QUANTITATIVE INVERTIBILITY OF NON-HERMITIAN RANDOM MATRICES

KONSTANTIN TIKHOMIROV

Dedicated to Prof. Nicole Tomczak-Jaegermann

ABSTRACT

The problem of estimating the smallest singular value of random square matrices is important in connection with matrix computations and analysis of the spectral distribution. In this survey, we consider recent developments in the study of quantitative invertibility in the non-Hermitian setting, and review some applications of this line of research.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 60B20; Secondary 65F05, 15A18

KEYWORDS

Random matrices, condition number, spectrum

INTERNATIONAL CONGRESS OF MATHEMATICIANS 2022 JULY 6–14

Proc. Int. Cong. Math. 2022, Vol. 4, pp. 3292–3313
DOI 10.4171/ICM2022/10

Published by EMS Press

1. INTRODUCTION

Given an $N \times n$ ($N \geq n$) matrix A, its singular values are defined as square roots of the eigenvalues of the positive semidefinite $n \times n$ matrix A^*A:

$$s_i(A) := \sqrt{\lambda_i(A^*A)}, \quad i = 1, 2, \dots, n,$$

where we assume the nonincreasing ordering $\lambda_1(A^*A) \geq \lambda_2(A^*A) \geq \cdots \geq \lambda_n(A^*A)$. The classical Courant–Fischer–Weyl theorem provides a variational formula

$$s_i(A) = \min_{E:\, \dim(E)=n-i+1} \; \max_{x \in E,\, \|x\|_2=1} \|Ax\|_2, \quad 1 \leq i \leq n,$$

where the minimum taken over all linear subspaces E of the specified dimension. In particular, *the smallest* and *the largest* singular values of A can be computed as

$$s_{\min}(A) = s_n(A) = \min_{x:\|x\|_2=1} \|Ax\|_2, \quad s_{\max}(A) = s_1(A) = \max_{x:\|x\|_2=1} \|Ax\|_2.$$

Additionally, if the matrix A is square ($N = n$) and invertible then $s_{\min}(A) = \frac{1}{s_{\max}(A^{-1})}$.

The magnitude of the smallest singular value of square random matrices has attracted much attention due to the special role it plays in several questions of theoretical significance and in applications. In particular, the ratio of the largest and smallest singular values of a square matrix – *the condition number* – is systematically used in numerical analysis as a measure of sensitivity to round-off errors. Further, for certain random matrix models, bounds on the spectral norm of the matrix' resolvent (or, equivalently, the smallest singular value of diagonal shifts of the matrix) is a crucial point in the study of the spectral distribution. We refer to Sections 2 and 3 of the survey for a discussion of those directions.

In this survey, we consider *quantitative invertibility* of random *non-Hermitian* square matrices, including matrices with independent entries and adjacency matrices of random regular digraphs. The main objective in this line of research is to obtain bounds on the probability $\mathbb{P}\{s_{\min}(A) \leq t\}$ as a function of t, of the dimension, and, possibly, of some parameters of the model under consideration, such as the variance profile of the matrix or its mean.

One approach to the problem, which can be named analytical, is based on comparing the distribution of $s_{\min}(A)$ with the distribution of the smallest singular value of a corresponding Gaussian random matrix. The latter is very well understood [25] since explicit formulas for the joint distribution of the singular values of Gaussian matrices are available [46]. We refer to [14,85] for results of that type.

Another approach, which is the focus of this survey, falls in the category of *non-asymptotic* methods [75] and is based on a combination of techniques originated within asymptotic geometric analysis. It often produces very strong probability estimates, although typically lacks the precision of the analytical methods. The major features of this approach are (a) reducing the estimation of $s_{\min}$ to estimating distances between random vectors and random linear subspaces associated with the matrix, and (b) using concentration (Bernstein-type) and anti-concentration (Littlewood–Offord-type) inequalities. Often, this approach also involves constructing discretizations of certain subsets of $\mathbb{R}^n$ or $\mathbb{C}^n$ (*ε-nets*) and estimating their cardinalities. We will give a description of the features by considering multiple examples from the literature.

Because of some differences in methodology, and because we wish to emphasize the importance of the matrix invertibility for numerical analysis and in the study of the spectral distribution, this survey does not cover nonquantitative results on the singularity of random matrices. We note that estimating the singularity probability for several models of *discrete* random matrices is a major topic within the combinatorial random matrix theory [10,21,47,50,63,83]. In the last few years there has been a significant progress in this research direction (also, as corollaries of quantitative results), in particular, the problem of estimating the singularity probability of adjacency matrices of random regular (di)graphs [38,62,65], of Bernoulli random matrices [43,57,92] and, more generally, discrete matrices with i.i.d. entries [44], as well as of random symmetric matrices [11–13,29]. We refer to a recent survey [97] for a discussion and further references.

The rest of the survey is organized as follows. Sections 2 and 3 provide motivation for studying quantitative invertibility of non-Hermitian random matrices, and a brief account of known results. In Section 4, we give an overview of the methodology, starting with the result of Rudelson and Vershynin [72] as a main illustration. We then discuss novel additions to the methodology made in the past ten years, which allowed making progress on several important problems in the random matrix theory. Finally, in Section 5, we discuss some open problems.

Let us recall some notions which will be used further.

A random variable X on $\mathbb{R}$ or $\mathbb{C}$ is called *subgaussian* if $\mathbb{E}\exp(|X|^2/K^2) < \infty$ for some number $K > 0$. The smallest value of K such that $\mathbb{E}\exp(|X|^2/K^2) - 1 \leq 1$, is called *the subgaussian moment* of X. Any gaussian random variable is also subgaussian; further, all bounded random variables are subgaussian.

Given a sequence of *random* Borel probability measures $(\mu_m)_{m=1}^{\infty}$ and a random probability measure μ on $\mathbb{C}$, we say that μ_m converge *weakly in probability* to μ if for every bounded continuous function f on $\mathbb{C}$,

$$\lim_{m\to\infty} \mathbb{P}\left\{\left|\int f\,d\mu_m - \int f\,d\mu\right| > \varepsilon\right\} = 0, \quad \forall \varepsilon > 0.$$

We will denote by $\|\cdot\|$ the spectral norm of a matrix. The standard Euclidean norm in $\mathbb{R}^n$ or $\mathbb{C}^n$ will be denoted by $\|\cdot\|_2$. We will write $\operatorname{dist}(S,T)$ for the Euclidean distance between two subsets S and T of $\mathbb{R}^n$ or $\mathbb{C}^n$. By $S^{n-1}(\mathbb{R})$ or $S^{n-1}(\mathbb{C})$ we denote the unit Euclidean sphere in $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively. The constants will be denoted by C, c', etc.

2. QUANTITATIVE INVERTIBILITY IN MATRIX COMPUTATIONS

In this section, we discuss the importance of estimating the smallest singular value in numerical analysis, and provide a brief overview of related results on random matrices.

2.1. The condition number in numerical analysis

For an $n \times n$ invertible matrix $A \in \mathbb{C}^{n\times n}$, *the condition number of* A is defined as

$$\kappa(A) := \|A\| \|A^{-1}\| = \frac{s_{\max}(A)}{s_{\min}(A)}.$$

Consider a system of n linear equations in n variables, represented in the matrix–vector form as $Ax = b$. If the system is *well conditioned*, i.e., the condition number of the coefficient matrix A is small, a perturbation of the matrix or the coefficient vector does not strongly affect the solution. In particular, the round-off errors in matrix computations such as the Gaussian elimination, do not significantly distort the solution vector.

As an example of well-known theoretical guarantees, we mention an estimate on *the relative distance* between the solution of $Ax = b$ and the solution of a perturbed system

$$(A + F)y = (b + f).$$

The terms $F \in \mathbb{C}^{n\times n}$ and $f \in \mathbb{C}^n$ can be thought of as consequences of measurement or round-off errors. It is not difficult to check that under the assumption that $\delta := \max(\frac{\|F\|}{\|A\|}, \frac{\|f\|_2}{\|b\|_2})$ is small, the relative distance $\frac{\|y-x\|_2}{\|x\|_2}$ satisfies

$$\frac{\|y - x\|_2}{\|x\|_2} = O\big(\delta\kappa(A)\big)$$

(see, in particular, [33, SECTION 2.6.2], [80, SECTION 4]). In the specific setting when the system $Ax = b$ is solved using the Gaussian elimination with partial pivoting and the perturbation of the system is due to round-off errors, Wilkinson [98] showed that the relative distance between the computed and actual solutions can be bounded above by $n^{O(1)}\varepsilon\kappa(A)\rho$. Here *the growth factor* ρ is defined as $\rho := \frac{\max_{k=0,1,\dots;\, i,j\le n} |a_{ij}^{(k)}|}{\max_{i,j\le n} |a_{ij}|}$, with $a_{ij}^{(k)}$ being the (i, j)th element of the matrix $A^{(k)}$ obtained from A after k iterations of the Gaussian elimination process, and ε is the precision of the machine (see also [77, 94]).

