Random Euclidean embeddings in finite dimensional Lorentz spaces

Daniel J. Fresen[*]

April 27, 2021

Abstract

Quantitative bounds for random embeddings of $\mathbb{R}^k$ into Lorentz sequence spaces are given, with improved dependence on ε.

1 Introduction

Our starting point is Milman's general Dvoretzky theorem [8]. The dependence on ε is due to Schechtman [14] following Gordon [5], and the Gaussian formulation due to Pisier [13]. We refer the reader to [9, 11, 15] for more details.

Theorem 1 *There exists a universal constant $c > 0$ such that the following is true. Consider any*

$$(n, k, \varepsilon) \in \mathbb{N} \times \mathbb{N} \times (0, 1)$$

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and set

$$M = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \|x\| \exp\left(-\frac{1}{2}|x|^2\right) dx \qquad b = \sup\left\{\|\theta\| : \theta \in S^{n-1}\right\}$$

where $|\cdot|$ denotes the standard Euclidean norm. Assume that $k \leq d$ where

$$d = c\left(\frac{M}{b}\right)^2 \varepsilon^2$$

and let G be an $n \times k$ random matrix with i.i.d. standard normal random variables as entries. With probability at least $1 - 2e^{-d}$, the following event occurs: for all $x \in \mathbb{R}^k$,

$$(1 - \varepsilon)M|x| \leq \|Gx\| \leq (1 + \varepsilon)M|x| \tag{1}$$

---

[*]University of Pretoria, Department of Mathematics and Applied Mathematics, daniel.fresen@up.ac.za
MSC: 46B06, 46B07, 46B09, 52A21, 52A23.

If $(X, \|\cdot\|)$ is a general n-dimensional normed space over $\mathbb{R}$, one can put coordinates on X by identifying it with $\mathbb{R}^n$ in such a way so that

$$b = 1 \qquad\qquad M \geq c\sqrt{\ln n}$$

and therefore if n is sufficiently large we can ensure that d is large. Logarithmic dependence on n is the worst case scenario (for the best possible choice of coordinates), which is the correct behavior in ℓ_∞^n but can be improved significantly for other spaces, such as ℓ_p^n for fixed $p \in [1, \infty)$, where one gets power dependence on n.

By rotational invariance of the normal distribution, $\mathrm{Range}(G)$ is a uniformly distributed random subspace in the Grassmannian $G_{n,k}$ of all k-dimensional linear subspaces of $\mathbb{R}^n$. The bounds in (1) mean that on $\mathrm{Range}(G)$, $\|\cdot\|$ approximates the pushforward norm $\|y\|_\sharp := M \left| G^{-1} y \right|$, where $G^{-1} : \mathrm{Range}(G) \mapsto \mathbb{R}^k$ denotes the inverse of the linear map associated to G. Since the sub-level sets of this pushforward norm are ellipsoids, Milman's general Dvoretzky theorem can be interpreted as follows: Assuming n is sufficiently large and we have chosen an appropriate coordinate system through which to identify a given normed space X with $\mathbb{R}^n$,

- Most k-dimensional subspaces of X are almost isometric to Hilbert spaces
- Most k-dimensional cross-sections of the unit ball $B = \{x : \|x\| \leq 1\}$ are approximately ellipsoidal.

For $\varepsilon = \varepsilon_0$, for any universal constant $\varepsilon_0 \in (0, 1)$, Theorem 1 is in a particular sense sharp (see [6, 10]), and in this sense dependence on n is understood. However the question of optimal dependence on ε is open. We refer the reader to [12, 16] for the best existing bounds of the form $\rho(\varepsilon) \ln n$ in the existential Dvoretzky theorem, where one is satisfied with a single subspace of this dimension, and bounds for general classes of spaces with symmetries are contained in [2, 17].

Paouris, Valettas and Zinn [11] studied dependence on ε in the randomized Dvoretzky theorem for the ℓ_p^n spaces, both in the range $1 \leq p \leq C \ln n$ and $p > C \ln n$, improving on the bounds in Theorem 1.

