Electronic Journal of Probability

Vol. 13 (2008), Paper no. 5, pages 107–134.

Journal URL
http://www.math.washington.edu/~ejpecp/

Concentration of the Spectral Measure for Large Random Matrices with Stable Entries

Christian Houdré *   Hua Xu †

Abstract

We derive concentration inequalities for functions of the empirical measure of large random matrices with infinitely divisible entries, in particular, stable or heavy tails ones. We also give concentration results for some other functionals of these random matrices, such as the largest eigenvalue or the largest singular value.

Key words: Spectral Measure, Random Matrices, Infinitely divisibility, Stable Vector, Concentration.

AMS 2000 Subject Classification: Primary 60E07, 60F10, 15A42, 15A52.

Submitted to EJP on June 12, 2007, final version accepted January 18, 2008.

*Georgia Institute of Technology, School of Mathematics, Atlanta, Georgia, 30332, houdre@math.gatech.edu
†Georgia Institute of Technology, School of Mathematics, Atlanta, Georgia, 30332, xu@math.gatech.edu

1 Introduction and Statements of Results:

Large random matrices have recently attracted a lot of attention in fields such as statistics, mathematical physics or combinatorics (e.g., see Mehta [24], Bai and Silverstein [4], Johnstone [18], Anderson, Guionnet and Zeitouni [2]). For various classes of matrix ensembles, the asymptotic behavior of the, properly centered and normalized, spectral measure or of the largest eigenvalue is understood. Many of these results hold true for matrices with independent entries satisfying some moment conditions (Wigner [35], Tracy and Widom [33], Soshnikov [28], Girko [8], Pastur [25], Bai [3], Götze and Tikhomirov [9]).

There is relatively little work outside the independent or finite second moment assumptions. Let us mention Soshnikov [30] who, using ideas from perturbation theory, studied the distribution of the largest eigenvalue of Wigner matrices with entries having heavy tails. (Recall that a real (or complex) Wigner matrix is a symmetric (or Hermitian) matrix whose entries $\mathbf{M}_{i,i}, 1 \leq i \leq N$, and $\mathbf{M}_{i,j}, 1 \leq i < j \leq N$, form two independent families of iid (complex valued in the Hermitian case) random variables.) In particular, (see [30]), for a properly normalized Wigner matrix with entries belonging to the domain of attraction of an α-stable law, $\lim_{N\to\infty} \mathbb{P}^N(\lambda_{max} \leq x) = \exp(-x^{-\alpha})$ (here λ_{max} is the largest eigenvalue of such a normalized matrix). Further, Soshnikov and Fyodorov [32], using the method of determinants, derived results for the largest singular value of $K \times N$ rectangular matrices with independent Cauchy entries, showing that the largest singular value of such a matrix is of order K^2N^2 (see also the survey article [31], where band and sparse matrices are studied).

On another front, Guionnet and Zeitouni [10], gave concentration results for functionals of the empirical spectral measure of, self-adjoint, random matrices whose entries are independent and either satisfy a logarithmic Sobolev inequality or are compactly supported. They obtained, for such matrices, the subgaussian decay of the tails of the empirical spectral measure when it deviates from its mean. They also noted that their technique could be applied to prove results for the largest eigenvalue or for the spectral radius of such matrices. Alon, Krivelevich and Vu [1] further obtained concentration results for any of the eigenvalues of a Wigner matrix with uniformly bounded entries (see, Ledoux [20] for more developments and references).

Our purpose in the present work is to deal with matrices whose entries form a general infinitely divisible vector, and in particular a stable one (without independence assumption). As well known, unless degenerated, an infinitely divisible random variable cannot be bounded. We obtain concentration results for functionals of the corresponding empirical spectral measure, allowing for any type of light or heavy tails. The methodologies developed here apply as well to the largest eigenvalue or to the spectral radius of such random matrices.

Following the lead of Guionnet and Zeitouni [10], let us start by setting our notation and framework.

Let $\mathcal{M}_{N\times N}(\mathbb{C})$ be the set of $N \times N$ Hermitian matrices with complex entries, which is throughout equipped with the Hilbert-Schmidt (or Frobenius or entrywise Euclidean) norm:

$$\|\mathbf{M}\| = \sqrt{tr(\mathbf{M}^*\mathbf{M})} = \sqrt{\sum_{i,j=1}^{N} |\mathbf{M}_{i,j}|^2}.$$

Let f be a real valued function on $\mathbb{R}$. The function f can be viewed as mapping $\mathcal{M}_{N\times N}(\mathbb{C})$ to $\mathcal{M}_{N\times N}(\mathbb{C})$. Indeed, for $\mathbf{M} = (\mathbf{M}_{i,j})_{1\leq i,j\leq N} \in \mathcal{M}_{N\times N}(\mathbb{C})$, so that $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, where $\mathbf{D}$ is a diagonal matrix, with real entries $\lambda_1, ..., \lambda_N$, and $\mathbf{U}$ is a unitary matrix, set

$$f(\mathbf{M}) = \mathbf{U}f(\mathbf{D})\mathbf{U}^*,\ f(\mathbf{D}) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{pmatrix}.$$

Let $tr(\mathbf{M}) = \sum_{i=1}^N \mathbf{M}_{i,i}$ be the trace operator on $\mathcal{M}_{N\times N}(\mathbb{C})$ and set also

$$tr_N(\mathbf{M}) = \frac{1}{N}\sum_{i=1}^N \mathbf{M}_{i,i}.$$

For a $N \times N$ random Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$, let $F_N(x) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{\lambda_i \le x\}}$ be the corresponding empirical spectral distribution function. As well known, if $\mathbf{M}$ is a $N \times N$ Hermitian Wigner matrix with $\mathbb{E}[\mathbf{M}_{1,1}] = \mathbb{E}[\mathbf{M}_{1,2}] = 0$, $\mathbb{E}[|\mathbf{M}_{1,2}|^2] = 1$, and $\mathbb{E}[\mathbf{M}_{1,1}^2] < \infty$, the spectral measure of $\mathbf{M}/\sqrt{N}$ converges to the semicircle law: $\sigma(dx) = \sqrt{4-x^2}\mathbf{1}_{\{|x|\le 2\}}dx/2\pi$ ([2]).

We study below the tail behavior of either the spectral measure or the linear statistic of $f(\mathbf{M})$ for classes of matrices $\mathbf{M}$. Still following Guionnet and Zeitouni, we focus on a general random matrix $\mathbf{X_A}$ given as follows:

$$\mathbf{X_A} = ((\mathbf{X_A})_{i,j})_{1\le i,j\le N},\ \mathbf{X_A} = \mathbf{X_A^*},\ (\mathbf{X_A})_{i,j} = \frac{1}{\sqrt{N}}A_{i,j}\omega_{i,j},$$

with $(\omega_{i,j})_{1\le i,j\le N} = (\omega_{i,j}^R + \sqrt{-1}\omega_{i,j}^I)_{1\le i,j\le N}$, $\omega_{i,j} = \overline{\omega_{j,i}}$, and where $\omega_{i,j}$, $1 \le i \le j \le N$ is a complex valued random variable with law $P_{i,j} = P_{i,j}^R + \sqrt{-1}P_{i,j}^I$, $1 \le i \le j \le N$, with $P_{i,i}^I = \delta_0$ (by the Hermite property). Moreover, the matrix $\mathbf{A} = (A_{i,j})_{1\le i,j\le N}$ is Hermitian with, in most cases, non-random complex valued entries uniformly bounded, say, by a.

Different choices for the entries of $\mathbf{A}$ allow to cover various types of ensembles. For instance, if $\omega_{i,j}, 1 \le i < j \le N$, and $\omega_{i,i}, 1 \le i \le N$, are iid $N(0,1)$ random variables, taking $A_{i,i} = \sqrt{2}$ and $A_{i,j} = 1$, for $1 \le i < j \le N$ gives the GOE (Gaussian Orthogonal Ensemble). If $\omega_{i,j}^R, \omega_{i,j}^I, 1 \le i < j \le N$, and $\omega_{i,i}^R, 1 \le i \le N$, are iid $N(0,1)$ random variables, taking $A_{i,i} = 1$ and $A_{i,j} = 1/\sqrt{2}$, for $1 \le i < j \le N$ gives the GUE (Gaussian Unitary Ensemble) (see [24]). Moreover, if $\omega_{i,j}^R, \omega_{i,j}^I, 1 \le i < j \le N$, and $\omega_{i,i}^R, 1 \le i \le N$, are two independent families of real valued random variables, taking $A_{i,j} = 0$ for $|i-j|$ large and $A_{i,j} = 1$ otherwise, gives band matrices.

Proper choices of non-random $A_{i,j}$ also make it possible to cover Wishart matrices, as seen in the later part of this section. In certain instances, A can also be chosen to be random, like in the case of diluted matrices, in which case $A_{i,j}, 1 \le i \le j \le N$, are iid Bernoulli random variables (see [10]).

On $\mathbb{R}^{N^2}$, let $\mathbb{P}^N$ be the joint law of the random vector $X = (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)$, $1 \le i < j \le N$, where it is understood that the indices for $\omega_{i,i}^R$ are $1 \le i \le N$. Let $\mathbb{E}^N$ be the corresponding expectation. Denote by $\hat{\mu}_A^N$ the empirical spectral measure of the eigenvalues of $\mathbf{X_A}$, and further note that

$$tr_N f(\mathbf{X_A}) = \frac{1}{N}tr(f(\mathbf{X_A})) = \int_{\mathbb{R}} f(x)\hat{\mu}_{\mathbf{A}}^N(dx),$$

for any bounded Borel function f. For a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, set

$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|},$$

where throughout $\|\cdot\|$ is the Euclidean norm, and where we write $f \in Lip(c)$ whenever $\|f\|_{Lip} \leq c$. Each element $\mathbf{M}$ of $\mathcal{M}_{N\times N}(\mathbb{C})$ has a unique collection of eigenvalues $\lambda = \lambda(\mathbf{M}) = (\lambda_1, \cdots, \lambda_N)$ listed in non increasing order according to multiplicity in the simplex

$$\mathcal{S}^N = \{\lambda_1 \geq \cdots \geq \lambda_N : \lambda_i \in \mathbb{R}, 1 \leq i \leq N\},$$

where throughout $\mathcal{S}^N$ is equipped with the Euclidian norm $\|\lambda\| = \sqrt{\sum_{i=1}^N \lambda_i^2}$. It is a classical result sometimes called Lidskii's theorem ([26]), that the map $\mathcal{M}_{N\times N}(\mathbb{C}) \to \mathcal{S}^N$ which associates to each Hermitian matrix its ordered list of real eigenvalues is 1-Lipschitz ([11], [19]). For a matrix $\mathbf{X_A}$ under consideration with eigenvalues $\lambda(\mathbf{X_A})$, it is then clear that the map $\varphi : (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)_{1\leq i<j\leq N} \mapsto \lambda(\mathbf{X_A})$ is Lipschitz, from $(\mathbb{R}^{N^2}, \|\cdot\|)$ to $(\mathcal{S}^N, \|\cdot\|)$, with Lipschitz constant bounded by $a\sqrt{2/N}$. Moreover, for any real valued Lipschitz function F on $\mathcal{S}^N$ with Lipschitz constant $\|F\|_{Lip}$, the map $F \circ \varphi$ is Lipschitz, from $(\mathbb{R}^{N^2}, \|\cdot\|)$ to $\mathbb{R}$, with Lipschitz constant at most $a\|F\|_{Lip}\sqrt{2/N}$. Appropriate choices of F ([19], [2]) ensure that the maximal eigenvalue $\lambda_{max}(\mathbf{X_A}) = \lambda_1(\mathbf{X_A})$ and, in fact, any one of the N eigenvalues is a Lipschitz function with Lipschitz constants at most $a\sqrt{2/N}$. Similarly, the spectral radius $\rho(\mathbf{X_A}) = \max_{1\leq i\leq N} |\lambda_i|$ and $tr_N(f(\mathbf{X_A}))$, where $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function, are themselves Lipschitz functions with Lipschitz constants at most $a\sqrt{2/N}$ and $\sqrt{2}a\|f\|_{Lip}/N$, respectively. These observations (and our results) are also valid for the real symmetric matrices, with proper modification of the Lipschitz constants.

