# Sharp bounds on the rate of convergence of the empirical covariance matrix ⋆

Radosław ADAMCZAK [a,1], Alexander E. LITVAK [b] Alain PAJOR [c]
Nicole TOMCZAK-JAEGERMANN [b,2]

[a] *Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland*
[b] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1*
[c] *Equipe d'Analyse et Mathématiques Appliquées, Université Paris Est, 5 boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallee, Cedex 2, France*

**Abstract**

Let $X_1, \ldots, X_N \in \mathbb{R}^n$ be independent centered random vectors with log-concave distribution and with the identity as covariance matrix. We show that with overwhelming probability one has

$$\sup_{x \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} \left( |\langle X_i, x\rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2 \right) \right| \leq C \sqrt{\frac{n}{N}},$$

where $C$ is an absolute positive constant. This result is valid in a more general framework when the linear forms $(\langle X_i, x\rangle)_{i\leq N, x\in S^{n-1}}$ and the Euclidean norms $(|X_i|/\sqrt{n})_{i\leq N}$ exhibit uniformly a sub-exponential decay. As a consequence, if $A$ denotes the random matrix with columns $(X_i)$, then with overwhelming probability, the extremal singular values $\lambda_{\min}$ and $\lambda_{\max}$ of $AA^\top$ satisfy the inequalities $1 - C\sqrt{\frac{n}{N}} \leq \frac{\lambda_{\min}}{N} \leq \frac{\lambda_{\max}}{N} \leq 1 + C\sqrt{\frac{n}{N}}$ which is a quantitative version of Bai-Yin theorem [4] known for random matrices with i.i.d. entries.

**Résumé**

**Ordre asymptotique des valeurs singulières extrêmes de la matrice de covariance empirique.** Soient $X_1, \ldots, X_N \in \mathbb{R}^n$ des vecteurs aléatoires indépendants centrés, de matrice de covariance l'identité et à densité log-concave. On démontre qu'avec une grande probabilité, on a

$$\sup_{x \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} \left( |\langle X_i, x\rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2 \right) \right| \leq C \sqrt{\frac{n}{N}},$$

où $C > 0$ est une constante numérique. Ce résultat reste vrai dans le cadre beaucoup plus général où les formes linéaires $(\langle X_i, x\rangle)_{,i\leq N}, x \in S^{n-1}$ et les normes euclidiennes $(|X_i|/\sqrt{n})_{i\leq N}$ vérifient des inégalités de type sous-exponentiel. Il en résulte que si $A$ désigne la matrice dont les colonnes sont $(X_i)$, alors avec grande probabilité, les valeurs singulières extrêmes $\lambda_{\min}$ et $\lambda_{\max}$ de $AA^\top$ vérifient $1 - C\sqrt{\frac{n}{N}} \leq \frac{\lambda_{\min}}{N} \leq \frac{\lambda_{\max}}{N} \leq 1 + C\sqrt{\frac{n}{N}}$, ce qui est une version quantitative du théorème de Bai-Yin [4] bien connu pour les matrices aléatoires à coefficients i.i.d.

*Preprint submitted to the Académie des sciences*

**Version française abrégée**

Soient $X \in \mathbb{R}^n$ un vecteur aléatoire centré de matrice de covariance l'identité et $X_1, \ldots, X_N \in \mathbb{R}^n$ un échantillon de taille $N$ de vecteurs aléatoires indépendants de même loi que $X$. Soit $A$ la matrice aléatoire $n \times N$ dont les colonnes sont les vecteurs $(X_i)$. On note $\lambda_{\min}$ (resp. $\lambda_{\max}$ ) la plus petite (resp. la plus grande) valeur singulière de la matrice de covariance empirique $AA^\top$. Dans l'étude du régime local en théorie des matrices aléatoires, on s'intéresse particulièrement au comportement asymptotique des valeurs extrêmes du spectre de $AA^\top$. Dans le cas de l'Ensemble de Wishart où le vecteur $X$ est à coordonnées i.i.d. le théorème de Bai-Yin assure la convergence de $\lambda_{\min}/N$ et $\lambda_{\max}/N$ lorsque $n, N \to \infty$ et $n/N \to \beta \in (0, 1)$, sous des hypothèses de moment d'ordre 4. Dans cette Note, on se place d'un point de vue "non-asymptotique" et dans un cadre où les coordonnées des vecteurs colonnes sont non i.i.d. et l'on cherche à quantifier le comportement des valeurs extrêmes lorsque $n$ et $N$ sont fixés. Dans le cas gaussien par exemple, on montre que les valeurs singulières extrêmes vérifient les inégalités

$$1 - C\sqrt{\frac{n}{N}} \leq \frac{\lambda_{\min}}{N} \leq \frac{\lambda_{\max}}{N} \leq 1 + C\sqrt{\frac{n}{N}} \tag{1}$$

avec grande probabilité, où $C > 0$ est une constante numérique.

