (\xi_i - \xi'_i).
$$

Let $\zeta$ be the number of $\eta_j$ which are equal to zero. If $J = [-\delta/2, \delta/2]$, we have

$$
\mathbb{P} \{ \varphi(X,Y) - \varphi(X',Y) \in J \} \leq \mathbb{P} \left\{ \varphi(X,Y) - \varphi(X',Y) \in J, \zeta \leq \frac{|S_2|}{2} \right\} \\
+ \mathbb{P} \left\{ \zeta > \frac{|S_2|}{2} \right\}.
$$
Since \( \zeta = \sum_{j=1}^{|S_2|} 1_{\{\eta_j = 0\}} \), using Lemma 3, we have

\[
E(\zeta) = \sum_{j=1}^{|S_2|} P\{ \eta_j = 0 \} = \sum_{j=1}^{|S_2|} O \left( \frac{\sum_{i=1}^r (1 - \pi(l_i)) \pi(l_i)}{\left\{ \sum_{i=1}^r [1 - \pi(l_i)] \right\}^{3/2}} \right)
\]

\[
= |S_2| O \left( \frac{\sum_{i=1}^r (1 - \pi(l_i)) \pi(l_i)}{\left\{ \sum_{i=1}^r [1 - \pi(l_i)] \right\}^{3/2}} \right),
\]

where \( \pi(l_i) \) and \( \pi(l_i) \) are the jumps associated with \( \xi_{l_i} - \xi_{l_i} \), \( l_i \in \{1, \ldots, |S_1|\} \) and we select those terms whose contribution to \( O(\cdot) \) are the smallest. By Markov’s inequality, we obtain

\[
P \left\{ \zeta > \frac{|S_2|}{2} \right\} \leq \frac{2}{|S_2|} E(\zeta) = O \left( \frac{\sum_{i=1}^r (1 - \pi(l_i)) \pi(l_i)}{\left\{ \sum_{i=1}^r [1 - \pi(l_i)] \right\}^{3/2}} \right).
\]

On the other hand, since \( v \leq |S_2|/2 \), we have at least \( |S_2|/2 \) for which \( \eta_j \neq 0 \). Hence

\[
P \left\{ \varphi(x, Y) - \varphi(x', Y) \in J, \zeta \leq \frac{|S_2|}{2} \right\} = O \left( \frac{\sum_{i=1}^{|S_2|/2} (1 - \kappa(l_i)) \kappa(l_i)}{\left\{ \sum_{i=1}^{|S_2|/2} [1 - \kappa(l_i)] \right\}^{3/2}} \right),
\]

where \( \kappa(l_i) \) and \( \kappa(l_i) \) are the jumps associated with \( \xi_{l_i}, \xi_{l_i} \in \{|S_1| + 1, \ldots, n\} \) we select those terms whose contribution to \( O(\cdot) \) are the smallest. So

\[
P \left\{ \varphi(X, Y) - \varphi(X', Y) \in J, \zeta \leq \frac{|S_2|}{2} \right\} =
\]

\[
= E \left( P \left\{ \varphi(X, Y) - \varphi(X', Y) \in J, \zeta \leq \frac{|S_2|}{2} \right\} \right)
\]

\[
= E \left( O \left( \frac{\sum_{i=1}^{|S_2|/2} (1 - \kappa(l_i)) \kappa(l_i)}{\left\{ \sum_{i=1}^{|S_2|/2} [1 - \kappa(l_i)] \right\}^{3/2}} \right) \right)
\]

\[
= O \left( \frac{\sum_{i=1}^{|S_2|/2} (1 - \kappa(l_i)) \kappa(l_i)}{\left\{ \sum_{i=1}^{|S_2|/2} [1 - \kappa(l_i)] \right\}^{3/2}} \right).
\]