Next, recall that X is a d-dimensional infinitely divisible random vector without Gaussian component, $X \sim ID(\beta, 0, \nu)$, if its characteristic function is given by,

$$\begin{aligned} \varphi_X(t) &= \mathbb{E}e^{i\langle t,X\rangle} \\ &= \exp\left\{ i\langle t, \beta\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle t,u\rangle} - 1 - i\langle t, u\rangle \mathbf{1}_{\|u\|\leq 1} \right) \nu(du) \right\}, \end{aligned} \tag{1.1}$$

where $t, \beta \in \mathbb{R}^d$ and $\nu \not\equiv 0$ (the *Lévy measure*) is a positive measure on $\mathcal{B}(\mathbb{R}^d)$, the Borel σ-field of $\mathbb{R}^d$, without atom at the origin, and such that $\int_{\mathbb{R}^d}(1 \wedge \|u\|^2)\nu(du) < +\infty$. The vector X has independent components if and only if its Lévy measure ν is supported on the axes of $\mathbb{R}^d$ and is thus of the form:

$$\nu(dx_1, \ldots, dx_d) = \sum_{k=1}^d \delta_0(dx_1) \ldots \delta_0(dx_{k-1}) \tilde{\nu}_k(dx_k) \delta_0(dx_{k+1}) \ldots \delta_0(dx_d), \tag{1.2}$$

for some one-dimensional Lévy measures $\tilde{\nu}_k$. Moreover, the $\tilde{\nu}_k$ are the same for all $k = 1, \ldots, d$, if and only if X has identically distributed components.

The following proposition gives an estimate on any median (or the mean, if it exists) of a Lipschitz function of an infinitely divisible vector X. It is used in most of the results presented in this paper. The first part is a consequence of Theorem 1 in [14], while the proof of the second part can be obtained as in [14].

Proposition 1.1. *Let* $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N} \sim ID(\beta, 0, \nu)$ *in* $\mathbb{R}^{N^2}$. *Let* $V^2(x) = \int_{\|u\|\le x} \|u\|^2 \nu(du)$, $\bar{\nu}(x) = \int_{\|u\|>x} \nu(du)$, *and for any* $\gamma > 0$, *let* $p_\gamma = \inf\big\{x > 0 : 0 < V^2(x)/x^2 \le \gamma\big\}$. *Let* $f \in Lip(1)$, *then for any* γ *such that* $\bar{\nu}(p_\gamma) \le 1/4$,

(i) any median $m(f(X))$ *of* $f(X)$ *satisfies*

$$|m(f(X)) - f(0)| \le G_1(\gamma) := p_\gamma\Big(\sqrt{\gamma} + 3k_\gamma(1/4)\Big) + E_\gamma,$$

(ii) the mean $\mathbb{E}^N[f(X)]$ *of* $f(X)$, *if it exists, satisfies*

$$|\mathbb{E}^N[f(X)] - f(0)| \le G_2(\gamma) := p_\gamma\Big(\sqrt{\gamma} + k_\gamma(1/4)\Big) + E_\gamma,$$

where $k_\gamma(x)$, $x > 0$, *is the solution, in* y, *of the equation*

$$y - (y+\gamma)\ln\left(1 + \frac{y}{\gamma}\right) = \ln x,$$

and where

$$E_\gamma = \left(\sum_{k=1}^{N^2}\Big(\langle e_k, \beta\rangle - \int_{p_\gamma<\|y\|\le 1}\langle e_k, y\rangle \nu(dy) + \int_{1<\|y\|\le p_\gamma}\langle e_k, y\rangle\nu(dy)\Big)^2\right)^{1/2}, \tag{1.3}$$

with $e_1, e_2, \ldots, e_{N^2}$ *being the canonical basis of* $\mathbb{R}^{N^2}$.

Our first result deals with the spectral measure of a Hermitian matrix whose entries on and above the diagonal form an infinitely divisible random vector with finite exponential moments. Below, for any $b > 0$, $c > 0$, let

$$Lip_b(c) = \Big\{f : \mathbb{R} \to \mathbb{R} \ : \|f\|_{Lip} \le c,\ \|f\|_\infty \le b\Big\},$$

while for a fixed compact set $\mathcal{K} \subset \mathbb{R}$, with diameter $|\mathcal{K}| = \sup_{x,y\in\mathcal{K}} |x - y|$, let

$$Lip_{\mathcal{K}}(c) := \{f : \mathbb{R} \to \mathbb{R} \ : \|f\|_{Lip} \le c,\ supp(f) \subset \mathcal{K}\},$$

where $supp(f)$ is the support of f.

Theorem 1.2. *Let* $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N}$ *be a random vector with joint law* $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ *such that* $\mathbb{E}^N[e^{t\|X\|}] < +\infty$, *for some* $t > 0$. *Let* $T = \sup\{t \ge 0 : \mathbb{E}^N\big[e^{t\|X\|}\big] < +\infty\}$ *and let* h^{-1} *be the inverse of*

$$h(s) = \int_{\mathbb{R}^{N^2}} \|u\|\big(e^{s\|u\|} - 1\big)\nu(du), \qquad 0 < s < T.$$

(i) For any compact set $\mathcal{K} \subset \mathbb{R}$,

$$\mathbb{P}^N\Big(\sup_{f\in Lip_{\mathcal{K}}(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N\left[tr_N(f(\mathbf{X_A}))\right]| \ge \delta\Big)$$

$$\le \frac{8|\mathcal{K}|}{\delta}\exp\left\{-\int_0^{\frac{N\delta^2}{8\sqrt{2}a|\mathcal{K}|}} h^{-1}(s)ds\right\}, \tag{1.4}$$

for all $\delta > 0$ *such that* $\delta^2 < 8\sqrt{2}a|\mathcal{K}|h\left(T^-\right)/N$.

(ii)

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N\left[tr_N(f(\mathbf{X_A}))\right]|\geq\delta\Big)$$
$$\leq\frac{C(\delta,b)}{\delta}\exp\bigg\{-\int_0^{\frac{N\delta^2}{\sqrt{2}aC(\delta,b)}}h^{-1}(s)ds\bigg\},\tag{1.5}$$

for all $\delta>0$ *such that* $\delta^2\leq\sqrt{2}aC(\delta,b)h(T^-)/N$, *where*

$$C(\delta,b)=C\bigg(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma)+h(t_0)\Big)+b\bigg),$$

with $G_2(\gamma)$ *as in Proposition 1.1,* C *a universal constant, and with* t_0 *the solution, in* t*, of* $th(t)-\int_0^t h(s)ds-\ln(12b/\delta)=0$.

Remark 1.3. *(i) The order of* $C(\delta,b)$ *in part (ii) can be made more specific. Indeed, it will be clear from the proof of this theorem (see (2.39)), that for any fixed* t^*, $0<t^*\leq T$,

$$C(\delta,b)\leq C\bigg(\frac{\sqrt{2}a}{\sqrt{N}}\Big(\frac{\ln\frac{12b}{\delta}}{t^*}+\frac{\int_0^{t^*}h(s)ds}{t^*}+G_2(\gamma)\Big)\bigg).$$

(ii) As seen from the proof (see (2.38)), in the statement of the above theorem, $G_2(\gamma)$ *can be replaced by* $\mathbb{E}^N\left[\|X\|\right]$. *Now* $\mathbb{E}^N\left[\|X\|\right]$ *is of order* N*, since*

$$N\min_{j=1,2,\ldots,N^2}\mathbb{E}^N\left[|X_j|\right]\leq\mathbb{E}^N\left[\|X\|\right]\leq N\max_{j=1,2,\ldots,N^2}\sqrt{\mathbb{E}^N\left[X_j^2\right]},\tag{1.6}$$

where the X_j, $j=1,2,\ldots,N^2$ *are the components of* X. *Actually, an estimate more precise than (1.6) is given by a result of Marcus and Rosiński [22] which asserts that if* $\mathbb{E}[X]=0$, *then*

$$\frac{1}{4}x_0\leq\mathbb{E}\left[\|X\|\right]\leq\frac{17}{8}x_0,$$

where x_0 *is the solution of the equation:*

$$\frac{V^2(x)}{x^2}+\frac{U(x)}{x}=1,\tag{1.7}$$

with $V^2(x)$ *as defined before and* $U(x)=\int_{\|u\|\geq x}\|u\|\nu(du)$, $x>0$.

(iii) As usual, one can easily pass from the mean $\mathbb{E}^N[tr_N(f)]$ *to any median* $m(tr_N(f))$ *in either (1.4) or (1.5). Indeed, for any* $0\leq\delta\leq 2b$, *if*

$$\sup_{f\in Lip_b(1)}|tr_N(f)-m(tr_N(f))|\geq\delta,$$

there exist a function $f\in Lip_b(1)$ *and a median* $m(tr_N(f))$ *of* $tr_N(f)$*, such that either* $tr_N(f)-m(tr_N(f))\geq\delta$ *or* $tr_N(f)-m(tr_N(f))\leq-\delta$. *Without loss of generality assuming the former, otherwise dealing with the latter with* $-f$*, consider the function* $g(y)=\min\left(d(y,A),\delta\right)/2$, $y\in\mathbb{R}^{N^2}$, *where* $A=\{tr_N(f)\leq m(tr_N(f))\}$. *Clearly*

$g \in Lip_b(1)$, $\mathbb{E}^N[tr_N(g)] \leq \delta/4$, and therefore $tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right] \geq \delta/4$, which indicates that

$$\sup_{g \in Lip_b(1)} \left| tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right] \right| \geq \frac{\delta}{4}.$$

Hence,

$$\mathbb{P}^N\left(\sup_{f \in Lip_b(1)} \left| tr_N(f) - m(tr_N(f)) \right| \geq \delta \right)$$
$$\leq \mathbb{P}^N\left(\sup_{g \in Lip_b(1)} \left| tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right] \right| \geq \frac{\delta}{4} \right). \tag{1.8}$$

(iv) When the entries of X are independent, and under a finite exponential moment assumption, the dependency in N of the function h (above and below) can sometimes be improved. We refer the reader to [15] where some of these generic problems are discussed and tackled.

Next, recall (see [7], [19]) that the Wasserstein distance between any two probability measures μ_1 and μ_2 on $\mathbb{R}$ is defined by

$$d_W(\mu_1, \mu_2) = \sup_{f \in Lip_b(1)} \left| \int_{\mathbb{R}} f d\mu_1 - \int_{\mathbb{R}} f d\mu_2 \right|. \tag{1.9}$$

Hence, Theorem 1.2 actually gives a concentration result, with respect to the Wasserstein distance, for the empirical spectral measure $\hat{\mu}^N_{\mathbf{A}}$, when it deviates from its mean $\mathbb{E}^N[\hat{\mu}^N_{\mathbf{A}}]$.
As in [10], we can also obtain a concentration result for the distance between any particular probability measure and the empirical spectral measure.