On considère l'espace $\mathbb{R}^n$ muni du produit scalaire naturel $\langle \cdot, \cdot \rangle$ ainsi que de la norme euclidienne associée $|\cdot|$. On utilisera cette même notation $|\cdot|$ pour désigner le cardinal d'un ensemble.

Dans cette Note, on démontre que les inégalités (1) sont vérifiées sous les hypothèses suivantes :

i) $X_1, \ldots, X_N$ sont des vecteurs centrés, indépendants, uniformément $\psi_1$, c'est-à dire qu'il existe $\psi > 0$ tel que

$$\sup_{i \leq N} \sup_{y \in S^{n-1}} \| \, |\langle X_i, y \rangle| \, \|_{\psi_1} \leq \psi,$$

où pour une variable aléatoire $Y \in \mathbb{R}$ on a posé $\|Y\|_{\psi_1} = \inf \{C > 0 \, ; \; \mathbb{E} \exp(|Y|/C) \leq 2\}$.

ii) il existe $K > 0$ tel que

$$\mathbb{P}\left( \max_{i \leq N} |X_i|/\sqrt{n} > K \max\{1, (N/n)^{1/4}\} \right) \leq \exp\left(-\sqrt{n}\right).$$

Un échantillon de vecteurs uniformément distribués dans la boule euclidienne de rayon $K\sqrt{n}$ vérifie ces hypothèses. Un exemple particulièrement intéressant de classe qui vérifie ces hypothèses est *l'Ensemble log-concave* (voir [1]) : on suppose que les vecteurs $(X_i)$ sont indépendants, centrés et log-concaves isotropes. Rappelons qu'un vecteur est isotrope quand sa matrice de covariance est l'identité et qu'il est log-concave si sa distribution est à densité log-concave. Les vecteurs $(X_i)$ vérifient i) et ii) avec $\psi$ et $K$ des constantes universelles. La propriété ii) résulte d'un résultat de Paouris [7] (voir [1], Lemma 3.1).

Des vecteurs aléatoires de $\mathbb{R}^n$ indépendants isotropes $(X_i)_{i \leq N}$ qui vérifient une inégalité de Poincaré avec constante $L$, c'est-à dire tels que $\mathrm{Var}(f(X_i)) \leq L^2 \mathbb{E}|\nabla f(X_i)|^2$ pour toute fonction régulière $f$ à support compact, remplissent les hypothèses i) et ii) avec $\psi = CL$ et $K = CL$ où $C$ est une constante numérique.

Dans le cadre qui précède, la matrice de covariance de $X$ étant l'identité, la relation (1) est une conséquence du théorème plus général suivant :

---

⋆. The research was conducted while the authors participated in the Thematic Program on Asymptotic Geometric Analysis at the Fields Institute in Toronto in Fall 2010.

*Email addresses:* R.Adamczak@mimuw.edu.pl (Radosław ADAMCZAK), alexandr@math.ualberta.ca (Alexander E. LITVAK), alain.pajor@univ-mlv.fr (Alain PAJOR), nicole@ellpspace.math.ualberta.ca (Nicole TOMCZAK-JAEGERMANN).

1 . Research partially supported by MNiSW Grant no. N N201 397437 and the Foundation for Polish Science.

2 . This author holds the Canada Research Chair in Geometric Analysis.

**Théorème 1** *Soient $N, n \geq 1$ des entiers et $\psi, K \geq 1$. Soient $X_1, \dots, X_N \in \mathbb{R}^n$ des vecteurs aléatoires indépendants vérifiant i) et ii) avec les paramètres $\psi$ et $K$. Alors avec une probabilité supérieure à $1 - 2\exp(-c\sqrt{n})$, on a*

$$\sup_{x \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} \left( |\langle X_i, x \rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2 \right) \right| \leq C\,(\psi + K)^2 \ \sqrt{\frac{n}{N}},$$

*où $C, c > 0$ sont des constantes numériques.*

Ce résultat améliore l'encadrement obtenu dans [1] où le facteur $\sqrt{n/N}$ du second membre était remplacé par $\sqrt{n/N}\ln(2N/n)$. Il est aussi optimal aux constantes numériques près puisque qu'il donne une estimation similaire au cas gaussien. L'estimation en $\sqrt{n/N}\ln(2N/n)$ de [1] était suffisante pour résoudre le problème de Kannan, Lovász et Simonovits [6] qui concernait le cas $N$ proportionel à $n$.