The results we now present extend results in [11] to the class of Lorentz spaces. These spaces have a structure that is more complicated than that of the ℓ_p^n spaces, and the Lorentz norm of a Gaussian random vector is typically not written in terms of the sum of i.i.d. random variables. Our approach is different to that in [11] and in the special case of ℓ_p^n it allows for a simpler proof without the use of Talagrand's $L_1 - L_2$ inequality, and with improved dependence on p (removing a factor of p^{-p}).

2   What's new?

Consider the finite dimensional Lorentz spaces, i.e. $\mathbb{R}^n$ endowed with the norm

$$|x|_{\omega,p} = \left(\sum_{i=1}^{n} \omega_i x_{[i]}^p \right)^{1/p}$$

where $1 \leq p < \infty$, $(\omega_i)_1^n$ is any non-increasing sequence in $[0, 1]$ with $\omega_1 = 1$, and $(x_{[i]})_1^n$ denotes the non-increasing rearrangement of $(|x_i|)_1^n$. These spaces are the finite dimensional counterparts to the infinite dimensional Lorentz spaces (see [7]) which play a classical role in analysis.

Before studying the general case in Section 9, we study the special case where $\omega_i = i^{-r}$ for $0 \leq r < \infty$, using the notation

$$|x|_{r,p} = \left(\sum_{i=1}^n i^{-r} x_{[i]}^p \right)^{1/p}$$

In this special case, using Lemmas 13 and 14, Theorem 1 applies with

$$d = \begin{cases} c_{r,p} n \varepsilon^2 & : \quad 0 \leq r \leq 1/2, \quad p < 2(1-r) \\ c_{r,p} n (\ln n)^{1-\frac{2}{p}} \varepsilon^2 & : \quad 0 \leq r \leq 1/2, \quad p = 2(1-r) \\ c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^2 & : \quad 0 \leq r \leq 1/2, \quad p > 2(1-r) \\ c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^2 & : \quad 1/2 < r < 1 \\ c_{r,p} (\ln n)^{1+\frac{2}{p}} \varepsilon^2 & : \quad r = 1 \end{cases}$$

where the coefficients $c_{r,p} > 0$ do not depend on anything except r and p and can be written explicitly. Our first main result, Theorem 17, is a little hard on the eye, and we have hidden it near the end of the paper. Its two main corollaries, however, are simpler (Corollaries 2 and 4). Both improve the dependence on ε.

Corollary 2 *In the case $\|\cdot\| = |\cdot|_{r,p}$, where $0 \leq r \leq 1$ and $1 \leq p < \infty$, Theorem 1 holds with the sufficient condition $k \leq d$ replaced with $k \leq d'$, where*

$$d' \geq \begin{cases} c_{r,p} n \varepsilon^2 & : \quad 0 \leq r < 1/2, p < 2 - 2r \\ c_{r,p} \min \left\{ n \varepsilon^2, n (\ln n)^{1-\frac{2}{p}} \varepsilon^{\frac{2}{p}} \right\} & : \quad 0 \leq r < 1/2, p = 2 - 2r \\ c_{r,p} \min \left\{ n \varepsilon^2, n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} \right\} & : \quad 0 \leq r < 1/2, p > 2 - 2r \\ c_{r,p} n (\ln n)^{-1} \varepsilon^2 & : \quad r = 1/2, p = 2 - 2r = 1 \\ c_{r,p} \min \left\{ n (\ln n)^{-p} \varepsilon^2, n^{\frac{1}{p}} \varepsilon^{\frac{2}{p}} \right\} & : \quad r = 1/2, p > 2 - 2r = 1 \\ c_{r,p} \min \left\{ n^{2(1-r)} (\ln n)^{-(p-1)} \varepsilon^2, n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} \right\} & : \quad 1/2 < r < 1 \\ c_{r,p} \min \left\{ (\ln n)^3 \varepsilon^2, (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}} \right\} & : \quad r = 1 \end{cases}$$

For the spaces under consideration, Corollary 2 improves on Theorem 1 in all cases except when either $p < 2 - 2r$ or when $p = 1$, and in those cases it reduces to the old bound.