Proposition 1.4. *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \leq i < j \leq N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that $\mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty$, for some $t > 0$. Let $T = \sup\{t > 0 : \mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty\}$ and let h^{-1} be the inverse of $h(s) = \int_{\mathbb{R}^{N^2}} \|u\|(e^{s\|u\|} - 1)\nu(du)$, $0 < s < T$. Then, for any probability measure μ,*

$$\mathbb{P}^N\left(d_W(\hat{\mu}^N_A, \mu) - \mathbb{E}^N[d_W(\hat{\mu}^N_A, \mu)] \geq \delta \right) \leq \exp\left\{ - \int_0^{\frac{N\delta}{\sqrt{2a}}} h^{-1}(s) ds \right\}, \tag{1.10}$$

for all $0 < \delta < \sqrt{2}ah\left(T^-\right)/N$.

Of particular importance is the case of an infinitely divisible vector having boundedly supported Lévy measure. We then have:

Corollary 1.5. *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \leq i < j \leq N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that ν has bounded support. Let $R = \inf\{r > 0 : \nu(x : \|x\| > r) = 0\}$, let $V^2\big(= V^2(R)\big) = \int_{\mathbb{R}^{N^2}} \|u\|^2 \nu(du)$, and for $x > 0$ let*

$$\ell(x) = (1 + x)\ln(1 + x) - x.$$

(i) For any $\delta > 0$,

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]| \geq \delta\Big)$$
$$\leq \frac{C(\delta,b)}{\delta}\exp\Bigg\{-\frac{V^2}{R^2}\ell\Big(\frac{NR\delta^2}{\sqrt{2}aC(\delta,b)V^2}\Big)\Bigg\}, \tag{1.11}$$

where

$$C(\delta,b) = C\bigg(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma) + \frac{V^2}{R}(e^{t_0R}-1)\Big) + b\bigg),$$

with $G_2(\gamma)$ as in Proposition 1.1, C a universal constant, and t_0 the solution, in t, of

$$\frac{V^2}{R^2}\Big(tRe^{tR} - e^{tR} + 1\Big) = \ln\frac{12b}{\delta}.$$

(ii) For any probability measure μ on $\mathbb{R}$, and any $\delta > 0$,

$$\mathbb{P}^N\big(d_W(\hat{\mu}_A^N,\mu) - \mathbb{E}^N[d_W(\hat{\mu}_A^N,\mu)] \geq \delta\big)$$
$$\leq \exp\Bigg\{\frac{N\delta}{\sqrt{2}aR} - \bigg(\frac{N\delta}{\sqrt{2}aR} + \frac{V^2}{R^2}\bigg)\ln\bigg(1 + \frac{NR\delta^2}{\sqrt{2}aV^2}\bigg)\Bigg\}. \tag{1.12}$$

Remark 1.6. *(i) As in Theorem 1.2, the dependency of $C(\delta,b)$ in δ and b can be made more precise. A key step in the proof of (1.11) is to choose τ such that*

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \delta/12b,$$

and then $C(\delta,b)$ is determined by τ. Minimizing, in t, the right hand side of (2.38), leads to the following estimate

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \exp\Bigg\{-\frac{V^2}{R^2}\ell\Bigg(\frac{R\Big(\frac{\sqrt{N}}{\sqrt{2}a}\tau - G_2(\gamma)\Big)}{V^2}\Bigg)\Bigg\},$$

where $\ell(x) = (1+x)\ln(1+x) - x$. For $x \geq 1$, $2\ell(x) \geq x\ln x$. Hence one can choose τ to be the solution, in x, of the equation

$$\frac{x}{R}\ln\frac{xR}{V^2} = 2\ln\frac{12b}{\delta}.$$

It then follows that $C(\delta,b)$ can be taken to be

$$C\bigg(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma)+\tau\Big) + b\bigg).$$

Without the finite exponential moment assumption, an interesting class of random matrices with infinitely divisible entries are the ones with stable entries, which we now analyze.

Recall that X in $\mathbb{R}^d$ is α-stable, $(0 < \alpha < 2)$, if its Lévy measure ν is given, for any Borel set $B \in \mathcal{B}(\mathbb{R}^d)$, by

$$\nu(B) = \int_{S^{d-1}}\sigma(d\xi)\int_0^{+\infty}\mathbf{1}_B(r\xi)\frac{dr}{r^{1+\alpha}}, \tag{1.13}$$

where σ, the spherical component of the Lévy measure, is a finite positive measure on S^{d-1}, the unit sphere of $\mathbb{R}^d$. Since the expected value of the spectral measure of a matrix with α-stable entries might not exist, we study the deviation from a median. Here is a sample result.

Theorem 1.7. *Let $0 < \alpha < 2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (1.13).*

(i) Let $f \in Lip(1)$, and let $m(tr_N(f(\mathbf{X_A})))$ be any median of $tr_N(f(\mathbf{X_A}))$. Then,

$$\mathbb{P}^N\big(tr_N(f(\mathbf{X_A})) - m(tr_N(f(\mathbf{X_A})))\ge\delta\big)\le C(\alpha)(\sqrt{2}a)^\alpha\frac{\sigma(S^{N^2-1})}{N^\alpha\delta^\alpha}, \tag{1.14}$$

whenever $\delta N > \sqrt{2}a\left[2\sigma(S^{N^2-1})C(\alpha)\right]^{1/\alpha}$, and where $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

(ii) Let $\lambda_{max}(\mathbf{X_A})$ be the largest eigenvalue of $\mathbf{X_A}$, and let $m(\lambda_{max}(\mathbf{X_A}))$ be any median of $\lambda_{max}(\mathbf{X_A})$, then

$$\mathbb{P}^N\left(\lambda_{max}(\mathbf{X_A}) - m(\lambda_{max}(\mathbf{X_A})) \ge \delta\right) \le C(\alpha)(\sqrt{2}a)^\alpha\frac{\sigma(S^{N^2-1})}{N^{\alpha/2}\delta^\alpha}, \tag{1.15}$$

whenever $\delta\sqrt{N} > \sqrt{2}a\left[2\sigma(S^{N^2-1})C(\alpha)\right]^{1/\alpha}$, and where $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Remark 1.8. *Let $\mathbf{M}$ be a Wigner matrix whose entries $\mathbf{M}_{i,i}, 1\le i\le N, \mathbf{M}^R_{i,j}, 1\le i<j\le N$, and $\mathbf{M}^I_{i,j}$, $1\le i<j\le N$, are iid random variables, such that the distribution of $|\mathbf{M}_{1,1}|$ belongs to the domain of attraction of an α-stable distribution, i.e., for any $\delta > 0$,*

$$\mathbb{P}(|\mathbf{M}_{1,1}| > \delta) = \frac{L(\delta)}{\delta^\alpha},$$

for some slowly varying positive function L such that $\lim_{\delta\to\infty} L(t\delta)/L(\delta) = 1$, for all $t > 0$. Soshnikov [30] showed that, for any $\delta > 0$,

$$\lim_{N\to\infty}\mathbb{P}^N(\lambda_{max}(b_N^{-1}\mathbf{M}) \ge \delta) = 1 - \exp(-\delta^{-\alpha}),$$

where b_N is a normalizing factor such that $\lim_{N\to\infty} N^2L(b_N)/b_N^\alpha = 2$ and where $\lambda_{max}(b_N^{-1}\mathbf{M})$ is the largest eigenvalue of $b_N^{-1}\mathbf{M}$. In fact $\lim_{N\to\infty} N^{\frac{2}{\alpha}-\epsilon}/b_N = 0$ and $\lim_{N\to\infty} b_N/N^{\frac{2}{\alpha}+\epsilon} = 0$, for any $\epsilon > 0$. As stated in [13], when the random vector X is in the domain of attraction of an α-stable distribution, concentration inequalities similar to (1.14) or (1.15) can be obtained for general Lipschitz function. In particular, if the Lévy measure of X is given by

$$\nu(B) = \int_{S^{N^2-1}}\sigma(d\xi)\int_0^{+\infty}\mathbf{1}_B(r\xi)\frac{L(r)dr}{r^{1+\alpha}}, \tag{1.16}$$

for some slowly varying function L on $[0, +\infty)$, and if we still choose the normalizing factor b_N such that $\lim_{N\to\infty}\sigma(S^{N^2-1})L(b_N)/b_N^\alpha$ is constant, then,

$$\mathbb{P}^N\big(\lambda_{max}(b_N^{-1}\mathbf{M}) - m(\lambda_{max}(b_N^{-1}\mathbf{M})) \geq \delta\big)$$
$$\leq \frac{C(\alpha)\sigma(S^{N^2-1})2^{\alpha/2}}{b_N^\alpha}\frac{L\Big(b_N\frac{\delta}{\sqrt{2}}\Big)}{\delta^\alpha}, \tag{1.17}$$

whenever

$$(\delta b_N)^\alpha \geq 2^{1+\alpha/2}C(\alpha)\sigma(S^{N^2-1})L\big(b_N\delta/\sqrt{2}\big).$$

Now, recall that for an N^2 dimensional vector with iid entries, $\sigma(S^{N^2-1}) = N^2(\hat{\sigma}(1) + \hat{\sigma}(-1))$, where $\hat{\sigma}(1)$ is short for $\sigma(1, 0, \ldots, 0)$ and similarly for $\hat{\sigma}(-1)$. Thus, for fixed N, our result gives the correct order of the upper bound for large values of δ, since for $\delta > 1$,

$$\frac{e-1}{e\delta^\alpha} \leq 1 - e^{-\delta^{-\alpha}} \leq \frac{1}{\delta^\alpha}.$$

Moreover, in the stable case, $L(\delta)$ becomes constant, and $b_N = N^{2/\alpha}$. Since $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$ is a Lipschitz function of the entries of the matrix $\mathbf{M}$ with Lipschitz constant at most $\sqrt{2}N^{-2/\alpha}$, for any median $m(\lambda_{max}(N^{-2/\alpha}\mathbf{M}))$ of $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$, we have,

$$\mathbb{P}^N\big(\lambda_{max}(N^{-\frac{2}{\alpha}}\mathbf{M}) - m(\lambda_{max}(N^{-\frac{2}{\alpha}}\mathbf{M})) \geq \delta\big) \leq C(\alpha)\frac{\big(\hat{\sigma}(1)+\hat{\sigma}(-1)\big)}{2^{\alpha/2}}\frac{1}{\delta^\alpha}, \tag{1.18}$$

whenever $\delta \geq \big[2C(\alpha)\big(\hat{\sigma}(1) + \hat{\sigma}(-1)\big)\big]^{1/\alpha}$. Furthermore, using Theorem 1 in [14], it is not difficult to see that $m(\lambda_{max}(N^{-2/\alpha}\mathbf{M}))$ can be upper and lower bounded independently of N. Finally, an argument as in Remark 1.15 below will give a lower bound on $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$ of the same order as (1.18).

The following proposition will give an estimate on any median of a Lipschitz function of X, where X is a stable vector. It is the version of Proposition 1.1 for α-stable vectors.