Whereas the condition number of A characterizes sensitivity of the corresponding system of linear equations to small perturbations, *the eigenvector condition number* quantifies the stability of the spectrum and eigenvectors of A. The eigenvector condition number of a diagonalizable matrix $A \in \mathbb{C}^{n\times n}$ is defined as

$$\kappa_V(A) := \min_{W\in\mathbb{C}^{n\times n}:\, W^{-1}AW \text{ is diagonal}} \kappa(W) = \min_{W\in\mathbb{C}^{n\times n}:\, W^{-1}AW \text{ is diagonal}} \frac{s_{\max}(W)}{s_{\min}(W)}.$$

Clearly, $\kappa_V(A) = 1$ if and only if A is *unitarily diagonalizable* (normal). A classical stability result for a matrix spectrum using the eigenvector condition number is the Bauer–Fike theorem [8]. According to the theorem, given a diagonalizable matrix A and its perturbation $A + F$, the distance between any eigenvalue μ of $A + F$ and the spectrum of A can be estimated as

$$\min_{\lambda\in\operatorname{Spec}(A)} |\mu - \lambda| \le \kappa_V(A)\|F\|.$$

Moreover, stability of matrix functions under perturbations of the argument can be quantified using the eigenvector condition number (see [36, SECTION 3.3]). Here, we refer to a related line of research dealing with *the approximate diagonalization* of matrices, namely approximating a matrix with one having a small eigenvector condition number (see [2,3,22,41] and references therein). A connection between $\kappa_V(A)$ and quantitative invertibility of diagonal shifts of A is established through the notion of a pseudospectrum. *An ε-pseudospectrum of A*, denoted $\mathrm{Spec}_\varepsilon(A)$, is defined as the set of all points $z \in \mathbb{C}$ with $s_{\min}(A - z\,\mathbf{Id}) < \varepsilon$. It can be shown (see [23, LEMMA 9.2.11]) that for a diagonalizable matrix A with D being a corresponding diagonal matrix, $\mathrm{Spec}(D) + \kappa_V(A)^{-1}\varepsilon U \subset \mathrm{Spec}_\varepsilon(A) \subset \mathrm{Spec}(D) + \kappa_V(A)\varepsilon U$, where U is the unit disk of the complex plane.

2.2. Related results on random matrices

Randomness is a natural approach to simulate typical matrices observed in applications. For example, the LINPACK benchmark for measuring the computing power involves systems of linear equations with a randomly generated coefficient matrix [24]. Condition numbers of random square matrices with the computational perspective were first considered by von Neumann and Goldstine [96]. Rigorous results were obtained much later, notably by Edelman [25] for Gaussian random matrices (see also Szarek [82]). We note here that for *sufficiently dense* random matrices with i.i.d. entries satisfying certain moment conditions, estimating the largest singular values up to a constant multiple can be accomplished by a simple combination of Bernstein-type inequalities and an ε-net argument (see, for example, [75]), and with precision up to $(1 \pm o(1))$ multiple via the trace method [30,79,100]. Further, we only discuss estimates for the smallest singular value.

The average-case quantitative analysis of the matrix invertibility, when a typical matrix is modeled as a random matrix with independent entries and with matching first two moments, has been developed in multiple works. We refer, in particular, to papers [14, 85] employing the analytical approach, as well as works [5, 7, 37, 43, 44, 57–60, 67, 69, 71, 72, 89–92] based on the reduction to distance estimates and use of concentration/anticoncentration inequalities. Some of those results are mentioned below.

In [72], Rudelson and Vershynin showed that given a random $n \times n$ matrix A with i.i.d. real entries of zero mean, unit variance, and a bounded *subgaussian moment*, the smallest singular value of A satisfies

$$\mathbb{P}\big\{s_{\min}(A) \le n^{-1/2}t\big\} \le C(t + c^n), \quad t > 0,$$

where the constants $C > 0$ and $c \in (0,1)$ may only depend on the subgaussian moment (in fact, the statement is preserved if A is shifted by a nonrandom matrix with the spectral norm of order $O(\sqrt{n})$). The moment assumptions and the requirement that the entries are equidistributed were relaxed in later works [58,59,67]. On the other hand, in the special case of a matrix A with i.i.d. entries taking values $+1$ and -1 with probability $1/2$, it was proved in [92] that for any $\varepsilon > 0$,

$$\mathbb{P}\big\{s_{\min}(A) \le n^{-1/2}t\big\} \le Ct + C(1/2 + \varepsilon)^n, \quad t > 0,$$

where $C > 0$ is only allowed to depend on ε (see the introduction to [92], as well as [97], for a discussion of this result in the context of the combinatorial random matrix theory). An even stronger result is available when A has i.i.d. *discrete* entries which are not uniformly distributed on their support [44]: for every $\varepsilon > 0$ and assuming n is sufficiently large,

$$\mathbb{P}\{s_{\min}(A) \le n^{-1/2}t\} \le Ct + (1+\varepsilon)\mathbb{P}\{\text{two rows or columns of } A \text{ are colinear}\}, \quad t > 0,$$

with $C > 0$ depending only on the individual entry's distribution (see [44] for the statement in its full strength). In the setting when A has i.i.d. Bernoulli(p) entries and p is allowed to depend on n, its was shown in [7, 37, 57] that, as long as $p \le c$ for a small universal constant $c > 0$, for every $\varepsilon > 0$ and assuming n is sufficiently large,

$$\mathbb{P}\{s_{\min}(A) \le n^{-C}t\} \le t + (1+\varepsilon)\mathbb{P}\{\text{a row or a column of } A \text{ is zero}\}, \quad t > 0,$$

where $C > 0$ is a universal constant. We refer to [37] for a generalization to matrix rank estimates, as well as work [43] for sharp bounds in the setting of constant $p \in (0, 1/2)$, and [5] for stronger quantitative estimates in a certain range for the parameter p.

Put forward by Spielman and Teng [81], *the smoothed analysis* of the condition number is concerned with quantitative invertibility of a typical matrix in a small neighborhood of a fixed matrix (with possibly a very large spectral norm). A basic model of that type is of the form $A + M$, where M is a nonrandom matrix, and A has i.i.d. entries. The result of Sankar–Spielman–Teng [78] provided a bound for the smallest singular value of a shifted *Gaussian* real random matrix with i.i.d. standard normal entries, *independent* from the shift:

$$\mathbb{P}\{s_{\min}(A+M) \le tn^{-1/2}\} \le Ct, \quad t > 0, \ M \in \mathbb{R}^{n\times n},$$

for a certain universal constant $C > 0$ (see also [3, SECTION 2.3]). Analogous estimates for a broader class of random matrices with continuous distribution were later obtained in [91]. On the other hand, it was observed that for certain discrete random matrices, such as random sign (Bernoulli) matrices, no shift-independent small ball probability bounds for $s_{\min}(A + M)$ are possible [42, 87, 91]. In particular, it is shown in [42] that, assuming A has i.i.d. entries taking values ± 1 with probability $1/4$ and zero with probability $1/2$, for every $L \ge 1$ and every positive integer K,

$$\sup_{M:\|M\|\le n^L} \mathbb{P}\{s_{\min}(A+M) \le Cn^{-KL}\} \ge cn^{-K(K-1)/4},$$

where $C, c > 0$ may only depend on L and K. The smoothed analysis of the matrix condition number for discrete distributions was carried out in works [39, 42, 87, 88] (see also references therein). The following result was proved in [87]. Let $K, B, \varepsilon > 0$ and $L \ge 1/2$ be arbitrary parameters. Then, for all sufficiently large n, given an $n \times n$ random matrix A with i.i.d. centered entries of unit variance and the subgaussian moment bounded above by B, and given a nonrandom matrix M with $\|M\| \le n^L$, one has

$$\mathbb{P}\{s_{\min}(A+M) \le n^{-(2K+1)L}\} \le n^{-K+\varepsilon}.$$

In [42], it is shown that the above small ball probability bound can be significantly improved to match the average-case result of Rudelson and Vershynin [72], under the assumption that a positive fraction of the singular values of M are of order $O(\sqrt{n})$. More specifically, for every $\tilde{c} \in (0,1)$ and $\tilde{C} > 0$, and any fixed matrix M with $s_{n-\lfloor \tilde{c}n \rfloor}(M) \le \tilde{C}\sqrt{n}$, one has

$$\mathbb{P}\{s_{\min}(A+M) \le tn^{-1/2}\} \le C(t + c^n),$$

where $C > 0$ and $c \in (0,1)$ may only depend on $\tilde{c}$, $\tilde{C}$, and the subgaussian moment B. Under much weaker assumptions on the shift M, though at a price of precision, quantitative bounds for $s_{\min}(A+M)$ were obtained in [39].

3. INVERTIBILITY AND SPECTRUM

Given a square $n \times n$ matrix A_n, denote by μ_{A_n} its normalized spectral measure (spectral distribution):

$$\mu_{A_n} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(A_n)}.$$

For real and complex Gaussian matrices with i.i.d. standard entries (*the Ginibre ensemble*), explicit formulas for the joint distribution of the eigenvalues are known [26, 31, 51]. Those, in turn, were used by Mehta [61], Silverstein (unpublished; see [9, SECTION 3]) and Edelman [26] to derive convergence results for the spectral distribution in the Gaussian case.