We have excluded the case $r > 1$ from Corollary 2 because as $n \to \infty$, the resulting Lorentz space is isomorphic to ℓ_∞^n, and Theorem 3 below gives better estimates. More generally, when

$$p > c \ln \left(1 + \frac{1 + n^{1-r}}{1 + |1-r| \ln n} \ln n \right)$$

3

for arbitrarily small universal constant $c > 0$, Lemma 10 (later in the paper) implies

$$|x|_\infty \le |x|_{r,p} \le C\,|x|_\infty \left(1 + \frac{1+n^{1-r}}{1+|1-r|\ln n}\ln n\right)^{1/p} \le C'\,|x|_\infty$$

In this case the estimates in Theorem 17 start to break down. The following result, which generalizes the case $p > c\ln n$ (and $r = 0$) in [11, Theorem 1.2] can then be used instead (with slightly improved probability bound from $1 - Cn^{-c\varepsilon/\ln(1/\varepsilon)}$ in [11]).

Theorem 3 *For all $0 < c_1 < C_1$ there exists $c_2 > 0$ such that the following is true: let $n \in \mathbb{N}$ and let $|\cdot|_\sharp$ be a norm on $\mathbb{R}^n$ that is invariant under coordinate permutations and satisfies $c_1\,|x|_\infty \le |x|_\sharp \le C_1\,|x|_\infty$ (for all $x \in \mathbb{R}^n$). Let $\varepsilon \in (0,1)$ and $0 < k \le c_2\varepsilon\,(\ln\varepsilon^{-1})^{-1}\ln n$. Let G be an $n \times k$ standard Gaussian random matrix. Then with probability at least $1 - Cn^{-c_2\varepsilon}$, for all $x \in \mathbb{R}^k$, $(1-\varepsilon)\,\mathbb{M}\,|Ge_1|_\sharp\,|x| \le |Gx|_\sharp \le (1+\varepsilon)\,\mathbb{M}\,|Ge_1|_\sharp\,|x|.$*

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the map that arranges the coordinates of a vector in non-decreasing order. Let X and Y be independent standard normal random vectors in $\mathbb{R}^n$. From estimates for the normal distribution, see e.g. (5) in Lemma 12, with probability at least $1 - C\exp\left(-ct^2\right)$,

$$\left|\,|X|_\sharp - |Y|_\sharp\right| = \left|\,|TX|_\sharp - |TY|_\sharp\right| \le |TX - TY|_\sharp \le C_1\,|TX - TY|_\infty \le \frac{CC_1 t^2}{\sqrt{\ln n}}$$

This can be converted to a deviation of $|X|_\sharp$ about $\mathbb{M}\,|X|_\sharp$ and is the same deviation estimate satisfied by $|X|_\infty$. The usual proof of the randomized Dvoretzky theorem (using the $(\varepsilon/4)$-net argument, see e.g. [9, 15]) then transfers to $|\cdot|_\sharp$. ∎

The second corollary of Theorem 17 applies to the classical ℓ_p^n spaces, with improved dependence on p compared to [11, Theorem 1.2].

Corollary 4 *For all $C_1 > 0$ there exists $c_2 > 0$ such that the following statement is true. In the case $\|\cdot\| = |\cdot|_{0,p} = |\cdot|_p$, and under the added assumption that $p < C_1\ln n$, Theorem 1 holds with the sufficient condition $k \le d$ replaced with $k \le d'$, where d' is defined as follows,*

$$d' = \begin{cases} cn\varepsilon^2 & : \quad 1 \le p \le 2 \\ c_2\min\left\{c^p n\varepsilon^2, pn^{\frac{2}{p}}\varepsilon^{\frac{2}{p}}\right\} & : \quad 2 < p < C_1\ln n \end{cases}$$

and $c > 0$ is a universal constant.

We end the paper with a result for general Lorentz norms.