Proposition 1.9. *Let $0 < \alpha < 2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\leq i<j\leq N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (1.13). Let $f \in Lip(1)$, then*

(i) any median $m(f(X))$ of $f(X)$ satisfies

$$|m(f(X)) - f(0)|$$
$$\leq J_1(\alpha) := \left(\frac{\sigma(S^{N^2-1})}{4\alpha}\right)^{1/\alpha}\left(\sqrt{\frac{\alpha}{4(2-\alpha)}} + 3k_{\frac{\alpha}{4(2-\alpha)}}(1/4)\right) + E, \tag{1.19}$$

(ii) the mean $\mathbb{E}^N[f(X)]$ of $f(X)$, if it exists, satisfies

$$|\mathbb{E}^N[f(X)] - f(0)|$$
$$\leq J_2(\alpha) := \left(\frac{\sigma(S^{N^2-1})}{4\alpha}\right)^{1/\alpha}\left(\sqrt{\frac{\alpha}{4(2-\alpha)}} + k_{\frac{\alpha}{4(2-\alpha)}}(1/4)\right) + E, \tag{1.20}$$

where $k_{\alpha/4(2-\alpha)}(x)$, $x > 0$, *is the solution, in* y, *of the equation*

$$y - \left(y + \frac{\alpha}{4(2-\alpha)}\right) \ln\left(1 + \frac{4(2-\alpha)y}{\alpha}\right) = \ln x,$$

and where

$$E = \left(\sum_{k=1}^{N^2} \left(\langle e_k, \beta \rangle - \int_{\left(\frac{4\sigma(S^{N^2-1})}{\alpha}\right)^{1/\alpha} < \|y\| \le 1} \langle e_k, y \rangle \nu(dy) \right.\right.$$
$$\left.\left. + \int_{1 < \|y\| \le \left(\frac{4\sigma(S^{N^2-1})}{\alpha}\right)^{1/\alpha}} \langle e_k, y \rangle \nu(dy) \right)^2 \right)^{1/2}, \tag{1.21}$$

with $e_1, e_2, \ldots, e_{N^2}$ *being the canonical basis of* $\mathbb{R}^{N^2}$.

Remark 1.10. *When the components of* X *are independent, a direct computation shows that, up to a constant,* E *in both* $J_1(\alpha)$ *and* $J_2(\alpha)$ *is dominated by* $\left(\frac{\sigma(S^{N^2-1})}{4\alpha}\right)^{1/\alpha}$, *as* $N \to \infty$.

In complete similarity to the finite exponential moments case, we can obtain concentration results for the spectral measure of matrices with α-stable entries.

Theorem 1.11. *Let* $0 < \alpha < 2$, *and let* $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \le i < j \le N}$ *be an* α-*stable random vector in* $\mathbb{R}^{N^2}$, *with Lévy measure* ν *given by (1.13).*

(i) Then,

$$\mathbb{P}^N\left(\sup_{f \in Lip_b(1)} \left| tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] \right| \ge \delta \right)$$
$$\le C(\delta, b, \alpha) \frac{a^\alpha \sigma(S^{N^2-1})}{N^\alpha \delta^\alpha} \wedge 1, \tag{1.22}$$

where

$$C(\delta, b, \alpha) = \left(C_1(\alpha) \left(\frac{\sqrt{2}a}{\sqrt{N}}\right)^{1+\alpha} \left(\frac{J_1(\alpha)+1}{\delta} + b\right)^{1+\alpha} + C_2(\alpha) \right),$$

with $C_1(\alpha)$ *and* $C_2(\alpha)$ *constants depending only on* α, *and with* $J_1(\alpha)$ *as in Proposition 1.9.*

(ii) For any probability measure μ,

$$\mathbb{P}^N\big(d_W(\hat{\mu}^N_A, \mu) - m(d_W(\hat{\mu}^N_A, \mu)) \ge \delta\big) \le C(\alpha)(\sqrt{2}a)^\alpha \frac{\sigma(S^{N^2-1})}{N^\alpha \delta^\alpha}, \tag{1.23}$$

whenever $\delta N \ge \sqrt{2}a \left[2\sigma(S^{N^2-1})C(\alpha)\right]^{1/\alpha}$ *and where* $C(\alpha) = 4^\alpha(2 - \alpha + e\alpha)/\alpha(2-\alpha)$.

It is also possible to obtain concentration results for smaller values of δ. Indeed, the lower and intermediate range for the stable deviation obtained in [5] provide the appropriate tools to achieve such results. We refer to [5] for complete arguments, and only provide below a sample result.

Theorem 1.12. *Let $0 < \alpha < 2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (1.13). For any $\epsilon > 0$, there exists $\eta(\epsilon)$, and constants $D_1 = D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$ and $D_2 = D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$, such that for all $0 < \delta < \eta(\epsilon)$,*

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} \big|tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]\big| \ge \delta\Big) \le (1+\epsilon)\frac{D_1}{\delta^{\frac{\alpha+1}{\alpha}}} \exp\Big(-D_2\delta^{\frac{2\alpha+1}{\alpha-1}}\Big). \tag{1.24}$$

Remark 1.13. *(i) In (1.14), (1.15) or (1.23), the constant $C(\alpha)$ is not of the right order as $\alpha \to 2$. It is, however, a simple matter to adapt Theorem 2 of [13] to obtain, at the price of worsening the range of validity of the concentration inequalities, the right order in the constants as $\alpha \to 2$.*

(ii) Let us now provide some estimation of D_1 and D_2, which are needed for comparison with the GUE results of [10] (see (iii) below). Let $C(\alpha) = 2^\alpha(e\alpha + 2 - \alpha)/(2(2-\alpha))$, $K(\alpha) = \max\big\{2^\alpha/(\alpha-1), C(\alpha)\big\}$, $L(\alpha) = \big((\alpha-1)/\alpha\big)^{\alpha/(\alpha-1)}(2-\alpha)/10$ and let

$$D^* = 2\left(\frac{\sqrt{2}a}{\sqrt{N}}\right)^{\frac{2\alpha-1}{\alpha}}\left(12\frac{C(\alpha)}{K(\alpha)}\right)^{\frac{1}{\alpha}} J_2(\alpha) b^{\frac{1}{\alpha}} + 2\left(\frac{\sqrt{2}a}{\sqrt{N}}\right)^{\frac{\alpha-1}{\alpha}}\left(12\frac{C(\alpha)}{K(\alpha)}\right)^{\frac{1}{\alpha}} b^{\frac{\alpha+1}{\alpha}} + \frac{2\sqrt{2}a}{\sqrt{N}}\Big(12C(\alpha)\sigma(S^{N^2-1})\Big)^{\frac{1}{\alpha}} b^{\frac{1}{\alpha}}. \tag{1.25}$$

As shown in the proof of the theorem, $D_1 = 24D^*$, while

$$D_2 = \frac{L(\alpha)}{\Big(\sigma(S^{N^2-1})\Big)^{\frac{1}{\alpha-1}}}\left(\frac{N}{\sqrt{2}a}\right)^{\frac{\alpha}{\alpha-1}}\frac{1}{(72D^*)^{\frac{\alpha}{\alpha-1}}}.$$

Thus, as $N \to +\infty$, D_1 is of order $N^{-1/2}\Big(\sigma(S^{N^2-1})\Big)^{1/\alpha}$, while D_2 is of order $N^{3\alpha/(2\alpha-2)}\Big(\sigma(S^{N^2-1})\Big)^{2/(1-\alpha)}$.

(iii) Guionnet and Zeitouni [10], obtained concentration results for the spectral measure of matrices with independent entries, which are either compactly supported or satisfy a logarithmic Sobolev inequality. In particular for the elements of the GUE, their upper bound of concentration for the spectral measure is

$$\frac{C_1 + b^{3/2}}{\delta^{3/2}} \exp\Big\{ -\frac{C_2}{8ca^2}N^2\frac{\delta^5}{(C_1+b^{3/2})^2}\Big\}, \tag{1.26}$$

where C_1 and C_2 are universal constants. In Theorem 1.12, the order, in b, of D_1 is at most $b^{\alpha+1/\alpha}$, while that of D_2 is at least $b^{-(\alpha+1)/(\alpha-1)}$. For α close to 2, this order is thus consistent with the one in (1.26). Taking into account part (ii) above, the order of the constants in (1.24) are correct when $\alpha \to 2$. Following [5] (see also Remark 4 in [21]), we can recover a suboptimal Gaussian result by considering a particular stable random vector $X^{(\alpha)}$ and letting $\alpha \to 2$. Toward this end, let $X^{(\alpha)}$ be the stable random vector whose Lévy measure has for spherical component σ, the uniform measure with total mass $\sigma(S^{N^2-1}) = N^2(2-\alpha)$. As α converges to 2, $X^{(\alpha)}$ converges in distribution to a standard normal random vector. Also, as $\alpha \to 2$, the range of δ in Theorem 1.12 becomes $(0, +\infty)$ while the constants in the concentration bound do converge. Thus, the right hand side of (1.24) becomes

$$\frac{D_1}{\delta^{3/2}} \exp\left\{ - D_2\delta^5 \right\},$$

which is of the same order, in δ, as (1.26). However our order in N is suboptimal.

(iv) *In the proof of Theorem 1.12, the desired estimate in (2.56) is achieved through a truncation of order $\delta^{-1/\alpha}$, which, when $\alpha \to 2$, is of the same order as the one used in obtaining (1.26). However, for the GUE result, using Gaussian concentration, a truncation of order $\sqrt{\ln(12b/\delta)}$ gives a slightly better bound, namely,*

$$\frac{C_1\sqrt{\ln \frac{12b}{\delta}}}{\delta} \exp\left\{ - \frac{C_2 N^2 \delta^4}{8ca^2 \ln \frac{12b}{\delta}} \right\},$$

where C_1 and C_2 are absolute constants (different from those of (1.26)).

Wishart matrices are of interest in many contexts, in particular as sample covariance matrices in statistics. Recall that $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$ is a complex Wishart matrix if $\mathbf{Y}$ is a $K \times N$ matrix, $K > N$, with entries $\mathbf{Y}_{i,j} = \mathbf{Y}^R_{i,j} + \sqrt{-1}\mathbf{Y}^I_{i,j}$ (a real Wishart matrix is defined similarly with $\mathbf{Y}^I_{i,j} = \delta_0$ and $\mathbf{M} = \mathbf{Y}^t\mathbf{Y}$). Recall also that if the entries of $\mathbf{Y}$ are iid centered random variables with finite variance σ^2, the empirical distribution of the eigenvalues of $\mathbf{Y}^*\mathbf{Y}/N$ converges as $K \to \infty$, $N \to \infty$, and $K/N \to \gamma \in (0, +\infty)$ to the Marčenko-Pastur law ([4], [23]) with density

$$p_\gamma(x) = \frac{1}{2\pi x\gamma\sigma^2}\sqrt{(c_2 - x)(x - c_1)}, \quad c_1 \le x \le c_2,$$

where $c_1 = \sigma^2(1-\gamma^{-1/2})^2$ and $c_2 = \sigma^2(1+\gamma^{-1/2})^2$. When the entries of $\mathbf{Y}$ are iid normal, Johansson [16] and Johnstone [17] showed, in the complex and real case respectively, that the properly normalized largest eigenvalue converges in distribution to the Tracy-Widom law ([33], [34]). Soshnikov [29] extended the results of Johansson and Johnstone to Wishart matrices with symmetric subgaussian entries under the condition that $K - N = O(N^{1/3})$. (As kindly indicated to us by a referee, this last condition has recently been removed by Péché). Soshnikov and Fyodorov [32] studied the distribution of the largest eigenvalue of the Wishart matrix $\mathbf{Y}^*\mathbf{Y}$, when the entries of $\mathbf{Y}$ are iid Cauchy random variables. We are interested here in concentration for the linear statistics of the spectral measure and for the largest eigenvalue of the Wishart matrix $\mathbf{Y}^*\mathbf{Y}$, where the entries of $\mathbf{Y}$ form an infinitely divisible vector and, in particular, a stable one. We restrict our work to the complex framework, the real framework being essentially the same.