In the non-Gaussian setting, where no similar formulas are available, Girko [32] proposed a *Hermitization* argument based on the identity

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\log\left|z-\lambda_i(A_n)\right| &= \frac{1}{n}\log\sqrt{\det\big((A_n - z\,\mathbf{Id})(A_n - z\,\mathbf{Id})^*\big)}\\ &= \frac{1}{n}\sum_{i=1}^{n}\log s_i(A_n - z\,\mathbf{Id}),\end{aligned}$$

which relates the spectrum to the singular values of the matrix resolvent. A modern form of the argument can be summarized as follows (see [9, LEMMA 4.3], as well as *the replacement principle* in [86]). Assume that a sequence of random matrices $(A_n)_{n=1}^{\infty}$ is such that for almost every $z \in \mathbb{C}$, the sequence of measures $\mu_{\sqrt{(A_n - z\,\mathbf{Id})(A_n - z\,\mathbf{Id})^*}}$ converges weakly in probability to a nonrandom probability measure μ_z. Assume further that the logarithm is *uniformly integrable in probability* with respect to $(\mu_{\sqrt{(A_n - z\,\mathbf{Id})(A_n - z\,\mathbf{Id})^*}})_{n=1}^{\infty}$ for almost every $z \in \mathbb{C}$, that is,

$$\lim_{t\to\infty}\sup_n \mathbb{P}\left\{\frac{1}{n}\sum_{i\le n:\,|\log s_i(A_n - z\,\mathbf{Id})|>t}\left|\log s_i(A_n - z\,\mathbf{Id})\right| > \varepsilon\right\} = 0, \quad \forall \varepsilon > 0. \tag{3.1}$$

Then there is a measure μ on $\mathbb{C}$ such that the sequence $(\mu_{A_n})_{n=1}^{\infty}$ converges to μ weakly in probability; moreover, the measure μ can be characterized in terms of $(\mu_z)_{z\in\mathbb{C}}$. We refer to [9] for proofs, as well as a detailed historical account of the study of the spectral distribution of non-Hermitian random matrices, up to 2000s.

In view of the uniform integrability requirement (3.1), strong quantitative estimates for small singular values of matrices $A_n - z\,\mathbf{Id}$ are an essential part of the Hermitization argument. In the setting of matrices with i.i.d. non-Gaussian entries, first rigorous estimates on the small singular values of $A_n - z\,\mathbf{Id}$ sufficient for the argument to go through were obtained for a class of continuous distributions by Bai [1], who applied the estimates to study the limiting spectral distribution in that setting. As the techniques to quantify invertibility of more general classes of matrices became available through the works of Tao–Vu [89], Rudelson [69], and Rudelson–Vershynin [72], the result of Bai was consequently generalized in works [34,66,84,86]. The *strong circular law* under minimal moment assumptions proved in [86] can be formulated as follows. Let ξ be a complex-valued random variable of zero mean and unit absolute second moment, and let $(A_n)_{n=1}^{\infty}$ be a sequence of random matrices, where each A_n is $n \times n$ with i.i.d. entries equidistributed with ξ. Then the sequence of spectral distributions $(\mu_{\frac{1}{\sqrt{n}}A_n})_{n=1}^{\infty}$ converges weakly almost surely to the uniform probability measure on the unit disk of the complex plane.

In the context of the circular law, the most studied model of *sparse* random matrices is of the form $A_n = B_n \odot M_n$, where B_n is the random matrix with i.i.d. Bernoulli(p_n) entries, M_n is independent from B_n and has i.i.d. entries equidistributed with a random variable ξ of unit variance, and "$\odot$" denotes the Hadamard (entrywise) product of matrices. In the regime $p_n \geq n^{-1+\varepsilon}$ for a fixed $\varepsilon > 0$, the (weak) circular law has been established in [99] following earlier works [34,84] dealing with additional moment assumptions.

In an even sparser regime, estimating the smallest singular value of $A_n - z\,\mathbf{Id}$ presents significant challenges, and further progress has only been made recently in [6,70]. In [70], it is proved that, assuming ξ is a real-valued random variable with unit variance, $np_n \leq n^{1/8}$, and np_n tends to infinity with n, and assuming the matrices $A_n = B_n \odot M_n$ are defined as in the previous paragraph, the sequence of spectral distributions $(\mu_{\frac{1}{\sqrt{np_n}}A_n})_{n=1}^{\infty}$ converges weakly in probability to the uniform measure on the unit disk of $\mathbb{C}$. A central technical result of [70] is the following quantitative bound for $s_{\min}(A_n - z\,\mathbf{Id})$: under the assumption that $|z| \leq np_n$ and $|\Im(z)| \geq 1$,

$$\mathbb{P}\left\{s_{\min}(A_n - z\,\mathbf{Id}) \leq \exp(-C\log^3 n)\right\} \leq C(np_n)^{-c},$$

where $C, c > 0$ may only depend on the c.d.f. of ξ.

Quantitative invertibility and spectrum of adjacency matrices of random regular directed graphs have been considered in multiple works in the past [4,15,17,18,53–55]. Given integers n and d, a *d-regular digraph* on vertices $\{1, 2, \dots, n\}$ is a directed graph in which every vertex has d incoming edges and d outgoing edges. Here, we focus on the model when no multiedges are allowed, but the graph may have loops (the latter condition is not conventional). For each n, denote by $A_{n,d}$ the adjacency matrix of a random graph uniformly distributed on the set of all d-regular digraphs on $\{1, 2, \dots, n\}$ (we allow d to depend on n). The first results on invertibility for this model were obtained by Cook [18]. The circular law for the sequence of spectral measures $(\mu_{\frac{1}{\sqrt{d(1-d/n)}}A_{n,d}})_{n=1}^{\infty}$ has been established in [17] under the assumption $\min(d, n-d) \geq \log^{96} n$. Later, in [53–55], the range $\omega(1) = d \leq \log^{96} n$ was treated. Either of the two results relies heavily on the estimates of the smallest singular values of $A_{n,d} - z\,\mathbf{Id}$. In particular, the main theorem of [53] is the following statement:

assuming $C \le d \le n/\log^2 n$ and $|z| \le d/6$,

$$\mathbb{P}\{s_{\min}(A_{n,d} - z\,\mathbf{Id}) < n^{-6}\} \le \frac{C\log^2 d}{\sqrt{d}},$$

where $C > 0$ is a universal constant.

The invertibility of *structured* random matrices and applications to the study of limiting spectral distribution have been considered, in particular, in [16, 19, 20, 40, 76]. A basic model of interest here is of the form $A_n = U_n \odot M_n - z\,\mathbf{Id}$, where M_n is a matrix with i.i.d. entries having zero mean and unit variance, $z \in \mathbb{C}$ is some complex number, U_n is a nonrandom matrix with nonnegative real entries encoding *the standard deviation profile*, and "$\odot$" denotes the Hadamard (entrywise) product of matrices. Note that A_n has mutually independent entries, with $\sqrt{\mathrm{Var}\, a_{ij}} = u_{ij}$, $1 \le i, j \le n$. In [76], the invertibility (and, more generally, the singular spectrum) of $U_n \odot M_n$ was studied in connection with the problem of estimation of matrix permanents. In particular, strong quantitative bounds on $s_{\min}(U_n \odot M_n)$ were obtained in the setting when M_n is the standard real Gaussian matrix, and U_n is a *broadly connected* profile (see [76, SECTION 2]). A significant progress in the study of structured random matrices was made by Cook in [16], who extended the result of [76] to non-Gaussian matrices, and obtained a polynomial lower bound on $s_{\min}(U_n \odot M_n - z\,\mathbf{Id})$ under very general assumptions on U_n. Namely, assuming that all entries of U_n are in the interval $[0, C]$, that $z \in [c\sqrt{n}, C\sqrt{n}]$ for some constants $c, C > 0$, and that the entries of M_n have a bounded $(4+\varepsilon)$-moment, the main result of [16] asserts that

$$\mathbb{P}\{s_{\min}(U_n \odot M_n - z\,\mathbf{Id}) \le n^{-\beta}\} \le n^{-\alpha}$$

for some $\alpha, \beta > 0$ depending only on c, C, ε, and the value of the $(4+\varepsilon)$-moment. In [19], this estimate was applied to derive limiting laws for the spectral distributions, under some additional assumptions on U_n. One of the results of [19] is *the circular law for doubly stochastic variance profiles*: provided that $\sum_{i=1}^n (U_n)_{ij}^2 = \sum_{i=1}^n (U_n)_{ji}^2 = n$, $1 \le j \le n$, and $\sup_n \max_{i,j} (U_n)_{ij} < \infty$, the sequence of spectral distributions $(\mu_{\frac{1}{\sqrt{n}} U_n \odot M_n})_{n=1}^\infty$ converges weakly in probability to the uniform measure on the unit disc of $\mathbb{C}$.