Hence

\[
P \{ \varphi = x \} = O \left( \frac{\sum_{i=1}^r (1 - \pi(l_i)) \pi(l_i)}{\left\{ \sum_{i=1}^r [1 - \pi(l_i)] \right\}^{3/2}} + \frac{\sum_{i=1}^{|S_2|/2} (1 - \kappa(l_i)) \kappa(l_i)}{\left\{ \sum_{i=1}^{|S_2|/2} [1 - \kappa(l_i)] \right\}^{3/2}} \right)^{1/2}.
\]
Remark 6. a) If $\xi_1, \ldots, \xi_n$ are iid random variables,

$$\sup_{x \in \mathbb{R}} \mathbb{P} \{ \varphi = x \} = O \left( \left[ \frac{(1 - \kappa) \kappa_\Delta}{r^{1/2} (1 - \kappa) \Delta^{3/2}} + \frac{(1 - \kappa) \kappa_\Delta}{(|S_2|/2)^{1/2} (1 - \kappa) \Delta^{3/2}} \right]^{1/2} \right),$$

and if they are discrete random variables and $x_\kappa$ is not an accumulation point of $DF$, we have

$$\sup_{x \in \mathbb{R}} \mathbb{P} \{ \varphi = x \} = O \left( \left[ \frac{\kappa^2}{r (1 - \kappa)} \right]^{1/2} + \left( \frac{\kappa}{(|S_2|/2) (1 - \kappa)} \right)^{1/2} \right)^{1/2}$$

$$= O \left( \left[ \frac{\kappa^2}{\min \{r, |S_2|/2\} (1 - \kappa)} \right]^{1/4} \right).$$

b) Lemma 5 holds when $r$ many of the random variables $\xi_1, \ldots, \xi_n$ are degenerate for some $1 \leq r < n$. Contribution to the bound of the concentration inequality is only provided by the non-degenerated random variables.

3 Proofs in the Ginibre case

Let us consider an $m \times n$ matrix and fix one of its row vectors. We have two possibilities: either this row is a linear combination of the remaining ones or it is linearly independent of them. In the first case consider a basis of the remaining vectors; the row is a uniquely determined by a linear combination of this basis.

Following the terminology in Komlós [16], the number of non-zero coefficients of this linear combination is called the degree of the row vector with respect to this basis. The row degree of the row is the smallest of its degrees with respect to all bases constructed from the remaining vectors. The definition of the column degree of a column is analogous.

Proof of Theorem 4. First we write the following lemmas where the role of the size of the biggest jump is exhibited. They are proved similar to Lemma 1 in [16] who worked the case of independent identically non-degenerated distribution but without given a rate of convergence of the probability of non-singularity nor exhibiting the role of the biggest jump.

Lemma 6. Let $G_n$ be an $n \times n$ Ginibre matrix as in Theorem 3. Let $B_1(n)$ be the event that there is a column of $G_n$, the column degree of which is at most $n^{1-\epsilon}$, and let $B_2(n)$ be the event that there
is a row of $G_n$, the row degree of which is at most $n^{1-\varepsilon}$, where $\varepsilon \in (0, 1)$. Then for all $n$ large,

$$\mathbb{P} \{ B_1(n) \} \leq \kappa^{(n-n^{1-\varepsilon})/2},$$  \hspace{1cm} (16)

$$\mathbb{P} \{ B_2(n) \} \leq \kappa^{(n-n^{1-\varepsilon})/2}. $$  \hspace{1cm} (17)

**Proof.** Since $G_n$ is square, it is enough to prove (16) for all $n$ large. We note that if $v_1, \ldots, v_{r-1}$ are linearly independent and $v_r$ is a linear combination of these vectors, with $v_i \in \mathbb{R}^m$, then the linear dependence of $v_r$ with $v_1, \ldots, v_{r-1}$ is determined only by the last $m-r+1$ entries of $v_r$. If there is a column of column-degree at most $N := n^{1-\varepsilon}$, then there are $1 \leq i \leq N$ columns of $G_n$ that are linearly independent and we can write the $(i+1)$-th column of $G_n$ as a linear combination of them. We denote this event by $D(n,i)$. Then, conditioning on the first $i$-th columns and using the stochastic independence

$$\mathbb{P} \{ B_1(n) \} \leq \sum_{i=1}^{N} n \binom{n-1}{i} \mathbb{P} \{ D(n,i) \} \leq \sum_{i=1}^{N} n \binom{n-1}{i} \mathbb{E} \{ \mathbb{P} \{ D(n,i) \} | v_1, \ldots, v_i \} \leq \sum_{i=1}^{N} n n^N \kappa^{n-N} = N n^{N+1} \kappa^{n-N},$$  \hspace{1cm} (18)

for all $n$ large, $N n^{N+1} \kappa^{n-N} \leq \kappa^{(n-n^{1-\varepsilon})/2}$. \hfill \blacksquare