Outre que le théorème 1 est optimal, c'est une version quantitative du théorème de Bai-Yin valable dans un cadre non i.i.d. (les coordonnées de $X$ ne sont pas a priori indépendantes). La question de certains principes universels de la théorie des matrices aléatoires peut être posé. En particulier, le résultat asymptotique de Bai-Yin pour les valeurs extrêmes du spectre de la matrice de covariance empirique est-il vrai pour l'Ensemble log-concave. La version quantitative démontrée dans cette Note permet à présent de supporter cette conjecture.

La démonstration du théorème 1 s'appuie sur le résultat suivant de [1] (Theorem 3.13), un choix judicieux des paramètres et un argument d'approximation.

**Théorème 2** *Soient $A$ une matrice $n \times N$ et $1 \leq m \leq N$. On considère le paramètre $A_m$ défini plus loin par (4). Sous les mêmes hypothèses que le théorème 1, pour tout $t \geq 1$ on a*

$$\mathbb{P}\left( \exists m \leq N \quad A_m \geq C\psi t \max\{\sqrt{m}\ln\frac{2N}{m}, \sqrt{n}\} + 6\max_{i \leq N}|X_i| \right) \leq \exp\left(-t\sqrt{n}\right),$$

*où $C > 0$ est une constante numérique.*

Let $X \in \mathbb{R}^n$ be a centered random vector whose covariance matrix is the identity and $X_1$, ..., $X_N$ be independent copies of $X$. Let $A$ be a random $n \times N$ matrix whose columns are $(X_i)$. By $\lambda_{\min}$ (resp. $\lambda_{\max}$) we denote the smallest (resp. the largest) singular number of the matrix of empirical covariance $AA^\top$. In the study of the local regime in the random matrix theory of particular interest is the limit behaviour of extremal values of the spectrum of $AA^\top$. In the case of Wishart Ensemble when the coordinates of $X$ are independent, the Bai-Yin theorem [4] establishes the convergence of $\lambda_{\min}/N$ and $\lambda_{\max}/N$ when $n, N \to \infty$ and $n/N \to \beta \in (0,1)$, under the assumption of a finite fourth moment. In this note we study the asymptotic non-limit behaviour (also called "non-asymptotic" in Statistics) i.e., we look for sharp upper and lower bounds for singular values in terms of $n$ and $N$, when $n$ and $N$ are sufficiently large. For example, for Gaussian matrices it is known that singular values satisfy inequalitites (1) with probability close to 1. We obtain the same estimates for large class of random matrices, which in particular do not require that entries of the matrix are independent or that $X_i$'s are identically distributed. Note that the natural question about convergence of singular values in such a case is still open (see [2] for the case of $X_i$ having uniform distribution on a rescaled $\ell_p^n$ ball).

The natural scalar product and Euclidean norm on $\mathbb{R}^n$ are denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$. We also denote by the same notation $|\cdot|$ the cardinality of a set. By $C$, $C_1$, $c$ etc. we will denote absolute positive constants.

Let $X_1, \dots, X_N$ be a sequence of random vectors in $\mathbb{R}^n$ (not necessarily identically distributed). We say that it is uniformly $\psi_1$ if for some $\psi > 0$,

$$\sup_{i \leq N} \sup_{y \in S^{n-1}} \|\, |\langle X_i, y\rangle|\, \|_{\psi_1} \leq \psi, \tag{2}$$

where for a random variable $Y \in \mathbb{R}$, $\|Y\|_{\psi_1} = \inf\{C > 0\,;\ \mathbb{E}\exp(|Y|/C) \le 2\}$. We say that it satisfies the boundedness condition with constant $K$ (for some $K \ge 1$) if

$$\mathbb{P}\left(\max_{i\le N} |X_i|/\sqrt{n} > K \max\{1, (N/n)^{1/4}\}\right) \le \exp\left(-\sqrt{n}\right). \tag{3}$$

The main result of this note is the following theorem.