The setting of *sparse* structured matrices is not well understood. For results in that direction, we refer to a recent paper [40] dealing with the invertibility and spectrum of *block band matrices*.

4. METHODOLOGY

We start this section with a brief outline of [72] which will serve as an illustration of nonasymptotic methods (at the same time, we note that the argument of [72] is strongly influenced by earlier works, in particular, by Tao–Vu [89] and Rudelson [69]). The proof of the main theorem in [72] relies on four major components: sphere partitioning, invertibility via distance, ε-net arguments, and Littlewood–Offord-type inequalities.

Let A be an $n \times n$ matrix with i.i.d. real entries having zero mean and unit variance, and assume for simplicity that the entries are K-subgaussian for some constant $K > 0$.

A vector $x \in \mathbb{R}^n$ is called *m-sparse* if the size of its support is at most m. We will denote the set of all m-sparse vectors by $\mathrm{Sparse}_n(m)$. The proof of **[72, THEOREM 3.1]** starts with splitting $S^{n-1}(\mathbb{R})$ into sets of *compressible* and *incompressible* vectors,

$$\mathrm{Comp}_n(\delta,\rho) := \big\{x \in S^{n-1}(\mathbb{R}) : \mathrm{dist}\big(x, \mathrm{Sparse}_n(\delta n)\big) < \rho\big\};$$
$$\mathrm{Incomp}_n(\delta,\rho) := \big\{x \in S^{n-1}(\mathbb{R}) : \mathrm{dist}\big(x, \mathrm{Sparse}_n(\delta n)\big) \geq \rho\big\}.$$

Here, $\delta, \rho \in (0,1)$ are small constants. The variational formula for $s_{\min}(A)$ allows writing

$$\mathbb{P}\big\{s_{\min}(A) \leq s\big\} \leq \mathbb{P}\big\{\|Ax\|_2 \leq s \text{ for some } x \in \mathrm{Comp}_n(\delta,\rho)\big\} + \mathbb{P}\big\{\|Ax\|_2 \leq s \text{ for some } x \in \mathrm{Incomp}_n(\delta,\rho)\big\}, \quad s > 0.$$

If both δ and ρ are sufficiently small, the set of compressible vectors has small *covering numbers*, which allows applying an ε-net argument. More specifically, it can be checked that for every $\varepsilon \in (3\rho, 1/2]$, there is a discrete subset $\mathcal{N} \subset \mathrm{Comp}_n(\delta,\rho)$ of size at most $(\frac{C}{\varepsilon\delta})^{\delta n}$ such that for every $x \in \mathrm{Comp}_n(\delta,\rho)$, we have $\mathrm{dist}(x,\mathcal{N}) \leq \varepsilon$ (i.e., $\mathcal{N}$ is an ε-net in $\mathrm{Comp}_n(\delta,\rho)$ with respect to the Euclidean metric). Consequently, for every $L > 0$,

$$\begin{aligned}
&\mathbb{P}\big\{\|Ax\|_2 \leq s \text{ for some } x \in \mathrm{Comp}_n(\delta,\rho)\big\} \\
&\quad \leq \mathbb{P}\big\{\|Ay\|_2 \leq s + \varepsilon L\sqrt{n} \text{ for some } y \in \mathcal{N}\big\} + \mathbb{P}\big\{\|A\| > L\sqrt{n}\big\} \\
&\quad \leq |\mathcal{N}| \sup_{z \in S^{n-1}(\mathbb{R})} \mathbb{P}\big\{\|Az\|_2 \leq s + \varepsilon L\sqrt{n}\big\} + \mathbb{P}\big\{\|A\| > L\sqrt{n}\big\}.
\end{aligned}$$

For any $z \in S^{n-1}(\mathbb{R})$, the vector Az has i.i.d. subgaussian components with unit variances, and a standard Laplace transform argument implies that, as long as $s + \varepsilon L\sqrt{n}$ is much less than $\sqrt{n}$, the probability $\mathbb{P}\{\|Az\|_2 \leq s + \varepsilon L\sqrt{n}\}$ is exponentially small in n. Moreover, for a sufficiently large constant L, the probability $\mathbb{P}\{\|A\| > L\sqrt{n}\}$ is exponentially small in n. Therefore, an appropriate choice of parameters $\delta, \rho, \varepsilon, L$ yields

$$\mathbb{P}\big\{\|Ax\|_2 \leq s \text{ for some } x \in \mathrm{Comp}_n(\delta,\rho)\big\} \leq 2\exp(-cn), \quad s = o(\sqrt{n}).$$

We refer to **[72]** and **[75]** for details regarding the above computations. Let us note also that the idea of sphere partitioning was applied a few years earlier in **[56]** dealing with rectangular random matrices.

The incompressible vectors are treated using the *invertibility via distance* argument, which is based on the observation that for any incompressible vector x, a constant proportion of its components are of order $\Omega(n^{-1/2})$ by the absolute value. For every $1 \leq i \leq n$, denote by $H_i(A)$ the linear span of columns of A except the ith,

$$H_i(A) := \mathrm{Span}\big\{\mathrm{Col}_j(A),\ j \neq i\big\}.$$

Then for arbitrary vector x and arbitrary "threshold" $\tau > 0$ with $\{i : |x_i| \geq \tau\} \neq \emptyset$, we have

$$\|Ax\|_2 \geq \max_{1 \leq i \leq n} \big(|x_i|\mathrm{dist}\big(\mathrm{Col}_i(A), H_i(A)\big)\big) \geq \tau \max_{i:|x_i| \geq \tau} \mathrm{dist}\big(\mathrm{Col}_i(A), H_i(A)\big),$$

and hence for any $s > 0$, $\mathbf{1}_{\{\|Ax\|_2 \leq s\}} \leq \frac{1}{|\{i:|x_i| \geq \tau\}|} \sum_{i=1}^n \mathbf{1}_{\{\mathrm{dist}(\mathrm{Col}_i(A), H_i(A)) \leq s/\tau\}}$. This, combined with Markov's inequality and the fact that every (δ,ρ)–ncompressible vector is *spread*, i.e., has at least δn components of magnitude at least $\rho n^{-1/2}$, gives for $t > 0$,

$$\mathbb{P}\big\{\exists\, x \in \mathrm{Incomp}_n(\delta,\rho) : \|Ax\|_2 \le tn^{-1/2}\big\} \le \frac{1}{\delta n}\sum_{i=1}^{n} \mathbb{P}\big\{\mathrm{dist}\big(\mathrm{Col}_i(A), H_i(A)\big) \le t/\rho\big\} \tag{4.1}$$

(see [72, LEMMA 3.5]). Since the distribution of A is invariant under column permutations, the last relation can be rewritten as

$$\begin{aligned}\mathbb{P}\big\{\|Ax\|_2 \le tn^{-1/2} \text{ for some } x \in \mathrm{Incomp}_n(\delta,\rho)\big\} &\le \frac{1}{\delta}\mathbb{P}\big\{\mathrm{dist}\big(\mathrm{Col}_n(A), H_n(A)\big) \le t/\rho\big\}\\ &\le \frac{1}{\delta}\mathbb{P}\big\{\big|\big\langle \mathrm{Col}_n(A), Y_n(A)\big\rangle\big| \le t/\rho\big\},\end{aligned}$$

where $Y_n(A)$ denotes a unit normal to $H_n(A)$ measurable with respect to $\sigma(H_n(A))$.

The most involved part of [72] is the analysis of anticoncentration of $\langle \mathrm{Col}_n(A), Y_n(A)\rangle$. Recall that *the Lévy concentration function* $\mathcal{L}(Z,t)$ of a real variable Z is defined as

$$\mathcal{L}(Z,t) := \sup_{r\in\mathbb{R}} \mathbb{P}\big\{|Z-r| \le t\big\}, \quad t \ge 0.$$

The relationship between the magnitude of $\mathcal{L}(\sum_{i=1}^n a_i Z_i, t)$ for a linear combination of random variables $\sum_{i=1}^n a_i Z_i$ and the structure of the coefficient vector $(a_1,\dots,a_n)$ has been studied in numerous works, starting from an inequality of Erdős–Littlewood–Offord [27, 52]; we refer, in particular, to works [28, 48, 49, 68], as well as [89] and a survey [64] for a more recent account of *the Littlewood–Offord theory* and its applications to the matrix invertibility.