Theorem 5 *(refer to Theorem 18 for a more detailed statement) A random embedding of $\left(\mathbb{R}^k, |\cdot|\right)$ into $\left(\mathbb{R}^n, |\cdot|_{\omega,p}\right)$ using a standard Gaussian matrix is, with probability at least $1 - 2e^{-d}$, an ε almost isometry in a sense similar to (1) provided $k \le d$, where d is defined as follows: If $1 < p < \infty$ set*

$$d = \left(1 + \frac{1}{p-1}\right)^{-1}\min\left\{\frac{c^p\left(\sum_{i=1}^n \omega_i\left(\ln\frac{n}{i}\right)^{p/2}\right)^2\varepsilon^2}{\sum_{i=1}^n \omega_i^2\left(\ln\frac{n}{i}\right)^{p-1}}, cB^{-1/p}\left(\sum_{i=1}^n \omega_i\left(\ln\frac{n}{i}\right)^{p/2}\right)^{2/p}\varepsilon^{2/p}\right\}$$

4

where

$$B = \begin{cases} \sum_{i=1}^{n} \omega_i^2 i^{-(p-1)} & : \quad 1 < p < 3/2 \\ \left(\sum_{i=1}^{n} \omega_i^{\frac{2}{2-p}} \right)^{2-p} & : \quad 3/2 \leq p < 2 \\ 1 & : \quad 2 \leq p < \infty \end{cases}$$

and if $p = 1$ *set*

$$d = \frac{c \left(\sum_{i=1}^{n} \omega_i \left(\ln \frac{n}{i} \right)^{1/2} \right)^2 \varepsilon^2}{\sum_{i=1}^{n} \omega_i^2}$$

3   The main engine: Gaussian concentration

We make extensive use of the classical Gaussian concentration inequality. The simple proof of Maurey and Pisier is contained in [13].

Theorem 6 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and let* X *be a random vector in* $\mathbb{R}^n$ *with the standard normal distribution. Then for all* $t \geq 0$,

$$\mathbb{P} \left\{ |X - \mathbb{M}f(X)| > Ct\mathrm{Lip}(f) \right\} \leq 2 \exp \left(-t^2 \right)$$

Assuming $\mathrm{Lip}(f) < \infty$, *the same holds true with* $\mathbb{M}$ *replaced by* $\mathbb{E}$.

If $\|\cdot\| : \mathbb{R}^n \to [0, \infty)$ is a norm, then

$$\mathrm{Lip} \left(\|\cdot\| \right) = \sup \left\{ \|\theta\| : \theta \in S^{n-1} \right\}$$

which can be seen by applying the triangle inequality. Denoting this supremum as $b \left(\|\cdot\| \right)$, Gaussian concentration implies that with probability at least $1 - 2 \exp \left(-t^2 \right)$,

$$\mathbb{M} \|X\| - Ctb \left(\|\cdot\| \right) \leq \|X\| \leq \mathbb{M} \|X\| + Ctb \left(\|\cdot\| \right)$$

For our purposes the right hand inequality $\|X\| \leq \mathbb{M} \|X\| + Ctb \left(\|\cdot\| \right)$ will be sufficient. The following result due to Schechtman[14] provides a uniform bound over a sphere rather than just at a point and reduces to Theorem 6 in the case $k = 1$.

Theorem 7 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and let* G *be an* $n \times k$ *random matrix with i.i.d standard normal random variables as entries. Let* $t \geq 0$ *and assume that* $k \leq ct^2$. *Then with probability at least* $1 - 2 \exp \left(-ct^2 \right)$ *the following event occurs: For all* $\theta \in S^{k-1}$,

$$|f \left(G\theta \right) - \mathbb{M}f \left(G\theta \right)| \leq t Lip(f)$$

4   Key methods

We will deal with functions that may not be Lipschitz, or whose Lipschitz constant is not representative of the typical behaviour of the function. For such functions we will need to prove deviation inequalities for $f(X)$ about $\mathbb{M}f(X)$, where X is a standard normal

random vector as in Section 3. In order to do so, it will be useful to restrict the function f to a set K with the following two properties:

- Lip $(f|_K)$ is nicely bounded
- $\mathbb{P}\{X \notin K\}$ is small (here X is normally distributed as in Theorem 6)

Exactly how one interprets 'nicely bounded' and 'small' may depend on the situation, and the reader will see the details in the proofs of Theorems 17 and 18. It follows from elementary metric space theory that $f|_K$ can be extended to a function $F : \mathbb{R}^n \to \mathbb{R}$ such that $\mathrm{Lip}(F) = \mathrm{Lip}(f|_K)$. One can then apply Gaussian concentration to F and transfer the result back to f since $f(X) = F(X)$ with high probability. This procedure appears in Bobkov, Nayar and Tetali [1] and is further explored in [3].

The original way of applying concentration of measure to prove Dvoretzky's theorem, as in the classical works of Milman and Schechtman, e.g. [8, 9, 14], was to study concentration of $\|X\|$ directly using Lipschitz properties of $\|\cdot\|$ (where $\|\cdot\|$ is the norm of the space in question). One actually considered a random point on the sphere as opposed to a Gaussian random vector, the Gaussian approach being made popular by Pisier [13], but the point is that regardles of the randomness used, the function under consideration was the norm. A trick that will be useful in the context of $\|\cdot\| = |\cdot|_{\omega,p}$ is to study concentration of $|X|_{\omega,p}^p = \sum \omega_i X_{[i]}^p$ instead of $|X|_{\omega,p}$ (using Gaussian concentration and the procedure from [1] just mentioned), and then to convert the result back to a bound on $|X|_{\omega,p}$ by transforming the distribution under the action of $s \mapsto s^{1/p}$. As functions on $\mathbb{R}^n$, $|\cdot|_{\omega,p}$ and $|\cdot|_{\omega,p}^p$ are fundamentally different in terms of their local-Lipschitz properties: the first function achieves its Lipschitz constant on any neighbourhood of the origin, while for the second function, the problematic points where the norm of the gradient is large have been moved far away from the origin so that a convexity argument can be used for points within a certain convex body containing the origin.

So, we will need to bound the distribution of a gradient, which comes down to bounding the expression $\sum i^{-2r} X_{[i]}^{2(p-1)}$. Since the terms of this sum are not independent, one cannot use the classical theory of sums of independent random variables. For $p \geq 3/2$ one can write such a quantity in terms of a norm and use Gaussian concentration applied to norms. For $1 \leq p < 3/2$ one cannot write the gradient as a function of a norm, and one therefore cannot use the equation $\mathrm{Lip}(\|\cdot\|) = b(\|\cdot\|)$. This causes difficulties, but one can bound the gradient above by a function of a norm, which is an interesting problem in its own right, especially in the case $p = 2 - 2r$. We postpone this discussion until Section 7.

5 Notation and once-off explanations

The symbols C and c denote positive universal constants that may take on different values at each appearance. $\mathbb{M}$ and $\mathbb{E}$ denote median and expected value. n and k will typically denote natural numbers and this will not always be stated explicitly but should be clear from the context. $\mathrm{Lip}(f) \in [0, \infty]$ denotes the Lipschitz constant of any function $f : \mathbb{R}^n \to \mathbb{R}$ with respect to the Euclidean norm. $\sum_{i=a}^{b} f(i)$ denotes summation over all $i \in \mathbb{N}$ such that $a \leq i \leq b$, regardless of whether $a, b \in \mathbb{N}$. $1_{\{.\}}$ denotes the indicator function of a set or condition. When proving a probability bound of the form $C \exp(-ct^2)$,

we may take C sufficiently large and assume in the proof that, say, $t \geq 1$, because for $t < 1$ the resulting probability bound is greater than 1 and the result holds trivially. After such a bound is proved we may replace C with 2 using the fact that there exists $c' > 0$ such that
$$\min\left\{1, C\exp\left(-ct^2\right)\right\} \leq 2\exp\left(-c't^2\right)$$
Lastly, by making an all-round change of variables we may take c' to be, say, 1 or $1/2$. The constants in the final probability bound will therefore (without further explanation) not always match what appears to come from the proof.