**Lemma 7.** Let $G_n$ be an $n \times n$ Ginibre matrix as in Theorem 4. Let $C_1(n)$ be the event that last column of $G_n$ is a linear combination of the other ones, and let $C_2(n)$ be the event that the last row of $G_n$ is a linear combination of the other ones. If $R(n)$ denotes the rank of $G_n$, we have

$$\mathbb{P} \{ C_1(n), R(n) < n \} = O_{\varepsilon} \left( \kappa^{(n-n^{1-\varepsilon})/2} + \frac{\kappa^2_{\Delta}}{n^{1-\varepsilon}(1-\kappa_{\Delta})} \right)^{1/2},$$  \hspace{1cm} (19)

$$\mathbb{P} \{ C_2(n), R(n) < n \} = O_{\varepsilon} \left( \kappa^{(n-n^{1-\varepsilon})/2} + \frac{\kappa^2_{\Delta}}{n^{1-\varepsilon}(1-\kappa_{\Delta})^3} \right)^{1/2},$$  \hspace{1cm} (20)

where the implicit constant in $O(\cdot)$ depends only on $\varepsilon \in (0, 1)$.

**Proof.** Since $G_n$ is square, it is only necessary to prove (19). We have

$$\mathbb{P} \{ C_1(n), R(n) < n \} = \mathbb{P} \{ C_1(n), R(n) < n, B_2(n) \} + \mathbb{P} \{ C_1(n), R(n) < n, B_2^c(n) \}.$$

The occurrence of the event $C_1(n) \cap \{ R(n) < n \} \cap B_2^c(n)$ implies that to preserve the linear
dependence, the entry \( \xi_{j_0,n}, 1 \leq j_0 \leq n \), of \( G_n \) should satisfy \( \xi_{j_0,n} = \sum_{i=1}^{t} \alpha_i \xi_{j_i,n}, j_i \neq j_0 \) and \( t < n \), where the values \( \alpha_i \) are determined only by entries in \( G_{n \times (n-1)} \) and the number of non-zero \( \alpha_i \) is at least \( N \). Hence, by Lemma 6 and Lemma 3,

\[
P \{ C_1(n), R(n) < n \} = \kappa^{(n-n^1-\varepsilon)/2} + O_{\varepsilon} \left( \left[ \frac{\kappa_{\Delta}^2}{n^{1-\varepsilon}(1 - \kappa_{\Delta})} \right]^{1/2} \right).
\]

Now we proceed with the proof of Theorem 4. Since the occurrence of either of \( C_1(n) \) or \( C_2(n) \) implies that \( \text{rank}(G_n) = \text{rank}(G_{n \times (n-1)}) \), from Lemmas 6 and 7, we can obtain the following inequalities

\[
P \{ R(n+1) = R(n) = n \} = O_{\varepsilon} \left( \kappa^{(n-n^1-\varepsilon)/2} + \left[ \frac{\kappa_{\Delta}^2}{n^{1-\varepsilon}(1 - \kappa_{\Delta})} \right]^{1/2} \right) \tag{21}
\]

\[
P \{ R(n+1) < R(n) + 2, R(n) < n \} = O_{\varepsilon} \left( \kappa^{(n-n^1-\varepsilon)/2} + \left[ \frac{\kappa_{\Delta}^2}{n^{1-\varepsilon}(1 - \kappa_{\Delta})} \right]^{1/2} \right) \tag{22}
\]

where \( R(n) \) denotes the rank of \( G_n \).

Finally, from (21) and (22),

\[
P \{ R(n+1) < n+1 \} = \]

\[
= \ P \{ R(n+1) < n+1, R(n) = n \} + P \{ R(n+1) < n+1, R(n) < n \}
\]

\[
\leq \ P \{ R(n+1) = R(n) = n \} + P \{ R(n+1) < R(n) + 2, R(n) < n \}
\]

\[
= \ O_{\varepsilon} \left( \kappa^{(n-n^1-\varepsilon)/2} + \left[ \frac{\kappa_{\Delta}^2}{n^{1-\varepsilon}(1 - \kappa_{\Delta})} \right]^{1/2} \right).
\]

\[
\text{Proof of (3) in Theorem 3} \quad \text{Since the constant in } O(\cdot) \text{ in Lemma 3 is universal, it is always possible to find an increasing sequence of natural numbers } \{ m_n : n \in \mathbb{N} \} \text{ such that }
\]

\[
\log(P \{ B_1(m_n) \}) \leq m_n \left[ \frac{(1 - \varepsilon) \log(m_n)}{m_n} + \frac{\log(m_n)}{m_n^\varepsilon} + \log(\kappa_n) \left( 1 - \frac{1}{m_n^\varepsilon} + \frac{1}{m_n} \right) \right] \to -\infty,
\]

and

\[
\frac{\kappa_{\Delta,n}^2}{m_n^{1-\varepsilon}(1 - \kappa_{\Delta,n})} \to 0.
\]