To characterize the structure of a coefficient vector in regard to anticoncentration, the notion of the *essential least common denominator* (LCD) has been introduced in [72]. We quote a slightly modified definition from [73]:

$$\mathrm{LCD}(a) := \inf\big\{\theta > 0 : \mathrm{dist}\big(\theta a, \mathbb{Z}^n\big) < \min\big(\gamma\|\theta a\|_2, \alpha\sqrt{n}\big)\big\}, \quad a \in \mathbb{R}^n.$$

Here, α,γ are small positive constants. The Littlewood–Offord-type inequality used in [72, 73] can be stated as follows. If $Z_1, Z_2, \dots, Z_n$ are i.i.d. real-valued random variables with $\mathbb{P}\{|Z_i - \mathbb{E}Z_i| < \beta\} \le 1-\beta$ for some $\beta > 0$ then for any unit vector $a \in \mathbb{R}^n$,

$$\mathcal{L}\left(\sum_{i=1}^n a_i Z_i, t\right) \le Ct + \frac{C}{\mathrm{LCD}(a)} + 2\exp(-cn), \quad t > 0, \tag{4.2}$$

where $C > 0$ may only depend on β,γ and $c > 0$ only on α,β (see [73] for a proof). Using an ε-net argument, the authors of [72] show that, with probability exponentially close to one, the random unit normal vector $Y_n(A)$ has an exponentially large LCD. This implies

$$\begin{aligned}\mathbb{P}\big\{\big|\big\langle \mathrm{Col}_n(A), Y_n(A)\big\rangle\big| \le s\big\} &\le \mathbb{P}\big\{\mathrm{LCD}\big(Y_n(A)\big) < \exp(c'n)\big\} + Cs + 2\exp(-c'n)\\ &\le Cs + 3\exp(-c''n), \quad s > 0.\end{aligned}$$

The combination of all the ingredients now gives the final estimate

$$\mathbb{P}\big\{s_{\min}(A) \le tn^{-1/2}\big\} \le \tilde{C}t + \tilde{C}\exp(-\hat{c}n), \quad t > 0,$$

matching, by the order of magnitude and up to the exponentially small additive term, the known asymptotics of $s_{\min}$ of Gaussian random matrices [25, 82].

In the remaining part of this section, we will consider some of the novel additions to the methodology made over the past years. To avoid technical details as much as possible, we will refer to compressible vectors, as well as all related notions from the literature, as *almost sparse* vectors, and to incompressible vectors and their relatives as *spread* vectors.

Invertibility over almost sparse vectors. In the setting of *dense* random matrices as described above, the set of almost sparse vectors AlSp_n can be treated by a simple ε-net argument since anticoncentration estimates for $\|Az\|_2$ for an *arbitrary* vector $z \in S^{n-1}$ are able to overpower the cardinality of the ε-net $\mathcal{N}$ in AlSp_n. In the case of sparse and certain models of structured random matrices, such an argument may not be sufficient since the product $|\mathcal{N}| \sup_{z \in S^{n-1}(\mathbb{R})} \mathbb{P}\{\|Az\|_2 \le s\}$ may become infinitely large even for small $s > 0$. We consider two (related) approaches to this problem from the literature.

The first is based on further subdividing AlSp_n into a few subsets $T_1, T_2, \ldots$ according to the size of set of vector's components of nonnegligible magnitude, and applying an ε-net argument within each subset. Anticoncentration estimates for Az for vectors $z \in T_i$ then compete with the cardinality of an ε-net on the set T_i rather than on the entire collection AlSp_n, which, for certain models, allows the proof to go through. We refer, in particular, to [76, SECTION 4] and [16, SECTION 3] for an application of this strategy to structured random matrices; as well as [17, PROPOSITION 3.1] dealing with adjacency matrices of random d-regular digraphs.

The second approach consists in identifying a class of nonrandom matrices $\mathcal{C}$ such that, for every $M \in \mathcal{C}$ and every almost sparse vector $z \in S^{n-1}$, Mz has a nonnegligible Euclidean norm, and then showing that, with probability close to one, $A \in \mathcal{C}$. As an example, consider a collection of matrices M such that for every nonempty subset $I \subset [n]$ with $|I| \le m$, there is a row $\mathrm{Row}_i(M)$ with $|\mathrm{supp}\,\mathrm{Row}_i(M) \cap I| = 1$. Then, it is not difficult to check that, for every nonzero m-sparse vector z, one has $Mz \neq 0$. It can further be verified that a random matrix A with i.i.d. Bernoulli(p) elements and $n^{-1}\mathrm{polylog}(n) \le p \le cm^{-1}$ belongs to this class with probability tending to one as $n \to \infty$ [5]. The construction can be made robust to treat almost sparse vectors, and can be further elaborated to deal with diagonal shifts of very sparse matrices [5, 53, 70].

Invertibility via distance. Relation (4.1) discovered in [72] can be applied to any model of randomness. However, this relation is not completely satisfactory when either (a) there are strong probabilistic dependencies between $\mathrm{Col}_i(A)$ and $H_i(A)$ which make estimating $\mathbb{P}\{\mathrm{dist}(\mathrm{Col}_i(A), H_i(A)) \le t\}$ challenging, or (b) invertibility over the almost sparse vectors cannot be treated with a desired precision using approaches based on ε-net arguments or on conditioning on a particular structure of the matrix. Here, we consider some developments of the invertibility via distance argument made in the contexts of d-regular random digraphs and smoothed analysis of the condition number.

Let $A_{n,d}$ be the adjacency matrix of a uniform random d-regular directed graph on n vertices. The regularity condition implies that for every $1 \le i \le n$, $\mathrm{Col}_i(A_{n,d})$ is a function of $\{\mathrm{Col}_j(A_{n,d})\}_{j\neq i}$, creating issues with applying the original version of the argument from [72]. In [18], Cook proposed a modification of the argument based on considering distances between the matrix columns and random subspaces of the form $H_{i_1,i_2,+}(A_{n,d}) := \mathrm{Span}\{\mathrm{Col}_j(A_{n,d}),\ j \neq i_1, i_2;\ \mathrm{Col}_{i_1}(A_{n,d}) + \mathrm{Col}_{i_2}(A_{n,d})\}$, for $i_1 \neq i_2$. This was later applied in [17,53]. Here, we quote [53, LEMMA 4.2]: denoting by $S(\rho,\delta)$ the collection of all unit vectors x in $\mathbb{C}^n$ with $\inf_{\lambda\in\mathbb{C}} |\{i \le n : |x_i - \lambda| > \rho n^{-1/2}\}| > \delta n$, one has

$$\begin{aligned}\mathbb{P}&\Big\{\inf_{x\in S(\rho,\delta)} \big\|(A_{n,d} - z\,\mathbf{Id})x\big\|_2 \le t n^{-1/2}\Big\}\\ &\le \frac{1}{\delta n^2} \sum_{\substack{i_1,i_2\in[n],\\ i_1\neq i_2}} \mathbb{P}\big\{\mathrm{dist}\big(\mathrm{Col}_{i_1}(A_{n,d} - z\,\mathbf{Id}), H_{i_1,i_2,+}(A_{n,d} - z\,\mathbf{Id})\big) \le t/\rho\big\}.\end{aligned}$$

Conditioned on a realization of $\mathrm{Col}_j(A_{n,d})$, $j \neq i_1, i_2$ (hence, also $Y := \mathrm{Col}_{i_1}(A_{n,d}) + \mathrm{Col}_{i_2}(A_{n,d})$), the support of the i_1th column of $A_{n,d}$ is uniformly distributed on the collection of d-subsets Q satisfying $\{j \le n : Y_j = 2\} \subset Q \subset \mathrm{supp}\, Y$. In the regime $d \to \infty$ as n tends to infinity, this is "sufficient randomness" for a satisfactory bound on $s_{\min}(A_{n,d} - z\,\mathbf{Id})$ required by the Hermitization argument [17,53].

We remark here that another version of the argument for matrices with dependencies based on evaluation of certain quadratic forms, introduced in [95], has been used in a non-Hermitian setting in [74] to estimate the smallest singular value of unitary and orthogonal perturbations of fixed matrices, which in turn is an important ingredient of *the single ring theorem* [35,74]. We refer to [74] for details.

In [91], a variant of the invertibility via the distance argument was developed to deal with nonrandom shifts of matrices with continuous distributions. The main observation of [91] is that the distances $\mathrm{dist}(\mathrm{Col}_i(A), H_i(A))$, $1 \le i \le n$, are highly correlated, which allows for a more efficient analysis than the first moment method estimate (4.1). The invertibility via distance is applied in [91] to the entire sphere rather than the set of spread vectors. As an illustration of the principle, we consider a simpler setting of centered random matrices when the argument is still able to produce new results. Assuming A is an $n \times n$ real random matrix with i.i.d. entries having zero mean, unit variance, and the distribution density bounded above by ρ, for every $t > 0$ and $1 \le k \le n$, one has $\mathbb{P}\{\exists\, I \subset [n] : |I| \ge k,\ \mathrm{dist}(\mathrm{Col}_i(A), H_i(A)) \le t \ \ \forall\, i \in I\} \le C_\rho t (n/k)^{5/11}$, where $C_\rho > 0$ may only depend on ρ (see [91, PROP. 3.8]). This, combined with the simple consequence of the *negative second moment identity*

$$s_{\min}(A) \ge \left(\sum_{i=1}^{n} \mathrm{dist}\big(\mathrm{Col}_i(A), H_i(A)\big)^{-2}\right)^{-1/2},$$

implies an estimate $\mathbb{P}\{s_{\min}(A) \le t n^{-1/2}\} \le C'_\rho t$, $t > 0$, which does not carry the c^n additive term inevitable when an ε-net-based approach is used. We refer to [91] for the more involved setting of noncentered random matrices.