**Theorem 1** *Let $N, n$ be positive integers and $\psi, K \ge 1$. Let $X_1, \dots, X_N$ be independent random vectors in $\mathbb{R}^n$ satisfying (2) and (3). Then with probability at least $1 - 2\exp(-c\sqrt{n})$ one has*

$$\sup_{x\in S^{n-1}} \Big| \frac{1}{N}\sum_{i=1}^{N} \left(|\langle X_i, x\rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2\right) \Big| \le C\,(\psi + K)^2\ \sqrt{\frac{n}{N}}.$$

Theorem 1 improves estimates obtained in [1] for log-concave isotropic vectors. There, we considered essentially the case of $N$ proportional to $n$, which was sufficient to answer the question of Kannan, Lovász and Simonovits [6], however, for bigger $N$, the results were off by a logarithmic factor. The theorem above removes this factor completely leading to the best possible estimate for an arbitrary $N$, that is to an estimate of the same order as in the Gaussian case.

As a consequence, we obtain in our setting, the following quantitative version of Bai-Yin theorem [4] known for random matrices with i.i.d. entries.

**Corollary 1** *Let $A$ be a random $n \times N$ matrix, whose columns $X_1, \dots, X_N$ are isotropic random vectors satisfying the assumptions of Theorem 1. Then with probability at least $1 - 2\exp(-c\sqrt{n})$,*

$$1 - C\,(\psi+K)^2\sqrt{\frac{n}{N}} \le \frac{\lambda_{\min}}{N} \le \frac{\lambda_{\max}}{N} \le 1 + C\,(\psi+K)^2\sqrt{\frac{n}{N}}.$$

To emphasize the strength of the above results we observe that conditions (2) and (3) are valid for many classes of distributions.

*Example 1* Random vectors uniformly distributed on the Euclidean ball of radius $K\sqrt{n}$ clearly satisfy (3). They also satisfy (2) with $\psi = CK$.

*Example 2* Log-concave isotropic random vectors in $\mathbb{R}^n$. Recall that a random vector is isotropic if its covariance matrix is the identity and it is log-concave if its distribution has a log-concave density. Such vectors satisfy (2) and (3) for appropriate absolute constants $\psi$ and $K$. The boundedness condition follows from Paouris' theorem ([7]) and is explicitly written e.g., in [1], Lemma 3.1. We would like to remark that a version of Theorem 1 with a weaker probability estimate was proved by Aubrun in the case of isotropic log-concave random vectors under an additional assumption of unconditionality (see [3]).

*Example 3* Any isotropic random vectors $(X_i)_{i\le N}$ in $\mathbb{R}^n$, satisfying the Poincaré inequality with constant $L$, i.e. such that $\mathrm{Var}(f(X_i)) \le L^2\mathbb{E}|\nabla f(X_i)|^2$ for all compactly supported smooth functions, satisfy (2) with $\psi = CL$ and (3) with $K = CL$. The question from [5] whether all log-concave isotropic random vectors satisfy the Poincaré inequality with an absolute constant is one of the major open problems in the theory of log-concave measures.

The proof of Theorem 1 is close to arguments in Section 4.3 of [1], however it uses a choice of parameters more appropriate for the case considered here, and a new approximation argument. We need additional notations. Let $1 \le m \le N$. By $U_m$ we denote the subset of all vectors in $S^{N-1}$ having at most $m$ non-zero coordinates. For an $n \times N$ matrix $A$ we let

$$A_m = \sup_{z\in U_m} |Az|. \tag{4}$$

The main technical tool is the following result which is the "in particular" part of Theorem 3.13 from [1] in which one needs to adjust corresponding constants and to take a union bound.

**Theorem 2** *Let $X_1, \ldots, X_N$ be as in Theorem 1, let $A$ be a random $n \times N$ matrix whose columns are the $X_i$'s. Then for every $t \geq 1$ one has*

$$\mathbb{P}\left(\exists m \quad A_m \geq C\psi t \max\{\sqrt{m}\ln\frac{2N}{m}, \sqrt{n}\} + 6\max_{i\leq N}|X_i|\right) \leq \exp\left(-t\sqrt{n}\right).$$

**Proof of Theorem 1.** We assume $N \geq n$, otherwise Theorem 2 implies Theorem 1. For $x \in S^{n-1}$ set

$$S(x) = \left|\frac{1}{N}\sum_{i=1}^{N}\left(|\langle X_i, x\rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2\right)\right|.$$