17
i.e.,
\[ \mathbb{P}\{ \text{rank}(G(m_n, \kappa_n)) < m_n \} = O \left( m_n^{-(1-\varepsilon)/2} \right). \]

Proof of Proposition 2 (i). We observe that if \( \kappa_n \to \kappa \in (0,1) \), then \( \kappa_n < \kappa + \Delta < 1 \) with \( \Delta \in (0,1) \) for \( n \) large and so it is possible to consider the same arguments as in the proof of Theorem 4.

4 Proofs in the Wigner case

Following the terminology introduced in Costello, Tao and Vu [8], given \( n \) vectors \( \{v_1, \ldots, v_n\} \), a linear combination of the \( v_i \)'s is a vector \( v = \sum_{i=1}^{n} c_i v_i \), where the \( c_i \) are real numbers. We say that a linear combination vanishes if \( v \) is the zero vector. A vanishing linear combination has degree \( k \) if exactly \( k \) among the \( c_i \) are nonzero.

A singular \( n \times n \) matrix is called normal if its row vectors do not admit a non-trivial vanishing linear combination with degree less than \( n^{1-\varepsilon} \) for a given \( \varepsilon \in (0,1) \). Otherwise it is said that the matrix is abnormal. Furthermore, a row of an \( n \times n \) non-singular matrix is called good if its exclusion leads to an \( (n-1) \times n \) matrix whose column vectors admit a non-trivial vanishing linear combination with degree at least \( n^{1-\varepsilon} \) (in fact, there is exactly one such combination as the rank of this \( (n-1) \times n \) matrix is \( n-1 \)). A row is said to be bad otherwise. Finally, an \( n \times n \) non-singular matrix \( A \) is perfect if every row in \( A \) is good row. If a non-singular matrix is not perfect, it is called imperfect.

For the proof of Theorem 5, we first present three lemmas which generalize results in [8] for Wigner matrices \( W_n = (\xi_{ij}) \) with independent entries which need not be identically distributed and the appropriate estimates in these new cases are found in terms of the size of the biggest jump of the distribution functions governing the entries. We also obtain a better rate of convergence which is universal. The proofs we give follow ideas in [8] but also take into account the size of the biggest jump.

Lemma 8. Let \( \varepsilon \in (0,1) \), then for all \( n \) large

\[ \mathbb{P}\{ W_n \text{ is singular and abnormal} \} \leq \kappa^{(n-n^{1-\varepsilon})/2} \] (23)
and

\[ \mathbb{P}\{W_n \text{ is non-singular and imperfect}\} \leq \kappa^{(n-n^{1-\varepsilon})/2}. \] (24)

**Proof.** If \( W_n \) is singular and abnormal the rows vectors of \( W_n \) admit a non-trivial vanishing linear combination with degree at most \( N := n^{1-\varepsilon} \). For \( i = 1, \ldots, N \), we have that if \( i = 1 \), there is a row of \( W_n \) that contains only zeros, and if \( i > 1 \), the \( i \)-th row is a linear combination of the first \( i-1 \) rows of \( W_n \) that are linearly independent. We denote by \( D(n, i) \) this last event and by \( T_{i-1} \) the upper triangular part of \( W_n \) until the row \( i-1 \) (included). The linear dependence of the \( i \)-th row of \( W_n \) with the \( i-1 \) rows of \( W_n \) is determined only by its last \( n^2 - i+1 \) entries. Then by the stochastic independence of \( T_{i-1} \) with the last \( n^2 - i+1 \) entries of the row \( i \)

\[
\mathbb{P}\{W_n \text{ is singular and abnormal}\} \leq \sum_{i=1}^{N} \binom{n}{i} \mathbb{P}\{D(n, i)\} \leq \sum_{i=1}^{N} \binom{n}{i} \mathbb{E}\{\mathbb{P}\{D(n, i)|T_{i-1}\}\}
\]

\[
\leq \sum_{i=1}^{N} n^N \kappa^{n-N+1} = Nn^N \kappa^{n-N+1},
\]

and for all \( n \) large,

\[
\mathbb{P}\{W_n \text{ is singular and abnormal}\} \leq \kappa^{\frac{3}{4}(n-n^{1-\varepsilon})} \leq \kappa^{\frac{1}{2}(n-n^{1-\varepsilon})}.
\]