Alternatives to the LCD. Functions of coefficient vectors different from the essential least common denominator have been introduced in the literature to deal with anticoncentration in the context of sparse and inhomogeneous random matrices, and matrices with dependencies. Here, we review some of them (for non-Hermitian models only).

The original notion of LCD is not applicable to the study of linear combinations of nonidentically distributed variables: in fact, given any vector $a \in S^{n-1}(\mathbb{R})$ with an exponentially large LCD, one can easily construct mutually independent variables $Z_1, \dots, Z_n$ with $\mathcal{L}(Z_i, 1) \le 1/2$, $1 \le i \le n$, and such that $\mathcal{L}(\sum_{i=1}^n a_i Z_i, 0) = \Omega(n^{-1/2})$. Given a random vector X in $\mathbb{R}^n$ and denoting by $\bar{X}$ the difference $X - X'$ (where X' is an independent copy of X), the *randomized least common denominator* with respect to X is defined by

$$\mathrm{RLCD}^X(a) := \inf\bigl\{\theta > 0 : \mathbb{E}\mathrm{dist}^2\bigl((\theta a_1 \bar{X}_1, \dots, \theta a_n \bar{X}_n), \mathbb{Z}^n\bigr) < \min\bigl(\gamma \|\theta a\|_2^2, \alpha n\bigr)\bigr\}, \quad a \in \mathbb{R}^n.$$

The notion was introduced in [59] to deal with inhomogeneous random matrices with different entry distributions. The small ball probability inequality (4.2) from [72,73] extends to the non-i.i.d. setting with the RLCD taking place of the original notion. We refer to [59] for details.

Strong quantitative invertibility results for matrices with fixed rowsums and adjacency matrices of d-regular digraphs obtained recently in [93] and [45], respectively, rely on a modification of the LCD which allows treating linear combinations of Bernoulli variables conditioned on their sum. Specifically, in [93] the notion of the *combinatorial least common denominator* CLCD is defined by

$$\mathrm{CLCD}(a) := \inf\bigl\{\theta > 0 : \mathrm{dist}\bigl(\theta(a_i - a_j)_{i<j}, \mathbb{Z}^{\binom{n}{2}}\bigr) < \min\bigl(\gamma \bigl\|\theta(a_i - a_j)_{i<j}\bigr\|_2, \alpha n\bigr)\bigr\}, \quad a \in \mathbb{R}^n,$$

where $(a_i - a_j)_{i<j}$ denotes a vector in $\mathbb{R}^{\binom{n}{2}}$ with the (i, j)th coordinate equal to $a_i - a_j$, $1 \le i < j \le n$. It is further shown that for the random vector $(Z_1, Z_2, \dots, Z_n)$ uniformly distributed on the collection of $0/1$ vectors with exactly $n/2$ ones, an analog of the anticoncentration inequality (4.2) holds, with LCD replaced by CLCD. A modification of the notion, called QCLCD, was further considered in [45]. We refer to that paper for details.

Another functional – *the degree of unstructuredness* UD – was introduced in [57] to study the invertibility of sparse Bernoulli random matrices. The main observation exploited in [57] is that, for $p = o(1)$, linear combinations of i.i.d. Bernoulli(p) random variables $\sum_{i=1}^n a_i Z_i$ are often more concentrated than corresponding linear combinations of *dependent* $0/1$ variables conditioned to sum to a fixed number of order $\Theta(pn)$. In [57], the argument proceeds by conditioning on the size of the support of a column of the matrix and estimating the anticoncentration of $\mathrm{dist}(\mathrm{Col}_i(A), H_i(A)) = |\langle \mathrm{Col}_i(A), Y_i(A)\rangle|$ in terms of the degree of unstructuredness of the unit random normal $Y_i(A)$. The definition of UD is technically involved, and we do not provide it here; see [57] for details.

Average-case analysis of anticoncentration. The average-case study of Littlewood–Offord-type inequalities for linear combinations $\sum_{i=1}^n a_i Z_i$, introduced in the random matrix context in [92], was a crucial element in some of recent advances on quantitative invertibility of random discrete matrices [43,44,92], which helped resolve some long-standing problems in the combinatorial random matrix theory. The main idea of [92] is, rather than attempting to obtain an explicit description of vectors a such that $\sum_{i=1}^{n} a_i Z_i$ is strongly anticoncentrated, to consider the linear combination for a *randomly chosen* coefficient vector (with an appropriately defined notion of randomness). This approach allowed strengthening the invertibility results available through the use of the LCD. As an example, we consider a simplified version of the main technical result of [92]. Let $\varepsilon \in (0, 1/2)$, $M \geq 1$. Then there exist $n_0 = n_0(\varepsilon, M)$ depending on ε, M and $L_0 = L_0(\varepsilon) > 0$ depending *only* on ε (and not on M) with the following property. Take $n \geq n_0$, $1 \leq N \leq (1/2 + \varepsilon)^{-n}$, and let $\mathcal{A} := (\{-2N, \ldots, -N-1\} \cup \{N+1, \ldots, 2N\})^n$. Assume that a random vector $a = (a_1, \ldots, a_n)$ is uniformly distributed on $\mathcal{A}$. Then

$$\mathbb{P}_a\left\{\mathcal{L}_Z(a_1 Z_1 + \cdots + a_n Z_n, \sqrt{n}) > L_0 N^{-1}\right\} \leq e^{-Mn}.$$

Here, $\mathcal{L}_Z(\cdot, \cdot)$ denotes the Lévy concentration function with respect to the randomness of $(Z_1, \ldots, Z_n)$, a vector with independent ± 1 components. The main point of the statement is that the parameter L_0 controlling the anticoncentration of the linear combination does not depend on M, i.e., the proportion of the coefficient vectors in $\mathcal{A}$ such that the anticoncentration of $a_1 Z_1 + \cdots + a_n Z_n$ is weak, becomes *superexponentially small* in n as $n \to \infty$.

Matrices with heavy entries. For the invertibility of (dense) random matrices with independent entries assuming only finite second moments, we refer to [58, 59, 67].

5. OPEN PROBLEMS

We conclude this survey with a selection of open research problems.

Refined smoothed analysis of invertibility. Recall that a standard model in the setting of the smoothed analysis of the condition number is of the form $A + M$, where A is an $n \times n$ random matrix with i.i.d. entries, and M is a nonrandom shift.

Problem 1 (Shift-independent estimates for matrices with continuous distributions). Let ξ be a real random variable with zero mean, unit variance, and bounded distribution density. Let A be an $n \times n$ matrix with i.i.d. entries equidistributed with ξ. It is true that for every nonrandom matrix M,

$$\mathbb{P}\left\{s_{\min}(A + M) \leq t n^{-1/2}\right\} \leq Ct, \quad t > 0,$$

where $C > 0$ may only depend on the c.d.f. of ξ (and not on n)?

For partial results on the above problem, see [78, 91].

Problem 2 (Optimal dependence of $s_{\min}(A + M)$ on the norm of the shift in the discrete setting). Let A be an $n \times n$ matrix with i.i.d. ± 1 entries, and let $T, t > 0$ be parameters. For any $\varepsilon, L > 0$, estimate $\sup_{M:\|M\| \leq T} \mathbb{P}\{s_{\min}(A + M) \leq t\}$ up to a multiplicative error $O(n^{\varepsilon})$ and an additive error $O(n^{-L})$, that is, find an explicit function $f(n, T, t)$ such that

$$n^{-\varepsilon} f(n, T, t) - Cn^{-L} \leq \sup_{M:\|M\| \leq T} \mathbb{P}\left\{s_{\min}(A + M) \leq t\right\} \leq n^{\varepsilon} f(n, T, t) + Cn^{-L},$$

where $C > 0$ may only depend on ε and L.

For the best known partial results on the above problem, see [42, 87].

Problem 3 (Dependence of $s_{\min}(A+M)$ on M in the Gaussian setting). Let A be an $n\times n$ matrix with i.i.d. standard real Gaussian entries. Find an estimate on $\mathbb{E}s_{\min}(A+M)$ in terms of the singular spectrum of M.

One can assume in the above problem that M is a diagonal matrix with the ith diagonal element $s_i(M)$, $1\le i\le n$. Note that A may either improve or degrade the invertibility of M.

Invertibility and spectrum of very sparse matrices. Here, we consider the problem of identifying the limiting spectral distribution for non-Hermitian matrices with *constant* average number of nonzero elements in a row/column.