It is not difficult to see that if $\mathbf{Y}$ has iid Gaussian entries, $\mathbf{Y}^*\mathbf{Y}$ has infinitely divisible entries, each with a Lévy measure without a known explicit form. However the dependence structure among the entries of $\mathbf{Y}^*\mathbf{Y}$ prevents the vector of entries to be, itself, infinitely divisible (this is a well known fact originating with Lévy, see [27]). Thus the methodology, we used till this point, cannot be directly applied to deal with functions of eigenvalues of $\mathbf{Y}^*\mathbf{Y}$. However, concentration results can be obtained when we consider the following facts, due to Guionnet and Zeitouni [10] and already used for that purpose in their paper.

Let

$$A_{i,j} = \begin{cases} 0 & \text{for } 1 \le i \le K, 1 \le j \le K \\ 0 & \text{for } N+1 \le i \le K+N, K+1 \le j \le K+N \\ 1 & \text{for } 1 \le i \le K, K+1 \le j \le K+N \\ 1 & \text{for } N+1 \le i \le K+N, 1 \le j \le K, \end{cases} \tag{1.27}$$

and

$$\omega_{i,j} = \begin{cases} 0 & \text{for } 1 \le i \le K, 1 \le j \le K \\ 0 & \text{for } N+1 \le i \le K+N, K+1 \le j \le K+N \\ \bar{Y}_{i,j} & \text{for } 1 \le i \le K, K+1 \le j \le K+N \\ Y_{i,j} & \text{for } N+1 \le i \le K+N, 1 \le j \le K, \end{cases} \tag{1.28}$$

then $\mathbf{X_A} = \begin{pmatrix} \mathbf{0} & \mathbf{Y}^* \\ \mathbf{Y} & \mathbf{0} \end{pmatrix} \in \mathcal{M}_{(K+N)\times(K+N)}(\mathbb{C})$, and

$$\mathbf{X}_\mathbf{A}^2 = \begin{pmatrix} \mathbf{Y}^*\mathbf{Y} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}\mathbf{Y}^* \end{pmatrix}.$$

Moreover, since the spectrum of $\mathbf{Y}^*\mathbf{Y}$ differs from that of $\mathbf{Y}\mathbf{Y}^*$ only by the multiplicity of the zero eigenvalue, for any function f, one has

$$tr(f(\mathbf{X}_\mathbf{A}^2)) = 2tr(f(\mathbf{Y}^*\mathbf{Y})) + (K-N)f(0),$$

and

$$\lambda_{max}(\mathbf{M}^{1/2}) = \max_{1\le i\le N} |\lambda_i(\mathbf{X_A})|,$$

where $\mathbf{M}^{1/2}$ is the unique positive semi-definite square root of $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$.

Next let $\mathbb{P}^{K,N}$ be the joint law of $(\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le K, 1\le j\le N}$ on $\mathbb{R}^{2KN}$, and let $\mathbb{E}^{K,N}$ be the corresponding expectation. We present below, in the infinitely divisible case, a concentration result for the largest eigenvalue $\lambda_{max}(\mathbf{M})$, of the Wishart matrices $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$. The concentration for the linear statistic $tr_N(f(\mathbf{M}))$ could also be obtained using the above observations.

Corollary 1.14. *Let $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$, with $\mathbf{Y}_{i,j} = \mathbf{Y}^R_{i,j} + \sqrt{-1}\mathbf{Y}^I_{i,j}$.*

(i) Let $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le N, 1\le j\le K}$ be a random vector with joint law $\mathbb{P}^{K,N} \sim ID(\beta, 0, \nu)$ such that $\mathbb{E}^{K,N}[e^{t\|X\|}] < +\infty$, for some $t > 0$. Let $T = \sup\{t > 0 : \mathbb{E}^{K,N}[e^{t\|X\|}] < +\infty\}$ and let h^{-1} be the inverse of

$$h(s) = \int_{\mathbb{R}^{2KN}} \|u\|(e^{s\|u\|} - 1)\nu(du), \qquad 0 < s < T.$$

Then,

$$\mathbb{P}^{K,N}\left(\lambda_{max}(\mathbf{M}^{1/2}) - \mathbb{E}^{K,N}[\lambda_{max}(\mathbf{M}^{1/2})] \geq \delta\right) \leq e^{-\int_0^{\delta/\sqrt{2}} h^{-1}(s)ds}, \tag{1.29}$$

for all $0 < \delta < h(T^-)$.

(ii) Let $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\leq i\leq K, 1\leq j\leq N}$ *be an* α*-stable random vector with Lévy measure* ν *given by* $\nu(B) = \int_{S^{2KN-1}} \sigma(d\xi) \int_0^{+\infty} \mathbf{1}_B(r\xi) dr/r^{1+\alpha}$. *Then,*

$$\mathbb{P}^{K,N}\Big(\lambda_{max}(\mathbf{M}^{1/2}) - m(\lambda_{max}(\mathbf{M}^{1/2})) \geq \delta\Big) \leq C(\alpha)(\sqrt{2})^\alpha \frac{\sigma(S^{2KN-1})}{\delta^\alpha},$$

whenever $\delta > \sqrt{2}a\left[2\sigma(S^{2KN-1})C(\alpha)\right]^{1/\alpha}$ *and where* $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Remark 1.15. *(i) As already mentioned, Soshnikov and Fyodorov ([32]) obtained the asymptotic behavior of the largest eigenvalue of the Wishart matrix* $\mathbf{Y}^*\mathbf{Y}$ *when the entries of, the* $K \times N$ *matrix,* $\mathbf{Y}$ *are iid Cauchy random variables. They further argue that although the typical eigenvalues of* $\mathbf{Y}^*\mathbf{Y}$ *are of order* KN*, the correct order for the largest one is* K^2N^2*. The above corollary combined with Remark 1.10 and the estimate* (1.19)*, shows that when the entries of* $\mathbf{Y}$ *form an* α*-stable random vector, the largest eigenvalue of* $\mathbf{Y}^*\mathbf{Y}$ *is of order at most* $\sigma(S^{2KN-1})^{2/\alpha}$*. There is also a lower concentration result, described next, which leads to a lower bound on the order of this largest eigenvalue. Thus, from these two estimates, if the entries of* $\mathbf{Y}$ *are iid* α*-stable, the largest eigenvalue of* $\mathbf{Y}^*\mathbf{Y}$ *is of order* $K^{2/\alpha}N^{2/\alpha}$.

(ii) Let $X \sim ID(\beta, 0, \nu)$ *in* $\mathbb{R}^d$*, then (see Lemma 5.4 in [6]) for any* $x > 0$*, and any norm* $\|\cdot\|_{\mathcal{N}}$ *on* $\mathbb{R}^d$*,*

$$\mathbb{P}\big(\|X\|_{\mathcal{N}} \geq x\big) \geq \frac{1}{4}\Big(1 - \exp\Big\{-\nu\big(\{u \in \mathbb{R}^d : \|u\|_{\mathcal{N}} \geq 2x\}\big)\Big\}\Big).$$

But, $\lambda_{max}(\mathbf{M}^{1/2})$ *is a norm of the vector* $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})$*, which we denote by* $\|X\|_\lambda$*, if* X *is a stable vector in* $\mathbb{R}^{2KN}$*.*

$$\begin{aligned}
&\mathbb{P}^{K,N}\Big(\lambda_{max}(\mathbf{M}^{1/2}) - m(\lambda_{max}(\mathbf{M}^{1/2})) \geq \delta\Big) \\
&= \mathbb{P}^{K,N}\Big(\lambda_{max}(\mathbf{M}^{1/2}) \geq \delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\Big) \\
&\geq \frac{1}{4}\Big(1 - \exp\Big\{-\nu\big(\{\lambda_{max}(\mathbf{M}^{1/2}) \geq 2\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)\}\big)\Big\}\Big) \\
&\geq \frac{1}{4}\Big(1 - \exp\Big\{-\nu\big(\{\|X\|_\lambda \geq 2\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)\}\big)\Big\}\Big) \\
&= \frac{1}{4}\Big(1 - \exp\Big\{-\frac{\tilde{\sigma}\big(S^{2KN-1}_{\|\cdot\|_\lambda}\big)}{\alpha\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)^\alpha}\Big\}\Big),
\end{aligned} \tag{1.30}$$

where $S^{2KN-1}_{\|\cdot\|_\lambda}$ *is the unit sphere relative to the norm* $\|\cdot\|_\lambda$ *and where* $\tilde{\sigma}$ *is the spherical part of the Lévy measure corresponding to this norm. Moreover, if the components of* X *are independent, in which case the Lévy measure is supported on the axes of* $\mathbb{R}^{2KN}$*,* $\tilde{\sigma}\big(S^{2KN-1}_{\|\cdot\|_\lambda}\big)$ *is of order* KN*, and so, as above, the largest eigenvalue of* $\mathbf{M}^{1/2}$ *is of order* $K^{1/\alpha}N^{1/\alpha}$*.*

(iii) *For any function f such that $g(x) = f(x^2)$ is Lipschitz with Lipschitz constant $\|g\|_{Lip} := |||f|||_{\mathcal{L}}$, $tr(g(\mathbf{X_A})) = tr(f(\mathbf{X_A^2}))$ is a Lipschitz function of the entries of $\mathbf{Y}$ with Lipschitz constant at most $\sqrt{2}|||f|||_{\mathcal{L}}\sqrt{K+N}$. Hence, under the assumptions of part (i) of Corollary 1.14,*

$$\mathbb{P}^{K,N}\Big(tr_N(f(\mathbf{M})) - \mathbb{E}^{K,N}[tr_N(f(\mathbf{M}))] \geq \delta\frac{K+N}{N}\Big) \leq \exp\Big\{-\int_0^{\sqrt{2(K+N)}\delta/|||f|||_{\mathcal{L}}} h^{-1}(s)ds\Big\}, \tag{1.31}$$

for all $0 < \delta < |||f|||_{\mathcal{L}} h(T^-)/\sqrt{2(K+N)}$.

(iv) *Under the assumptions of part (ii) of Corollary 1.14, for any function f such that $g(x) = f(x^2)$ is Lipschitz with $\|g\|_{Lip} = |||f|||_{\mathcal{L}}$, and any median $m(tr_N(f(\mathbf{M})))$ of $tr_N(f(\mathbf{M}))$ we have:*

$$\mathbb{P}^{K,N}\left(tr_N(f(\mathbf{M})) - m(tr_N(f(\mathbf{M}))) \geq \delta\frac{K+N}{N}\right) \leq C(\alpha)\frac{|||f|||_{\mathcal{L}}^{\alpha}}{\sqrt{2^{\alpha}(K+N)^{\alpha}}}\frac{\sigma(S^{2KN-1})}{\delta^{\alpha}}, \tag{1.32}$$

whenever $\delta > |||f|||_{\mathcal{L}}\left[2\sigma(S^{2KN-1})C(\alpha)\right]^{1/\alpha}/\sqrt{2(K+N)}$, and where $C(\alpha) = 4^{\alpha}(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Remark 1.16. *In the absence of finite exponential moments, the methods described in the present paper extend beyond the heavy tail case and apply to any random matrix whose entries on and above the main diagonal form an infinitely divisible vector X. However, to obtain explicit concentration estimates, we do need explicit bounds on V^2 and on $\bar{\nu}$. Such bounds are not always available when further knowledge on the Lévy measure of X is lacking.*

2 Proofs:

We start with a proposition, which is a direct consequence of the concentration inequalities obtained in [12] for general Lipschitz function of infinitely divisible random vectors with finite exponential moment.