Now, we consider the case when \( W_n \) is non-singular and imperfect. As \( W_n \) is symmetric and non-singular, \( W_n \) has only one bad row. If we consider the matrix \( W_{(n-1)\times(n-1)} \) obtained from \( W_n \) by removing the bad row and its transpose, again by the symmetry of \( W_n \), we have that \( W_{(n-1)\times(n-1)} \) is abnormal and singular, hence for all \( n \) large

\[
\mathbb{P}\{W_n \text{ is non-singular and imperfect}\} \leq n^2 \kappa^{\frac{3}{4}((n-1)-(n-1)^{1-\varepsilon})} \leq \kappa^{\frac{1}{2}(n-n^{1-\varepsilon})}.
\]

\[ \blacksquare \]

**Lemma 9.** Let \( A \) be a deterministic \( n \times n \) singular normal matrix, then

\[
\mathbb{P}\{\text{rank}(W_{n+1}) - \text{rank}(W_n) < 2 | W_n = A\} = O_{\varepsilon}\left( \frac{\kappa^2}{n^{1-\varepsilon}(1-\kappa\Delta)} \right)^{1/2}.
\]

**Proof.** Since \( r := \text{rank}(A) < n \), without loss of generality it is possible to suppose that the first \( r \) rows of \( A \) are linearly independent. If \( v_1, \ldots, v_r \) are the first rows of \( A \), then \( v_n = \sum_{i=1}^{r} \alpha_i v_i \),
and as $A$ is normal, the numbers of coefficients in this linear combination is at least $n^{1-\varepsilon}$. If it does not hold that $\xi_n = \sum_{i=1}^{r} \alpha_i \xi_i$, where $\xi_i$ are entries of the last column of $W_{n+1}$, by symmetry of $W_{n+1}$ we have $\text{rank}(W_{n+1}) = \text{rank}(A) + 2$. Hence

$$\mathbb{P}\{\text{rank}(W_{n+1}) - \text{rank}(W_n) < 2 \mid W_n = A\} \leq \mathbb{P}\left\{\xi_n = \sum_{i=1}^{r} \alpha_i \xi_i\right\} = O_\varepsilon\left(\left[\frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)}\right]^{1/2}\right).$$

The last expression follows from Lemma 3.

\begin{lemma}
Let $A$ be a deterministic $n \times n$ non-singular perfect symmetric matrix, then

$$\mathbb{P}\{\text{rank}(W_{n+1}) = n \mid W_n = A\} = O_\varepsilon\left(\left[\frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)}\right]^{1/4}\right).$$

\end{lemma}

\begin{proof}
If $\text{rank}(W_{n+1}) = n$, then $\det(W_{n+1}) = 0$, and we have

$$0 = \det(W_{n+1}) = (\det A)\xi_{n+1} + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \xi_i \xi_j,$$

where $\xi_i$ are entries of the last column of $W_{n+1}$ and its transpose, and $c_{ij}$ are cofactors of $A$. Since $A$ is normal, when we eliminate the $i$-th row of $A$, the columns of the matrix thus obtained admit a vanishing linear combination of degree smaller than $n^{1-\varepsilon}$. When the column $j$ is selected, where $j$ is the index of a non-zero coefficient in this linear combination, we obtain an $(n-1) \times (n-1)$ non-singular matrix. Hence

$$\mathbb{P}\{\text{rank}(W_{n+1}) = n \mid W_n = A\} \leq \mathbb{P}\left\{(\det A)\xi_{n+1} + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \xi_i \xi_j = 0\right\}$$

$$= \mathbb{E}\left\{(\det A)\xi_{n+1} + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \xi_i \xi_j = 0 \mid \xi_{n+1}\right\}\right.$$}

$$= \mathbb{E}\left(\mathbb{E}\left[\left[\frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)}\right]^{1/4}\right]\right)$$

$$= O_\varepsilon\left(\left[\frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa_\Delta)}\right]^{1/4}\right).$$

The last expression follows from Lemma 5.

\end{proof}
Now we consider the discrete stochastic process

\[
X_n = \begin{cases} 
0 & \text{if } \text{rank}(W_n) = n \\
\left(\kappa^{-1/8}\right)^{n - \text{rank}(W_n)} & \text{if } \text{rank}(W_n) < n,
\end{cases}
\]

for which we can prove the following result.