Let $B > 0$ be a parameter which we specify later and observe that

$$\sup_{x\in S^{n-1}} S(x) \leq \sup_{x\in S^{n-1}}\Big(\Big|\frac{1}{N}\sum_{i=1}^{N}\Big((|\langle X_i, x\rangle| \wedge B)^2 - \mathbb{E}\left(|\langle X_i, x\rangle| \wedge B\right)^2\Big)\Big|$$

$$+\frac{1}{N}\sum_{i=1}^{N}\left(|\langle X_i, x\rangle|^2 - B^2\right)\mathbf{1}_{\{|\langle X_i, x\rangle|\geq B\}} + \frac{1}{N}\mathbb{E}\sum_{i=1}^{N}\left(|\langle X_i, x\rangle|^2 - B^2\right)\mathbf{1}_{\{|\langle X_i, x\rangle|\geq B\}}\Big).$$

We denote the summands under the supremum by $S_1(x)$, $S_2(x)$, and $S_3(x)$, respectively.

**Estimate for $S_1$:** Given $x \in S^{n-1}$ and $i \leq N$ let $Z_i = Z_i(x) = (|\langle X_i, x\rangle| \wedge B)^2 - \mathbb{E}\left(|\langle X_i, x\rangle| \wedge B\right)^2$. Then $|Z_i| \leq B^2$. Moreover, since

$$\mathrm{Var}(Z_i) \leq \mathbb{E}\left(|\langle X_i, x\rangle| \wedge B\right)^4 \leq \mathbb{E}|\langle X_i, x\rangle|^4 \leq C_1\psi^4,$$

we observe that $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\mathrm{Var}(Z_i) \leq C_1\psi^4$. Thus, by Bernstein's inequality

$$\mathbb{P}\left(S_1(x) \geq \theta\right) = \mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N} Z_i \geq \theta\right) \leq \exp\left(-\frac{\theta^2 N}{2(C_1\psi^4 + B^2\theta/3)}\right).$$

It is well known that $S^{n-1}$ admits a $(1/3)$-net $\mathcal{N}$ in the Euclidean metric such that $|\mathcal{N}| \leq 7^n$. Then by the union bound we obtain that if

$$\theta^2 N > 8C_1\psi^4 n\ln 7 \quad \text{and} \quad \theta N > (8/3)B^2 n\ln 7 \tag{5}$$

then

$$\mathbb{P}\left(\sup_{x\in\mathcal{N}} S_1(x) \geq \theta\right) \leq \exp\left(n\ln 7 - \frac{\theta^2 N}{2(C_1\psi^4 + B^2\theta/3)}\right) \leq \exp\left(-\frac{\theta^2 N}{4(C_1\psi^4 + B^2\theta/3)}\right). \tag{6}$$

**Estimates for $S_2$ and $S_3$:** By Hölder's inequality and (2) we have, for some absolute constant $C_2 \geq 1$,

$$\sup_{x\in S^{n-1}} S_3(x) \leq \frac{1}{N}\sum_{i=1}^{N}\sup_{x\in S^{n-1}}\|\langle X_i, x\rangle\|_4^2\,\mathbb{P}\left(|\langle X_i, x\rangle| \geq B\right)^{1/2} \leq C_2\psi^2\exp(-B/\psi). \tag{7}$$

To estimate $S_2$, we will use the following notation

$$M = \max\left\{\psi^2 n, \max_{i\leq N}|X_i|^2\right\}, \quad E_B = E_B(x) = \{i \leq N\colon |\langle X_i, x\rangle| \geq B\}, \quad m = \sup_{x\in S^{n-1}}|E_B(x)|.$$

By the definition of $A_m$, we have for every $x \in S^{n-1}$

$$B^2|E_B| \leq \sum_{i\in E_B}|\langle X_i, x\rangle|^2 \leq \sup_{|E|\leq m}\sum_{i\in E}|\langle X_i, x\rangle|^2 \leq A_m^2,$$

which yields $B^2 m \le A_m^2$ and $NS_2(x) \le A_m^2$. Theorem 2 implies that for some absolute constant $C \ge C_2$, with probability at least $1-\exp(-\sqrt{n})$ one has

$$B^2 m \le C\left(M + \psi^2 m \ln^2 \frac{2N}{m}\right) \quad \text{and} \quad \sup_{x\in S^{n-1}} S_2(x) \le C\left(\frac{M}{N} + \psi^2 \frac{m}{N}\ln^2\frac{2N}{m}\right). \tag{8}$$