Problem 4 (The oriented Kesten–McKay law; see [9, SECTION 7]). Let $d\ge 3$. For each n, let $A_{n,d}$ be the adjacency matrix of a uniform random d-regular directed graph on n vertices. Prove that the sequence of spectral distributions $(\mu_{A_{n,d}})_{n=1}^{\infty}$ converges weakly to the probability measure on $\mathbb{C}$ with the density function

$$\rho_d(z):=\frac{1}{\pi}\frac{d^2(d-1)}{(d^2-|z|^2)^2}\mathbf{1}_{\{|z|<\sqrt{d}\}}.$$

Assuming the standard Hermitization approach to the above problem, the following is the crucial (perhaps the main) step of the argument:

Problem 5. Let $d\ge 3$ and let $(A_{n,d})_{n=1}^{\infty}$ be as above. Prove that for almost every $z\in\mathbb{C}$ and every $\varepsilon>0$,

$$\lim_{n\to\infty}\mathbb{P}\big\{s_{\min}(A_{n,d}-z\,\mathbf{Id})\le\exp(-\varepsilon n)\big\}=0.$$

Problem 6 (Spectrum of directed Erdős–Renyi graphs of constant average degree). Let $\alpha>0$. For each $n\ge\alpha$, let A_n be an $n\times n$ random matrix with i.i.d. Bernoulli(α/n) entries. Does a sequence of spectral distributions (μ_{A_n}) converge weakly to a nonrandom probability measure?

As in the case of regular digraphs, assuming the Hermitization argument, the following problem constitutes an important step in understanding the asymptotics of the spectrum:

Problem 7. For each $n\ge\alpha$, let A_n be an $n\times n$ random matrix with i.i.d. Bernoulli(α/n) entries. Is it true that for almost every $z\in\mathbb{C}$ and every $\varepsilon>0$,

$$\lim_{n\to\infty}\mathbb{P}\big\{s_{\min}(A_n-z\,\mathbf{Id})\le\exp(-\varepsilon n)\big\}=0?$$

Invertibility and spectrum of structured random matrices. The spectrum of structured random matrices in the absence of expansion-like properties (such as *broad connectivity* [76] or *robust irreducibility* [19, 20]) is not well understood as of now. In particular, a full description of the class of inhomogeneous matrices with independent entries with spectral convergence to the circular law seems to be out of reach of modern methods.

Problem 8. Give a complete description of sequences of standard deviation profiles $(U_n)_{n=1}^{\infty}$ satisfying the following condition: assuming that ξ is any random variable with zero mean and unit variance, and that for each n, M_n is an $n \times n$ matrix with i.i.d. entries equidistributed with ξ, the sequence of spectral distributions $(\mu_{U_n \odot M_n})$ converges weakly in probability to the uniform measure on the unit disc of $\mathbb{C}$.

A natural class of profiles considered, in particular, in [19,20], are *doubly stochastic* profiles. One may expect that those profile sequences, under some weak assumption on the magnitude of the maximal entry, should be sufficient for the circular law to hold:

Problem 9. Assume that for each n, the standard deviation profile U_n satisfies

$$\sum_{i=1}^{n} (U_n)_{ij}^2 = \sum_{i=1}^{n} (U_n)_{ji}^2 = 1, \quad 1 \le j \le n,$$

and that for some $\varepsilon > 0$, $\limsup_n \max_{ij} ((U_n)_{ij} n^{\varepsilon}) = 0$. Is it true that, with M_n as in the above problem, the sequence $(\mu_{U_n \odot M_n})$ converges weakly in probability to the uniform measure on the unit disc of $\mathbb{C}$?

Note that the above setting allows sparse matrices (cf. [19, THEOREM 2.4]). Solving the above problem, if approached with Girko's Hermitization procedure, requires satisfactory bounds on the smallest singular values of $U_n \odot M_n - z\,\mathbf{Id}$.

FUNDING

This work was partially supported by the Sloan Research Fellowship and by NSF grant DMS 2054666.

REFERENCES

[1] Z. D. Bai, Circular law. *Ann. Probab.* **25** (1997), no. 1, 494–529.

[2] J. Banks, J. Garza Vargas, A. Kulkarni, and N. Srivastava, Overlaps, eigenvalue gaps, and pseudospectrum under real Ginibre and absolutely continuous perturbations. 2020, arXiv:2005.08930.

[3] J. Banks, A. Kulkarni, S. Mukherjee, and N. Srivastava, *Gaussian regularization of the seudospectrum and Davies' conjecture*. CPAM, to appear.

[4] A. Basak, N. Cook, and O. Zeitouni, Circular law for the sum of random permutation matrices. *Electron. J. Probab.* **23** (2018), Paper No. 33, 51 pp.

[5] A. Basak and M. Rudelson, Invertibility of sparse non-Hermitian matrices. *Adv. Math.* **310** (2017), 426–483.

[6] A. Basak and M. Rudelson, The circular law for sparse non-Hermitian matrices. *Ann. Probab.* **47** (2019), no. 4, 2359–2416.

[7] A. Basak and M. Rudelson, Sharp transition of the invertibility of the adjacency matrices of sparse random graphs. *Probab. Theory Related Fields* **180** (2021), no. 1–2, 233–308.

[8] F. L. Bauer and C. T. Fike, Norms and exclusion theorems. *Numer. Math.* **2** (1960), 137–141.

[9] C. Bordenave and D. Chafaï, Around the circular law. *Probab. Surv.* **9** (2012), 1–89.

[10] J. Bourgain, V. H. Vu, and P. M. Wood, On the singularity probability of discrete random matrices. *J. Funct. Anal.* **258** (2010), no. 2, 559–603.

[11] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe, Singularity of random symmetric matrices revisited. 2020, arXiv:2011.03013.

[12] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe, The singularity probability of a random symmetric matrix is exponentially small. 2021, arXiv:2105.11384.

[13] M. Campos, L. Mattos, R. Morris, and N. Morrison, On the singularity of random symmetric matrices. *Duke Math. J.* **170** (2021), no. 5, 881–907.

[14] Z. Che and P. Lopatto, Universality of the least singular value for sparse random matrices. *Electron. J. Probab.* **24** (2019), Paper No. 9, 53 pp.

[15] N. Cook, The circular law for random regular digraphs with random edge weights. *Random Matrices Theory Appl.* **6** (2017), no. 3, 1750012, 23 pp.

[16] N. Cook, Lower bounds for the smallest singular value of structured random matrices. *Ann. Probab.* **46** (2018), no. 6, 3442–3500.

[17] N. Cook, The circular law for random regular digraphs. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (2019), no. 4, 2111–2167.

[18] N. A. Cook, On the singularity of adjacency matrices for random regular digraphs. *Probab. Theory Related Fields* **167** (2017), no. 1–2, 143–200.

[19] N. Cook, W. Hachem, J. Najim, and D. Renfrew, Non-Hermitian random matrices with a variance profile (I): deterministic equivalents and limiting ESDs. *Electron. J. Probab.* **23** (2018), Paper No. 110, 61 pp.

[20] N. Cook, W. Hachem, J. Najim, and D. Renfrew, Non-Hermitian random matrices with a variance profile (II): properties and examples. 2020, arXiv:2007.15438.

[21] K. P. Costello, T. Tao, and V. Vu, Random symmetric matrices are almost surely nonsingular. *Duke Math. J.* **135** (2006), no. 2, 395–413.

[22] E. B. Davies, Approximate diagonalization. *SIAM J. Matrix Anal. Appl.* **29** (2007), no. 4, 1051–1064.

[23] E. B. Davies, *Linear operators and their spectra*. Cambridge Stud. Adv. Math. 106, Cambridge University Press, Cambridge, 2007.

[24] J. J. Dongarra, P. Luszczek, and A. Petitet, The LINPACK Benchmark: past, present and future. *Concurrency Comput.: Pract. Exp.* **15** (2003), 803–820.

[25] A. Edelman, Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.* **9** (1988), no. 4, 543–560.

[26] A. Edelman, The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.* **60** (1997), no. 2, 203–232.

[27] P. Erdős, On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.* **51** (1945), 898–902.

[28] C. G. Esseen, On the Kolmogorov–Rogozin inequality for the concentration function. *Z. Wahrsch. Verw. Gebiete* **5** (1966), 210–216.

[29] A. Ferber and V. Jain, Singularity of random symmetric matrices—a combinatorial approach to improved bounds. *Forum Math. Sigma* **7** (2019), Paper No. e22, 29 pp.

[30] S. Geman, A limit theorem for the norm of random matrices. *Ann. Probab.* **8** (1980), no. 2, 252–261.

[31] J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** (1965), 440–449.

[32] V. L. Girko, The circular law. *Teor. Veroyatn. Primen.* **29** (1984), no. 4, 669–679.

[33] G. H. Golub and C. F. Van Loan, *Matrix computations*. 4th edn., Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, 2013.

[34] F. Götze and A. Tikhomirov, The circular law for random matrices. *Ann. Probab.* **38** (2010), no. 4, 1444–1491.

[35] A. Guionnet, M. Krishnapur, and O. Zeitouni, The single ring theorem. *Ann. of Math. (2)* **174** (2011), no. 2, 1189–1217.

[36] N. J. Higham, *Functions of matrices*. SIAM, Philadelphia, 2008.

[37] H. Huang, Rank of sparse Bernoulli matrices. 2020, arXiv:2009.13726.

[38] J. Huang, Invertibility of adjacency matrices for random d-regular graphs. 2018, arXiv:1807.06465.

[39] V. Jain, Quantitative invertibility of random matrices: a combinatorial perspective. *Discrete Anal.*, to appear.