**Proposition 3.**

\[
E(X_n) = O_\varepsilon \left( \frac{\kappa^{\frac{3}{8}n - \frac{1}{2}\varepsilon}}{\kappa(1 - \kappa)} + \left[ \frac{\kappa^2}{n^{1-\varepsilon}(1 - \kappa\Delta)} \right]^{1/4} \right)
\]

**Proof.** For \( j = 0, \ldots, n \), write \( A_j = \{\text{rank}(W_n) = n - j\} \) and let \( 1 + \gamma = \kappa^{-1/8} \). We have

\[
E(X_n) = \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j\}
\]

\[
\leq \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j, W_n \text{ normal}\} + \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j, W_n \text{ abnormal}\} + \mathbb{P}\{A_0, W_n \text{ imperfect}\}
\]

\[
= \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j, W_n \text{ normal}\} + (1 + \gamma)^{n-1}\kappa^{n-1-\varepsilon}/2
\]

\[
\leq \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j, W_n \text{ normal}\} + 8\frac{\kappa^{\frac{3}{8}n - \frac{1}{2}\varepsilon}}{1 - \kappa}.
\]

On the other hand,

\[
E(X_{n+1}) = E(X_{n+1} | A_0, W_n \text{ perfect}) \mathbb{P}\{A_0, W_n \text{ perfect}\}
\]

\[
+ E(X_{n+1} | A_0, W_n \text{ imperfect}) \mathbb{P}\{A_0, W_n \text{ imperfect}\}
\]

\[
+ \sum_{j=1}^{n} E(X_{n+1} | A_0, W_n \text{ normal}) \mathbb{P}\{A_0, W_n \text{ normal}\}
\]

\[
+ \sum_{j=1}^{n} E(X_{n+1} | A_0, W_n \text{ abnormal}) \mathbb{P}\{A_0, W_n \text{ abnormal}\}
\]

\[
\leq E(X_{n+1} | A_0, W_n \text{ perfect}) \mathbb{P}\{A_0, W_n \text{ perfect}\}
\]

\[
+ \sum_{j=1}^{n} E(X_{n+1} | A_0, W_n \text{ normal}) \mathbb{P}\{A_0, W_n \text{ normal}\}
\]

\[
+ 8\frac{\kappa^{3/8}n - \frac{1}{2}\varepsilon}{1 - \kappa}.
\]

21
Since
\[ \mathbb{E}(X_{n+1} | A_0, W_n \text{ normal}) = O_\varepsilon \left( \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa \Delta)} \right]^{1/4} \right), \]
and, for \( j = 1, \ldots, n \), \( W_n \) has rank \( n - j \) and it is normal, we have that the rank of \( W_n \) can equal either \( n - j \) or \( n - j + 2 \). Thus
\[ \mathbb{E}(X_{n+1} | A_j, W_n \text{ normal}) \leq (1 + \gamma)^{j-1} + (1 + \gamma)^{j+1} O_\varepsilon \left( \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa \Delta)} \right]^{1/2} \right), \]
and, for \( n \) large,
\[ \mathbb{E}(X_{n+1} | A_j, W_n \text{ normal}) \leq \kappa^{1/8}(1 + \gamma)^{j}. \]
Finally,
\[ \mathbb{E}(X_{n+1}) \leq \kappa^{1/8} \sum_{j=1}^{n} (1 + \gamma)^j \mathbb{P}\{A_j, W_n \text{ normal}\} + 8 \kappa^2 n - \frac{1}{2} n^{1-\varepsilon} \frac{\kappa_\Delta^2}{\kappa(1 - \kappa)} + O_\varepsilon \left( \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa \Delta)} \right]^{1/4} \right). \]
This proves the proposition. □

**Proof of Theorem 5** By Markov’s inequality,
\[ \mathbb{P}\{\text{rank}(W_n) < n\} = \mathbb{P}\{X_n \geq 1\} \leq \mathbb{E}(X_n) \leq \mathbb{E}(X_n) \leq O_\varepsilon \left( \frac{\kappa^{n-\frac{1}{2}} n^{1-\varepsilon}}{\kappa(1 - \kappa)} + \left[ \frac{\kappa_\Delta^2}{n^{1-\varepsilon}(1 - \kappa \Delta)} \right]^{1/4} \right). \]
where we have used Proposition 3. □

The proofs of (4) in Theorem 3 and Proposition 2(ii) are similar to the proofs of (3) in Theorem 3 and Proposition 2(i) in the Ginibre case but now we need to use the expression (25).