Now we choose the parameters. Let $B = 2\sqrt{2C}\psi\ln(5N/n)$. Then (7) gives $S_3(x) \le C\psi^2\frac{n}{N} \le C\frac{M}{N}$ for all $x \in S^{n-1}$ and together with (8) it yields that with probability at least $1-\exp(-\sqrt{n})$ one has

$$\sup_{x\in S^{n-1}} (S_2(x)+S_3(x)) \le C\left((2M/N) + \psi^2(m/N)\ln^2(2N/m)\right).$$

It is easy to check that $M \ge \psi^2 m \ln^2(2N/m)$ on the set where (8) holds. Indeed, assume it is not so, thus $M < \psi^2 m \ln^2(2N/m)$. Then by (8) we observe that $B^2 \le 2C\psi^2\ln^2(2N/m)$, which implies

$$m \le 2N\exp(-B/\psi\sqrt{2C}) = 2n^2/25N.$$

By our hypothetical upper bound for $M$ and since $f(m) = m\ln^2(2N/m)$ increases on $[1, 2N/e^2]$, we get

$$\psi^2 n \le M \le \psi^2(8n^2/25N)\ln^2(5N/n),$$

which is impossible.

It follows that $\mathbb{P}\left(\sup_{x\in S^{n-1}}(S_2(x)+S_3(x)) \le 3C(M/N)\right) \ge 1-\exp(-\sqrt{n})$. Combining this estimate with (6), we get $\mathbb{P}\left(\sup_{x\in\mathcal{N}} S(x) \le \theta + 3C\frac{M}{N}\right) \ge 1-\exp(-\sqrt{n}) - \exp\left(-\frac{\theta^2 N}{4(C_1\psi^4 + B^2\theta/3)}\right)$. We now set $\theta = C_3\psi^2\sqrt{n/N}$, where $C_3$ is a sufficiently large absolute positive constant so that (5) is satisfied. Then using boundedness condition with constant $K$ we obtain

$$\mathbb{P}\left(\sup_{x\in\mathcal{N}} S(x) \le \left(C_3\psi^2 + 3CK^2\right)\sqrt{n/N}\right) \ge 1-\exp(-\sqrt{n}) - \exp(-cn) \ge 1-2\exp(-c\sqrt{n}),$$

where $c$ is a sufficiently small positive constant. It proves the desired estimate on the (1/3)-net.

To pass from $\mathcal{N}$ to the whole sphere note that $S(x)$ can be written as $|\langle Tx, x\rangle|$, where $T$ is a self-adjoint operator on $\mathbb{R}^n$. Thus, writing for each $x \in S^{n-1}$, $x = y+z$ with $y \in \mathcal{N}$ and $|z| \le 1/3$, we get

$$\|T\| = \sup_{x\in S^{n-1}} |\langle Tx, x\rangle| \le \sup_{y\in\mathcal{N}} |\langle Ty, y\rangle| + \frac{2}{3}\sup_{y\in\mathcal{N}} |Ty| + \sup_{|z|\le 1/3} |\langle Tz, z\rangle| \le \sup_{y\in\mathcal{N}} S(y) + \frac{7}{9}\|T\|,$$

which implies the desired estimate on the whole sphere $S^{n-1}$. $\square$

## References


[1] R. Adamczak, A. E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Quantitative estimates of the convergence of the empirical covariance matrix in log-concave Ensembles, *Journal of AMS*, 234 (2010), 535–561.

[2] G. Aubrun, Random points in the unit ball of $\ell_p^n$, *Positivity*, 10 (2006), 755–759.

[3] G. Aubrun, Sampling convex bodies: a random matrix approach, *Proc. Amer. Math. Soc.*, 135 (2007), 1293–1303.

[4] Z. D. Bai and Y. Q. Yin, Limit of the smallest eigenvalue of a large dimensional sample covariance matrix, *Ann. Probab.* 21 (1993), 1275–1294.

[5] R. Kannan, L. Lovász and M. Simonovits, Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* 13 (1995), no. 3-4, 541–559.

[6] R. Kannan, L. Lovász and M. Simonovits, Random walks and $O^*(n^5)$ volume algorithm for convex bodies, *Random structures and algorithms*, 2 (1997), no. 1, 1–50.

[7] G. Paouris, Concentration of mass on convex bodies. *Geom. Funct. Anal.* 16 (2006), no. 5, 1021–1049.