Proposition 2.1. *Let $X = (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)_{1\leq i<j\leq N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that $\mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty$, for some $t > 0$ and let $T = \sup\{t > 0 : \mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty\}$. Let h^{-1} be the inverse of*

$$h(s) = \int_{\mathbb{R}^{N^2}} \|u\|\big(e^{s\|u\|} - 1\big)\nu(du), \quad 0 < s < T.$$

(i) *For any Lipschitz function f,*

$$\mathbb{P}^N\big(tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] \geq \delta\big) \leq \exp\left\{-\int_0^{\frac{N\delta}{\sqrt{2}a\|f\|_{Lip}}} h^{-1}(s)ds\right\},$$

for all $0 < \delta < \sqrt{2}a\|f\|_{Lip} h(T^-)/N$.

(ii) Let $\lambda_{max}(\mathbf{X_A})$ be the largest eigenvalue of the matrix $\mathbf{X_A}$. Then,

$$\mathbb{P}^N\left(\lambda_{max}(\mathbf{X_A}) - \mathbb{E}^N[\lambda_{max}(\mathbf{X_A})] \geq \delta\right) \leq \exp\left\{ - \int_0^{\frac{\sqrt{N}\delta}{\sqrt{2}a}} h^{-1}(s)ds \right\},$$

for all $0 < \delta < \sqrt{2}ah\left(T^-\right)/\sqrt{N}$.

Proof of Theorem 1.2:

For part (i), following the proof of Theorem 1.3 of [10], without loss of generality, by shift invariance, assume that $\min\{x : x \in \mathcal{K}\} = 0$. Next, for any $v > 0$, let

$$g_v(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < v \\ v & \text{if } x \geq v. \end{cases} \tag{2.33}$$

Clearly $g_v \in Lip(1)$ with $\|g_v\|_\infty = v$. Next for any function $f \in Lip_{\mathcal{K}}(1)$, any $\Delta > 0$, define recursively $f_\Delta(x) = 0$ for $x \leq 0$, and for $(j-1)\Delta \leq x \leq j\Delta$, $j = 1, \ldots, \lceil \frac{x}{\Delta} \rceil$, let

$$f_\Delta(x) = \sum_{j=1}^{\lceil \frac{x}{\Delta} \rceil} g_\Delta^{(j)},$$

where $g_\Delta^{(j)} := (2\mathbf{1}_{\{f(j\Delta) > f_\Delta((j-1)\Delta)\}} - 1)g_\Delta(x - (j-1)\Delta)$. Then $|f - f_\Delta| \leq \Delta$ and the 1-Lipschitz function f_Δ is the sum of at most $|\mathcal{K}|/\Delta$ functions $g_\Delta^{(j)} \in Lip(1)$, regardless of the function f. Now, for $\delta > 2\Delta$,

$$\begin{aligned}
&\mathbb{P}^N\left(\sup_{f \in Lip_{\mathcal{K}}(1)} \left|tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]\right| \geq \delta\right) \\
&\leq \mathbb{P}^N\Bigg(\sup_{f \in Lip_{\mathcal{K}}(1)} \Big\{\left|tr_N(f_\Delta(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\Delta(\mathbf{X_A})))\right| + \big|tr_N(f(\mathbf{X_A})) \\
&\qquad - tr_N(f_\Delta(\mathbf{X_A}))\big| + \left|\mathbb{E}^N[tr_N(f(\mathbf{X_A}))] - \mathbb{E}^N[tr_N(f_\Delta(\mathbf{X_A}))]\right|\Big\} \geq \delta\Bigg) \\
&\leq \mathbb{P}^N\left(\sup_{f_\Delta}\left|tr_N(f_\Delta(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\Delta(\mathbf{X_A})))\right| > \delta - 2\Delta\right) \\
&\leq \frac{|\mathcal{K}|}{\Delta} \sup_{g_\Delta^{(j)} \in Lip(1)} \mathbb{P}^N\left(\left|tr_N(g_\Delta^{(j)}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_\Delta^{(j)}(\mathbf{X_A}))]\right| \geq \frac{\Delta(\delta - 2\Delta)}{|\mathcal{K}|}\right) \\
&\leq \frac{8|\mathcal{K}|}{\delta} \exp\left\{ - \int_0^{\frac{N\delta^2}{8\sqrt{2}a|\mathcal{K}|}} h^{-1}(s)ds \right\},
\end{aligned} \tag{2.34}$$

whenever $0 < \delta < \sqrt{8\sqrt{2}a|\mathcal{K}|h\left(T^-\right)/N}$, and where the last inequality follows from part (i) of the previous proposition by taking also $\Delta = \delta/4$.

In order to prove part (ii), for any $f \in Lip_b(1)$, i.e, such that $\|f\|_{Lip} \le 1$, $\|f\|_\infty \le b$, and any $\tau > 0$, let f_τ be given via:

$$f_\tau(x) = \begin{cases} f(x) & \text{if } |x| < \tau \\ f(\tau) - \text{sign}(f(\tau))(x-\tau) & \text{if } \tau \le x < \tau + |f(\tau)| \\ f(-\tau) + \text{sign}(f(-\tau))(x+\tau) & \text{if } -\tau - |f(-\tau)| < x \le -\tau \\ 0 & \text{otherwise.} \end{cases} \tag{2.35}$$

Clearly $f_\tau \in Lip(1)$ and $supp(f_\tau) \subset [-\tau - |f(-\tau)|, \tau + |f(\tau)|]$. Moreover,

$$\begin{aligned} &\sup_{f\in Lip_b(1)} \Big| tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A}))) \Big| \\ &\le \sup_{f\in Lip_b(1)} \Big| tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A}))) \Big| \\ &\qquad + \sup_{f\in Lip_b(1)} \Big| tr_N(f(\mathbf{X_A}) - f_\tau(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}) - f_\tau(\mathbf{X_A}))] \Big| \\ &\le \sup_{f\in Lip_b(1)} \Big| tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A}))) \Big| \\ &\qquad + 2tr_N(g_b(|\mathbf{X_A}| - \tau)) + 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))], \end{aligned} \tag{2.36}$$

with g_b given as in (2.33). Now,

$$\begin{aligned} &\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A})))| \ge \delta \Big) \\ &\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\ &\qquad + \mathbb{P}^N\Big(2tr_N(g_b(|\mathbf{X_A}| - \tau)) + 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{2\delta}{3} \Big) \\ &\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\ &+ \mathbb{P}^N\Big(tr_N(g_b(|\mathbf{X_A}| - \tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{\delta}{3} - 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \Big) \\ &\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\ &+ \mathbb{P}^N\Big(tr_N(g_b(|\mathbf{X_A}| - \tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}| \ge \tau\}})] \Big). \end{aligned} \tag{2.37}$$

Let us first bound the second probability in (2.37). Recall that the spectral radius $\rho(\mathbf{X_A}) = \max_{1\le i\le N} |\lambda_i|$ is a Lipschitz function of X with Lipschitz constant at most $a\sqrt{2/N}$. Hence, for any

$0 < t \leq T$, and $\gamma > 0$ such that $\bar{\nu}(p_\gamma) \leq 1/4$,

$$\begin{aligned}
\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{P}^N\Big(|\lambda_i(\mathbf{X_A})| \geq \tau\Big) \\
&\leq \mathbb{P}^N\big(\rho(\mathbf{X_A}) \geq \tau\big) \\
&\leq \mathbb{P}^N\left(\frac{\sqrt{N}}{\sqrt{2}a}\rho(\mathbf{X_A}) - \frac{\sqrt{N}}{\sqrt{2}a}\mathbb{E}^N\big[\rho(\mathbf{X_A})\big] \geq \frac{\sqrt{N}}{\sqrt{2}a}\tau - G_2(\gamma)\right) \\
&\leq \exp\left\{H(t) - \left(\frac{\sqrt{N}}{\sqrt{2}a}\tau - G_2(\gamma)\right)t\right\} \qquad (2.38)
\end{aligned}$$

where, above, we have used Proposition 1.1 in the next to last inequality and where the last inequality follows from Theorem 1 in [12] (p. 1233) with

$$H(t) = \int_0^t h(s)ds = \int_{\mathbb{R}^{N^2}} \big(e^{t\|u\|} - t\|u\| - 1\big)\nu(du).$$

We want to choose τ, such that $\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \delta/12b$. This can be achieved if

$$\frac{\sqrt{N}}{\sqrt{2}a}\tau - G_2(\gamma) \geq \frac{\ln\frac{12b}{\delta} + H(t)}{t}. \qquad (2.39)$$

Since

$$\frac{d}{dt}\left(\frac{\ln\frac{12b}{\delta} + H(t)}{t}\right) = \frac{th(t) - \ln\frac{12b}{\delta} - H(t)}{t^2},$$

and

$$\frac{d^2}{dt^2}\left(\frac{\ln\frac{12b}{\delta} + H(t)}{t}\right) = \frac{t^3H''(t) - 2t(th(t) - \ln\frac{12b}{\delta} - H(t))}{t^4},$$

it is clear that the right hand side of (2.39) is minimized when $t = t_0$, where t_0 is the solution of

$$th(t) - H(t) - \ln\frac{12b}{\delta} = 0,$$

and the minimum is then $h(t_0)$.
Thus, if

$$\tau = C_0(\delta, b) := \frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma) + h(t_0)\Big), \qquad (2.40)$$

then

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \frac{\delta}{12b},$$

and so,

$$\begin{aligned}
\mathbb{P}^N&\left(tr_N(g_b(|\mathbf{X_A}|-\tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}|-\tau))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\right) \\
&\leq \mathbb{P}^N\left(tr_N(g_b(|\mathbf{X_A}|-\tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}|-\tau))] \geq \frac{\delta}{6}\right) \\
&\leq \exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}} h^{-1}(s)ds\right\}, \qquad (2.41)
\end{aligned}$$

for all $0 < \delta < 6\sqrt{2}ah\left(T^{-}\right)/N$, where Proposition 2.1 is used in the last inequality.

For τ chosen as in (2.40), letting $\mathcal{K} = [-\tau - b, \tau + b]$, it follows that for any $f \in Lip_b(1)$, $f_\tau \in Lip_{\mathcal{K}}(1)$. By part (i), the first term in (2.37) is such that

$$\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \geq \frac{\delta}{3}\Big)\\
&\qquad\qquad \leq \mathbb{P}^N\left(\sup_{f_\tau\in Lip_{\mathcal{K}}(1)}|tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f_\tau(\mathbf{X_A}))]| \geq \frac{\delta}{3}\right)\\
&\qquad\qquad \leq \frac{48(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta^2}{144\sqrt{2}a(C_0(\delta,b)+b)}} h^{-1}(s)ds\right\}, \qquad (2.42)
\end{aligned}$$

for all $0<\delta^2\leq 144\sqrt{2}a\big(C_0(\delta,b)+b\big)h(T^-)/N$.