[40] V. Jain, I. Jana, K. Luh, and S. O'Rourke, Circular law for random block band matrices with genuinely sublinear bandwidth. 2020, arXiv:2008.03850.

[41] V. Jain, A. Sah, and M. Sawhney, On the real Davies' conjecture. 2020, arXiv:2005.08908.

[42] V. Jain, A. Sah, and M. Sawhney, On the smoothed analysis of the smallest singular value with discrete noise. *Bull. Lond. Math. Soc.*, to appear.

[43] V. Jain, A. Sah, and M. Sawhney, Sharp invertibility of random Bernoulli matrices. 2020, arXiv:2010.06553.

[44] V. Jain, A. Sah, and M. Sawhney, Singularity of discrete random matrices. 2020, arXiv:2010.06554.

[45] V. Jain, A. Sah, and M. Sawhney, The smallest singular value of dense random regular digraphs. 2020, arXiv:2008.04755.

[46] A. T. James, The distribution of the latent roots of the covariance matrix. *Ann. Math. Stat.* **31** (1960), 151–158.

[47] J. Kahn, J. Komlós, and E. Szemerédi, On the probability that a random ± 1-matrix is singular. *J. Amer. Math. Soc.* **8** (1995), no. 1, 223–240.

[48] H. Kesten, A sharper form of the Doeblin–Lévy–Kolmogorov–Rogozin inequality for concentration functions. *Math. Scand.* **25** (1969), 133–144.

[49] A. Kolmogorov, Sur les propriétés des fonctions de concentrations de M. P. Lévy. *Ann. Inst. Henri Poincaré* **16** (1958), 27–34.

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

[51] N. Lehmann and H.-J. Sommers, Eigenvalue statistics of random real matrices. *Phys. Rev. Lett.* **67** (1991), no. 8, 941–944.

[52] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation. III. *Rec. Math. [Mat. Sbornik] N.S.* **12** (1943), no. 54, 277–286.

[53] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, The smallest singular value of a shifted d-regular random square matrix. *Probab. Theory Related Fields* **173** (2019), no. 3–4, 1301–1347.

[54] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, Structure of eigenvectors of random regular digraphs. *Trans. Amer. Math. Soc.* **371** (2019), no. 11, 8097–8172.

[55] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, Circular law for sparse random regular digraphs. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 2, 467–501.

[56] A. E. Litvak, A. Pajor, M. Rudelson, and N. Tomczak-Jaegermann, Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.* **195** (2005), no. 2, 491–523.

[57] A. E. Litvak and K. Tikhomirov, Singularity of sparse Bernoulli matrices. *Duke Math. J.*, to appear.

[58] G. V. Livshyts, The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding. *J. Anal. Math.* **145** (2021), 257–306.

[59] G. V. Livshyts, K. Tikhomirov, and R. Vershynin, The smallest singular value of inhomogeneous square random matrices. *Ann. Probab.* **49** (2021), no. 3, 1286–1309.

[60] K. Luh, Complex random matrices have no real eigenvalues. *Random Matrices Theory Appl.* **7** (2018), no. 1, 1750014, 17 pp.

[61] M. L. Mehta, *Random matrices and the statistical theory of energy levels*. Academic Press, New York, 1967.

[62] A. Mészáros, The distribution of sandpile groups of random regular graphs. *Trans. Amer. Math. Soc.* **373** (2020), no. 9, 6529–6594.

[63] H. H. Nguyen, Inverse Littlewood–Offord problems and the singularity of random symmetric matrices. *Duke Math. J.* **161** (2012), no. 4, 545–586.

[64] H. H. Nguyen and V. H. Vu, Small ball probability, inverse theorems, and applications. In *Erdős centennial*, pp. 409–463, Bolyai Soc. Math. Stud. 25, János Bolyai Math. Soc., Budapest, 2013.

[65] H. H. Nguyen and M. M. Wood, Cokernels of adjacency matrices of random r-regular graphs. 2018, arXiv:1806.10068.

[66] G. Pan and W. Zhou, Circular law, extreme singular values and potential theory. *J. Multivariate Anal.* **101** (2010), no. 3, 645–656.

[67] E. Rebrova and K. Tikhomirov, Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries. *Israel J. Math.* **227** (2018), no. 2, 507–544.

[68] B. A. Rogozin, On the increase of dispersion of sums of independent random variables. *Teor. Veroyatn. Primen.* **6** (1961), 106–108.

[69] M. Rudelson, Invertibility of random matrices: norm of the inverse. *Ann. of Math. (2)* **168** (2008), no. 2, 575–600.

[70] M. Rudelson and K. Tikhomirov, The sparse circular law under minimal assumptions. *Geom. Funct. Anal.* **29** (2019), no. 2, 561–637.

[71] M. Rudelson and R. Vershynin, The least singular value of a random square matrix is $O(n^{-1/2})$. *C. R. Math. Acad. Sci. Paris* **346** (2008), no. 15–16, 893–896.

[72] M. Rudelson and R. Vershynin, The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** (2008), no. 2, 600–633.

[73] M. Rudelson and R. Vershynin, Smallest singular value of a random rectangular matrix. *Comm. Pure Appl. Math.* **62** (2009), no. 12, 1707–1739.

[74] M. Rudelson and R. Vershynin, Invertibility of random matrices: unitary and orthogonal perturbations. *J. Amer. Math. Soc.* **27** (2014), no. 2, 293–338.

[75] M. Rudelson and R. Vershynin, Non-asymptotic theory of random matrices: extreme singular values. In *Proceedings of the International Congress of Mathematicians. Volume III*, pp. 1576–1602, Hindustan Book Agency, New Delhi, 2010.

[76] M. Rudelson and O. Zeitouni, Singular values of Gaussian matrices and permanent estimators. *Random Structures Algorithms* **48** (2016), no. 1, 183–212.

[77] A. Sankar, *Smoothed analysis of Gaussian elimination*. PhD Thesis, 2004.

[78] A. Sankar, D. A. Spielman, and S.-H. Teng, Smoothed analysis of the condition numbers and growth factors of matrices. *SIAM J. Matrix Anal. Appl.* **28** (2006), no. 2, 446–476.

[79] Y. Seginer, The expected norm of random matrices. *Combin. Probab. Comput.* **9** (2000), no. 2, 149–166.

[80] S. Smale, On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc. (N.S.)* **13** (1985), no. 2, 87–121.

[81] D. A. Spielman and S.-H. Teng, Smoothed analysis of algorithms. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pp. 597–606, Higher Ed. Press, Beijing.

[82] S. J. Szarek, Condition numbers of random matrices. *J. Complexity* **7** (1991), no. 2, 131–149.

[83] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.* **20** (2007), no. 3, 603–628.

[84] T. Tao and V. Vu, Random matrices: the circular law. *Commun. Contemp. Math.* **10** (2008), no. 2, 261–307.

[85] T. Tao and V. Vu, Random matrices: the distribution of the smallest singular values. *Geom. Funct. Anal.* **20** (2010), no. 1, 260–297.

[86] T. Tao and V. Vu, Random matrices: universality of ESDs and the circular law. *Ann. Probab.* **38** (2010), no. 5, 2023–2065.

[87] T. Tao and V. Vu, Smooth analysis of the condition number and the least singular value. *Math. Comp.* **79** (2010), no. 272, 2333–2352.

[88] T. Tao and V. Vu, The condition number of a randomly perturbed matrix. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 248–255, ACM, New York, 2007.

[89] T. Tao and V. H. Vu, Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)* **169** (2009), no. 2, 595–632.

[90] K. Tatarko, An upper bound on the smallest singular value of a square random matrix. *J. Complexity* **48** (2018), 119–128.

[91] K. Tikhomirov, Invertibility via distance for noncentered random matrices with continuous distributions. *Random Structures Algorithms* **57** (2020), no. 2, 526–562.

[92] K. Tikhomirov, Singularity of random Bernoulli matrices. *Ann. of Math. (2)* **191** (2020), no. 2, 593–634.

[93] T. Tran, The smallest singular value of random combinatorial matrices. 2019, arXiv:1909.04219.

[94] L. N. Trefethen and R. S. Schreiber, Average-case stability of Gaussian elimination. *SIAM J. Matrix Anal. Appl.* **11** (1990), no. 3, 335–360.

[95] R. Vershynin, Invertibility of symmetric random matrices. *Random Structures Algorithms* **44** (2014), no. 2, 135–182.

[96] J. von Neumann and H. H. Goldstine, Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.* **53** (1947), 1021–1099.

[97] V. H. Vu, Recent progress in combinatorial random matrix theory. *Probab. Surv.* **18** (2021), 179–200.

[98] J. H. Wilkinson, Error analysis of direct methods of matrix inversion. *J. Assoc. Comput. Mach.* **8** (1961), 281–330.

[99] P. M. Wood, Universality and the circular law for sparse random matrices. *Ann. Appl. Probab.* **22** (2012), no. 3, 1266–1300.

[100] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah, On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** (1988), no. 4, 509–521.

KONSTANTIN TIKHOMIROV

School of Mathematics, 686 Cherry street, Atlanta GA 30332, ktikhomirov6@gatech.edu