**Acknowledgments**

Rahul Roy wants to thank CIMAT for the warm hospitality he received during his visits. The work of Paulo Manrique was supported by the Ph.D. Conacyt grant 210223.
References

[1] G. W. Anderson, A. Guionnet and O. Zeitouni, An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics, **118** (2010), Cambridge University Press, Cambridge.

[2] Z. D. Bai and J. W. Silverstein, Spectral Analysis of Large Dimensional Random Matrices. Mathematics Monograph Series 2, 2006, Science Press, Beijing.

[3] B. Bollobás, Random Graphs, Academic Press, New York, 1985.

[4] L. Bruneau and F. Germinet, On the singularity of random matrices with independent entries, *Proc. Amer. Math. Soc.* **137** (2009), 787-792.

[5] Bourgain, Vu and Wood, On the singularity probability of discrete random matrices, *J. Funct. Anal.* **258** (2010), 559-603.

[6] K. Costello and V. Vu, The rank of random graphs, *Random Struct. Alg.* **33** (2008), 269-285.

[7] K. Costello and V. Vu, On the rank of random sparse matrices, *Combinat. Probab. Comput.* **19** (2010), 321-342.

[8] K. Costello, T. Tao and Van Vu, Random symmetric matrices are almost surely non-singular, *Duke Math J.* **135** (2006), 395-413.

[9] L. Erdös, Universality of Wigner random matrices: A survey of recent results, *Russian Math. Surv.* **66** (2011), 507-626.

[10] L. Erdös, H.T. Yan and J. Yin, Bulk universality for generalized Wigner matrices. *Probab. Theory Relat. Fields* **154** (2012), 341-407.

[11] C. G. Essen, On the concentration function of a sum of independent random variables, *Z. Wahrscheinlinchkeitstheorie verw. Geb.* **9** (1968), 290-308.

[12] J. Kahn, J. Komlós and E. Szemerédi, On the probability a random ±1 matrix is singular, *J. Amer. Math Soc.* **8** (1995), 223-240.

[13] H. Kesten, A sharper form of the Doeblin-Lévy-Kolmogorov-Rogozin inequality for concentration function, *Math. Scand.* **25**, 133-144 (1969).
[14] H. Kesten, Sums of Independent Random Variables—Without Moment Conditions (The 1971 Rietz Lecture), *Ann. Math. Statist.* **43** (1972), 701-732.

[15] J. Komlós, On the determinant of (0, 1) matrices, *Studia. Sci. Math. Hungar.* **2** (1967), 7-21.

[16] J. Komlós, On the determinant of random matrices, *Studia. Sci. Math. Hungar.* **3** (1968), 387-399.

[17] H. H Nguyen, On the singularity of random combinatorial matrices, *SIAM J. Discrete Math.*, **27** (2013), no. 1, 447–458.

[18] M. Rudelson, Invertibility of random matrices: norm of the inverse, *Ann. Math.* **168**, 575-600.

[19] M. Rudelson and R. Vershynin, The Littlewood-Offord problem and invertibility of random matrices, *Adv. Math.* **218** (2008), 600-633.

[20] M. Rudelson and R. Vershynin, Non-asymptotic theory of random matrices: Extreme singular values, *Proceedings of the International Congress of Mathematicians*, Hyderabad, India (2010) 1576-1602.

[21] A. Slinko, A generalization of Komlos theorem on random matrices, *New Zealand J. Math.* **30** (2001), 81-86.

[22] T. Spencer, Random banded and sparse matrices, *The Oxford Handbook on Random Matrix Theory*, Oxford University Press (2011), 471-488.

[23] S. J. Szarek, Spaces with large distance to $l^n_\infty$ and random matrices, *Amer. J. Math.* **112** (1990), 899-942.

[24] T. Tao and V. Vu, On Random ±1 matrices: Singularity and Determinant, *Random Struct. Alg.* **28** (2006), 1-23.

[25] T. Tao and V. Vu, On the singularity of random Bernoulli matrices, *J. Amer. Math. Soc.* **20** (2007), 603-628.

[26] R. Vershynin, Invertibility of symmetric random matrices, *Random Struct. Alg.* **44** (2014), 135-182.