Hence, returning to (2.37), using (2.41) and (2.42) and for

$$\delta < \min\left\{6\sqrt{2}ah\left(T^-\right)/N, \sqrt{144\sqrt{2}a\big(C_0(\delta,b)+b\big)h(T^-)/N}\right\},$$

we have

$$\begin{aligned}
&\mathbb{P}^N\left(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A})))| \geq \delta\right)\\
&\leq 2\frac{24(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}\frac{\delta}{24(C_0(\delta,b)+b)}} h^{-1}(s)ds\right\}+\exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}} h^{-1}(s)ds\right\}\\
&\leq \left(2+\frac{1}{12}\right)\frac{24(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta^2}{144\sqrt{2}a(C_0(\delta,b)+b)}} h^{-1}(s)ds\right\}, \qquad (2.43)
\end{aligned}$$

since only the case $\delta \leq 2b$ presents some interest (otherwise the probability in the statement of the theorem is zero). Part (ii) is then proved.

□

Proof of Proposition 1.4:

As a function of $x \in \mathbb{R}^{N^2}$, $d_W(\hat{\mu}_A^N, \mu)(x)$ is Lipschitz with Lipschitz constant at most $\sqrt{2}a/N$. Indeed, for $x, y \in \mathbb{R}^{N^2}$,

$$\begin{aligned}
d_W(\hat{\mu}_A^N,\mu)(x) &= \sup_{f\in Lip_b(1)}\left|tr_N\big(f(\mathbf{X_A})(x)\big) - \int_{\mathbb{R}} fd\mu\right|\\
&\leq \sup_{f\in Lip_b(1)}\left|tr_N(f(\mathbf{X_A})(x)) - tr_N(f(\mathbf{X_A})(y))\right|\\
&\qquad\qquad + \sup_{f\in Lip_b(1)}\left|tr_N(f(\mathbf{X_A})(y)) - \int_{\mathbb{R}} fd\mu\right|\\
&\leq \frac{\sqrt{2}a}{N}\|x-y\| + d_W(\hat{\mu}_A^N,\mu)(y). \qquad (2.44)
\end{aligned}$$

Theorem 1.4 then follows from Theorem 1 in [12]. □

Proof of Corollary 1.5:

For Lévy measures with bounded support, $\mathbb{E}^N\big[e^{t\|X\|}\big] < +\infty$, for all $t \geq 0$, and moreover

$$h(t) \leq V^2\bigg(\frac{e^{tR}-1}{R}\bigg).$$

Hence

$$H(t) = \int_0^t h(s)ds \leq \frac{V^2}{R^2}\big(s^{tR} - 1 - tR\big),$$

and

$$\exp\bigg\{-\int_0^x h^{-1}(s)ds\bigg\} \leq \exp\bigg\{\frac{x}{R} - \bigg(\frac{x}{R}+\frac{V^2}{R^2}\bigg)\ln\bigg(1+\frac{Rx}{V^2}\bigg)\bigg\}.$$

Thus, one can take

$$C(\delta, b) = C\bigg(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma) + \frac{V^2}{R}\big(e^{t_0R}-1\big)\Big) + b\bigg),$$

where t_0 is the solution, in t, of

$$\frac{V^2}{R^2}\Big(tRe^{tR} - e^{tR} + 1\Big) = \ln\frac{12b}{\delta}.$$

Applying Theorem 1.2 (ii) yields the result.

□

In order to prove Theorem 1.11, we first need the following lemma, whose proof is essentially as the proof of Theorem 1 in [13].

Lemma 2.2. *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\leq i<j\leq N}$ be an α-stable vector, $0 < \alpha < 2$, with Lévy measure ν given by (1.13). For any $x_0, x_1 > 0$, let $g_{x_0,x_1}(x) = g_{x_1}(x - x_0)$, where $g_{x_1}(x)$ is defined as in (2.33). Then,*

$$\mathbb{P}^N\bigg(\Big|tr_N(g_{x_0,x_1}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(\mathbf{X_A}))]\Big| \geq \delta\bigg) \leq C(\alpha)\frac{a^\alpha\sigma(S^{N^2-1})}{N^\alpha\delta^\alpha},$$

whenever $\delta^{1+\alpha} > \big(2\sqrt{2}a\big)^{1+\alpha}\sigma(S^{N^2-1})x_1/\alpha N^{1+\alpha}$ and where $C(\alpha) = 2^{5\alpha/2}(2e\alpha + 2 - \alpha)/\alpha(2-\alpha)$.

Proof of Theorem 1.11

For part (i), first consider $f \in Lip_{\mathcal{K}}(1)$. Using the same approximation as in Theorem 1.2, any function $f \in Lip_{\mathcal{K}}(1)$ can be approximated by f_Δ, which is the sum of at most $|\mathcal{K}|/\Delta$ functions $g^{(j)}_\Delta \in Lip(1)$, regardless of the function f. Now, and as before, for $\delta > 2\Delta$,

$$
\begin{aligned}
&\mathbb{P}^N\left(\sup_{f\in Lip_{\mathcal{K}}(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\right)\\
&\leq\frac{|\mathcal{K}|}{\Delta}\sup_{\substack{g_\Delta^{(j)}\in Lip_b(1)\\ j=1,\cdots,\lceil\frac{|\mathcal{K}|}{\Delta}\rceil}}\mathbb{P}^N\left(\left|tr_N(g_\Delta^{(j)}(\mathbf{X_A}))-\mathbb{E}^N[tr_N(g_\Delta^{(j)}(\mathbf{X_A}))]\right|\geq\frac{\Delta(\delta-2\Delta)}{|\mathcal{K}|}\right)\\
&\leq\frac{4|\mathcal{K}|}{\delta}\frac{8^\alpha a^\alpha C_2(\alpha)\sigma(S^{N^2-1})|\mathcal{K}|^\alpha}{N^\alpha\delta^{2\alpha}},
\end{aligned}
\tag{2.45}
$$

whenever

$$
\frac{\delta^2}{8|\mathcal{K}|}>\frac{2\sqrt{2}a}{N}\Big(\frac{\sigma(S^{N^2-1})\delta}{4\alpha}\Big)^{\frac{1}{1+\alpha}},
\tag{2.46}
$$

and where the last inequality follows from Lemma 2.2, taking also $\Delta=\delta/4$.

For any $f\in Lip_b(1)$, and any $\tau>0$, let f_τ be given as in (2.35). Then, $f_\tau\in Lip_{\mathcal{K}}(1)$, where $\mathcal{K}=[-\tau-b,\tau+b]$, and moreover,

$$
\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\Big)\\
&\leq\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|))-\mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))]\geq\frac{\delta}{3}-2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big)\\
&\qquad+\mathbb{P}^N\Big(\sup_{f_\tau\in Lip_{\mathcal{K}}(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))|\geq\frac{\delta}{3}\Big).
\end{aligned}
\tag{2.47}
$$

The spectral radius $\rho(\mathbf{X_A})$ is a Lipschitz function of X with Lipschitz constant at most $\sqrt{2}a/\sqrt{N}$. Then by Theorem 1 in [13],

$$
\begin{aligned}
\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})]&=\frac{1}{N}\sum_{i=1}^N\mathbb{P}^N\Big(|\lambda_i(\mathbf{X_A})|\geq\tau\Big)\\
&\leq\mathbb{P}^N\Big(\rho(\mathbf{X_A})>\tau\Big)\\
&\leq\mathbb{P}^N\left(\rho(\mathbf{X_A})-m(\rho(\mathbf{X_A}))>\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\right)\\
&\leq\frac{C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}\big(\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\big)^\alpha},
\end{aligned}
\tag{2.48}
$$

whenever

$$
\left(\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\right)^\alpha\geq\frac{2C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}},
\tag{2.49}
$$

and where $C_1(\alpha)=4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$. Now, if τ is chosen such that

$$
\frac{C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}\big(\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\big)^\alpha}\leq\frac{\delta}{12b},
$$

that is, if

$$\Big(\tau - \frac{\sqrt{2}a}{\sqrt{N}} J_1(\alpha)\Big)^{\alpha} \geq \frac{12bC_1(\alpha)2^{\alpha/2}a^{\alpha}\sigma(S^{N^2-1})}{\delta N^{\alpha/2}}, \tag{2.50}$$

it then follows that

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \frac{\delta}{12b}.$$

Since $g_{\tau,b}(|\mathbf{X_A}|)$ is the sum of two functions of the type studied in Lemma 2.2 with $x_1 = b$, we have,

$$\begin{aligned}
\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|)) - \mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big)& \\
\leq 2\mathbb{P}^N\Big(tr_N(g_{\tau,b}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_{\tau,b}(\mathbf{X_A}))] \geq \frac{\delta}{12}\Big)& \\
\leq 2C_2(\alpha)\frac{12^{\alpha}a^{\alpha}\sigma(S^{N^2-1})}{N^{\alpha}\delta^{\alpha}},&
\end{aligned} \tag{2.51}$$

whenever

$$\delta^{1+\alpha} > \Big(\frac{2\sqrt{2}a}{N}\Big)^{1+\alpha}\frac{12^{1+\alpha}\sigma(S^{N^2-1})b}{\alpha}, \tag{2.52}$$

and where $C_2(\alpha) = 2^{5\alpha/2}(2e\alpha + 2 - \alpha)/\alpha(2-\alpha)$. The respective ranges (2.50) and (2.52) suggest that one can choose, for example,

$$\tau = \frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha) + \frac{\sqrt{2}a}{\sqrt{N}}\delta.$$

Then, there exists $\delta(\alpha, a, N, \nu)$ such that for $\delta > \delta(\alpha, a, N, \nu)$,

$$\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]| \geq \delta\Big) \\
&\leq \mathbb{P}^N\Big(\sup_{f_\tau Lip_{\mathcal{K}}(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f_\tau(\mathbf{X_A}))]| \geq \frac{\delta}{3}\Big) \\
&\qquad + \mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|)) - \mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big) \\
&\leq \frac{C_3(\alpha)a^{\alpha}\sigma(S^{N^2-1})\Big(\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha) + b + \frac{\sqrt{2}a}{\sqrt{N}}\delta\Big)^{1+\alpha}}{N^{\alpha}\delta^{1+2\alpha}} + \frac{C_4(\alpha)a^{\alpha}\sigma(S^{N^2-1})}{N^{\alpha}\delta^{\alpha}},
\end{aligned}$$

where $C_3(\alpha) = 2^{4+2\alpha}12^{\alpha}C_2(\alpha)$, $C_4(\alpha) = 2(12^{\alpha})C_2(\alpha)$ and $\delta(\alpha, a, N, \nu)$ is such that (2.46) and (2.52) hold.

Part (ii) is a direct consequence of Theorem 1 of [13], since $d_W(\hat{\mu}_A^N, \mu) \in Lip(\sqrt{2}a/N)$ as shown in the proof of Proposition 1.4. □

Proof of Theorem 1.12. For any $f \in Lip(1)$, Theorem 1 in [13] gives a concentration inequality for $f(X)$, when it deviates from one of its medians. For $1 < \alpha < 2$, a completely similar (even simpler) argument gives the following result,

$$\mathbb{P}^N\big(f(X) - \mathbb{E}^N[f(X)] \geq x\big) \leq \frac{C(\alpha)\sigma(S^{N^2-1})}{x^{\alpha}}, \tag{2.53}$$

whenever $x^\alpha \geq K(\alpha)\sigma(S^{N^2-1})$, where $C(\alpha) = 2^\alpha(e\alpha+2-\alpha)/(\alpha(2-\alpha))$ and $K(\alpha) = \max\{2^\alpha/(\alpha-1), C(\alpha)\}$.

Next, following the proof of Theorem 1.2, approximate any function $f \in Lip_b(1)$ by $f_\tau \in Lip_{[-\tau-b,\tau+b]}(1)$ defined via (2.35). Hence,

$$\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A})))| \geq \delta\Big)\\
&\leq \mathbb{P}^N\Big(\sup_{f_\tau\in Lip_K(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \geq \frac{\delta}{3}\Big)\\
&+\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|)) - \mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})]\Big). \qquad (2.54)
\end{aligned}$$

For $\rho(\mathbf{X_A})$ the spectral radius of the matrix $\mathbf{X_A}$, and for any τ, such that $\tau - \mathbb{E}^N[\rho(\mathbf{X_A})] \geq \Big(\frac{\sqrt{2}a}{\sqrt{N}}K(\alpha)\sigma(S^{N^2-1})\Big)^{1/\alpha}$,

$$\begin{aligned}
\mathbb{E}^N\big(tr_N(\mathbf{1}_{\{|\mathbf{X_A}|>\tau\}})\big) &\leq \mathbb{P}^N\Big(\rho(\mathbf{X_A}) - \mathbb{E}^N[\rho(\mathbf{X_A})] \geq \tau - \mathbb{E}^N[\rho(\mathbf{X_A})]\Big)\\
&\leq \frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{\big(\tau - \mathbb{E}^N[\rho(\mathbf{X_A})]\big)^\alpha}, \qquad (2.55)
\end{aligned}$$

where we have used, in the last inequality, (2.53) and the fact that $\rho(\mathbf{X_A}) \in Lip(\sqrt{2}a/\sqrt{N})$. For $Q > 0$, let $\tau = \mathbb{E}^N[\rho(\mathbf{X_A})] + Q\delta^{-1/\alpha}$. With this choice, we then have:

$$\begin{aligned}
\mathbb{E}^N\big(tr_N(\mathbf{1}_{\{|\mathbf{X_A}|>\tau\}})\big) &\leq \frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{\big(\tau - \mathbb{E}^N[\rho(\mathbf{X_A})]\big)^\alpha}\\
&\leq \delta\frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{Q^\alpha}\\
&\leq \frac{\delta}{12b}, \qquad (2.56)
\end{aligned}$$

provided $Q^\alpha/\delta > \sqrt{2}aK(\alpha)\sigma(S^{N^2-1})/\sqrt{N}$, and $\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})/Q^\alpha \leq 1/(12b)$. Now, taking $Q = \sqrt{2}a\big(12bC(\alpha)\sigma(S^{N^2-1})\big)^{1/\alpha}/\sqrt{N}$, and recalling, for $1 < \alpha < 2$, the lower range concentration result for stable vectors (Theorem 1 and Remark 3 in [5]): For any $\epsilon > 0$, there exists $\eta_0(\epsilon)$, such that for all $0 < \delta < \sqrt{2}a\|f\|_{Lip}\eta_0(\epsilon)/N$,

$$\begin{aligned}
&\mathbb{P}^N\big(tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A}))) \geq \delta\big)\\
&\qquad\leq (1+\epsilon)\exp\left\{-\frac{\frac{2-\alpha}{10}\big(\frac{\alpha-1}{\alpha}\big)^{\frac{\alpha}{\alpha-1}}}{(\sigma(S^{N^2-1}))^{1/(\alpha-1)}}\left(\frac{N}{\sqrt{2}a\|f\|_{Lip}}\right)^{\frac{\alpha}{\alpha-1}}\delta^{\frac{\alpha}{\alpha-1}}\right\}. \qquad (2.57)
\end{aligned}$$

With arguments as in the proof of Theorem 1.2, if

$$\delta < \eta(\epsilon) := \left(\frac{72\sqrt{2}a}{N} \left(\frac{\sqrt{2}a}{\sqrt{N}} J_2(\alpha) + b + \left(\frac{\sqrt{2}a}{\sqrt{N}} K(\alpha)\sigma(S^{N^2-1}) \right)^{1/\alpha} \right) \eta_0(\epsilon) \right)^{1/2},$$

there exist constants $D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$ and $D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$, such that the first term in (2.54) is bounded above by

$$(1+\epsilon)\frac{D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)}{\delta^{\frac{\alpha+1}{\alpha}}} \exp\Big(-D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)\delta^{\frac{2\alpha+1}{\alpha-1}}\Big). \tag{2.58}$$

Indeed, with the choice of τ above and D^* as in (1.25), $2(\tau + b) \le D^*/\delta^{1/\alpha}$. Moreover, as in obtaining (2.34), D_1 can be chosen to be $24D^*$, while D_2 can be chosen to be

$$\frac{\frac{2-\alpha}{10}\left(\frac{\alpha-1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}}{\Big(\sigma(S^{N^2-1})\Big)^{\frac{1}{\alpha-1}}} \left(\frac{N}{\sqrt{2}a}\right)^{\frac{\alpha}{\alpha-1}} \frac{1}{(72D^*)^{\frac{\alpha}{\alpha-1}}}.$$

Next, as already mentioned, $J_2(\alpha)$ can be replaced by $\mathbb{E}^N[\|X\|]$. In fact, according to (1.7) and an estimation in [14], if $\mathbb{E}^N[X] = 0$, then

$$\frac{1}{4(2-\alpha)^{1/\alpha}}\sigma(S^{N^2-1})^{1/\alpha} \le \mathbb{E}^N[\|X\|] \le \frac{17}{8\big((2-\alpha)(\alpha-1)\big)^{1/\alpha}}\sigma(S^{N^2-1})^{1/\alpha}.$$

Finally, note that, as in the proof of Theorem 1.2 (ii), the second term in (2.54) is dominated by the first term. The theorem is then proved, with the constant $D_1\big(a, N, \sigma(S^{N^2-1})\big)$ magnified by 2. □

Proof of Corollary 1.14. As a function of $(\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le K, 1\le j\le N}$, with the choice of A made in (1.27), $\lambda_{max}(\mathbf{X_A}) \in Lip(\sqrt{2})$. Hence part(i) is a direct application of Theorem 1 in [12], while part(ii) can be obtained by applying Theorem 1.7.

□

Acknowledgements Both authors would like to thank the organizers of the Special Program on High-Dimensional Inference and Random Matrices at SAMSI. Their hospitality and support, through the grant DMS-0112069, greatly facilitated the completion of this paper.

References

[1] N. Alon, M. Krivelevich, V. Vu, On the concentration of eigenvalues of random symmetric matrices. *Isr. J. Math.*, **131**, (2002), 259–267. MR1942311

[2] G.W. Anderson, A. Guionnet, O. Zeitouni, Lecture notes on random matrices. *SAMSI.* September, (2006).

[3] Z.D. Bai, Circular law. *Ann. Probab.*, **25**, (1997), 494-529. MR1428519

[4] Z.D. Bai, J.W. Silverstein, Spectral analysis of large dimensional random matrices. Science Press, Beijing, (2006).

[5] J.-C. Breton, C. Houdré, On finite range stable-type concentration. To appear *Theory Probab. Appl.* (2007).

[6] J.-C. Breton, C. Houdré, N. Privault, Dimension free and infinite variance tail estimates on Poisson space. *Acta Appli. Math.*, **95**, (2007), 151–203. MR2317609

[7] R.M. Dudley, Real analysis and probability. Cambridge University Press, (2002). MR1932358

[8] V.L. Girko, Spectral theory of random matrices. Moscow, Nauka, (1988). MR0955497

[9] F. Götze, A. Tikhomirov, On the circular law. Preprint. Avalaible at Math arXiv 0702386, (2007).

[10] A. Guionnet, O. Zeitouni, Concentration of spectral measure for large matrices. *Electron. Comm. Probab.* **5** (2000), 119-136. MR1781846

[11] R.A. Horn, C. Johnson, Topics in matrix analysis. Cambridge Univ. Press, (1991). MR1091716

[12] C. Houdré, Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.* **30** (2002), 1223-1237. MR1920106

[13] C. Houdré, P. Marchal, On the concentration of measure phenomenon for stable and related random vectors. *Ann. Probab.* **32** (2004), 1496-1508. MR2060306

[14] C. Houdré, P. Marchal, Median, concentration and fluctuations for Lévy processes. *Stochastic Processes and their Applications* (2007).

[15] C. Houdré, P. Marchal, P. Reynaud-Bouret, Concentration for norms of infinitely divisible vectors with independent components. *Bernoulli* (2008).

[16] K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices. *Duke math. J.* **91** (1998), 151-204. MR1487983

[17] I.M. Johnstone, On the distribution of the largest principal component. *Ann. Statist.*, **29**, (2001), 295-327. MR1863961

[18] I.M. Johnstone, High dimensional statistical inference and random matrices. International Congress of Mathematicians. Vol. I, 307–333, Eur. Math. Soc., Zürich, (2007) MR2334195

[19] M. Ledoux, The concentration of measure phenomenon. *Math. Surveys Monogr.*, **89** American Mathematical Society, Providence, RI, (2001). MR1849347

[20] M. Ledoux, Deviation inequalities on largest eigenvalues. *Geometric Aspects of Functional Analysis, Israel Seminar 2004-2005*, Lecture Notes in Math. **1910**, 167-219. Springer (2007). MR2349607

[21] P. Marchal, Measure concentration for stable laws with index close to 2. *Electron. Comm. Probab.* **10** (2005), 29-35. MR2119151

[22] M.B. Marcus, J. Rosiński, L^1-norms of infinitely divisible random vectors and certain stochastic integrals. *Electron. Comm. Probab.* **6** (2001), 15-29. MR1817886

[23] V.A. Marčenko, L.A. Pastur. Distributions of eigenvalues of some sets of random matrices. *Math. USSR-Sb.*, **1** (1967), 507-536.

[24] M.L. Mehta, Random matrices, 2nd ed. Academic Press, San Diego, (1991). MR1083764

[25] L.A. Pastur, On the spectrum of random matrices. *Teor. Mat. Fiz.* **10** (1972), 102-112. MR0475502

[26] B. Simon, Trace ideals and their applications. Cambridge University Press, (1979). MR0541149

[27] K-I. Sato, Lévy processes and infinitely divisible distributions. Cambridge University Press, Cambridge, (1999). MR1739520

[28] A. Soshnikov, Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207** (1999), 697-733. MR1727234

[29] A. Soshnikov, A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. *J. Stat. Phys.*, **108, Nos 5/6**, (2002), 1033-1056. MR1933444

[30] A. Soshnikov, Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails. *Elec. Commun. Probab.* **9** (2004), 82-91. MR2081462

[31] A. Soshnikov, Poisson statistics for the largest eigenvalues in random matrix ensembles. *Mathematical physics of quantum mechanics*, 351–364, Lecture Notes in Phys., **690**, Springer, Berlin, 2006. MR2234922

[32] A. Soshnikov, Y. V. Fyodorov, On the largest singular values of random matrices with independent Cauchy entries. *J. Math. Phys.* **46** (2005), 033302. MR2125574

[33] C. Tracy, H. Widom, Level-spacing distribution and the Airy kernel. *Comm. Math. Phys.* **159** (1994), 151-174. MR1257246

[34] C. Tracy, H. Widom, On orthorgonal and symplectic random matrix ensembles. *Comm. Math. Phys.* **177** (1996), 724-754. MR1385083

[35] E.P. Wigner, On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)* **67** (1958), 325-327. MR0095527