6   Lemmas

We start with basic estimates for the lower incomplete gamma function suited to our purposes.

Lemma 8 *For all $b, q \in [0, \infty)$,*

$$c^{1+q}\min\left\{1+q, b\right\}^{1+q} \leq \int_0^b e^{-\omega}\omega^q d\omega \leq C^{1+q}\min\left\{1+q, b\right\}^{1+q}$$

$$\frac{ce^b b^{1+q}}{1+q+b} \leq \int_0^b e^{\omega}\omega^q d\omega \leq \frac{Ce^b b^{1+q}}{1+q+b}$$

Proof. The first integrand increases on $[0, q]$ and decreases on $[q, \infty)$, so for $b \leq q$, comparing the integral to the area of a large rectangle,

$$\int_0^b e^{-\omega}\omega^q d\omega \leq be^{-b}b^q \leq b^{1+q}$$

while for $b \geq 1 + q$,

$$\int_0^b e^{-\omega}\omega^q d\omega \leq \Gamma(1+q) \leq C^{1+q}(1+q)^{1+q} \leq C^{1+q}b^{1+q}$$

If $q \geq 1$ this also holds for $q < b < 1 + q$, since in that case $(1+q)^{1+q} \leq C^{1+q}b^{1+q}$. So all that remains for the upper bound is the case where $0 \leq q < 1$ and $q < b < 1 + q$, which implies

$$\int_0^b e^{-\omega}\omega^q d\omega \leq \int_0^b \omega^q d\omega \leq b^{1+q}$$

We now consider the lower bound. For $b \leq q$, comparing the integral to the area of a smaller rectangle,
$$\int_0^b e^{-\omega}\omega^q d\omega \geq \frac{b}{2}e^{-b/2}\left(\frac{b}{2}\right)^q \geq c^{1+q}b^{1+q}$$

For $b \geq q$ and $q \geq 1$, using what we have just proved,

$$\int_0^b e^{-\omega}\omega^q d\omega \geq \int_0^q e^{-\omega}\omega^q d\omega \geq c^{1+q}q^{1+q} \geq c^{1+q}(1+q)^{1+q}$$

7

For $b \geq q$ and $0 \leq q < 1$, we consider two sub-cases: firstly $b \leq 1 + q$, in which case

$$\int_0^b e^{-\omega} \omega^q d\omega \geq c \int_0^b \omega^q d\omega \geq c^{1+q} b^{1+q}$$

and secondly $b > 1 + q$, in which case

$$\int_0^b e^{-\omega} \omega^q d\omega \geq c \int_0^1 \omega^q d\omega \geq c \geq c^{1+q} (1 + q)^{1+q}$$

The second integral with e^ω instead of $e^{-\omega}$ can be estimated by writing $e^\omega \omega^q = \exp(\omega + q \ln \omega)$ and using $\ln \omega \leq (\omega - b)/b + \ln b$, valid for all $\omega \in (0, b]$, and $\ln \omega \geq 2(\omega - b)/b + \ln b$, valid for all $\omega \in [b/2, b]$. Here we also use $ce^z/(1 + z) \leq (e^z - 1)/z \leq Ce^z/(1 + z)$ valid for all $z > 0$. $\blacksquare$

We will use the fact that for any non-increasing function $f : [1, n] \to \mathbb{R}$,

$$\frac{1}{2} \left(f(1) + \int_1^n f(x)dx \right) \leq \sum_{i=1}^n f(i) \leq f(1) + \int_1^n f(x)dx$$

Lemma 9 *For all $a, q \in [0, \infty)$ and all $n \geq 2$ the following is true: If $a \in [0, 1]$ then*

$$\sum_{i=1}^n i^{-a} \left(\ln \frac{n}{i} \right)^q \leq \frac{C^{1+q} n^{1-a} (1 + q)^{1+q} (\ln n)^{1+q}}{((1 - a) \ln n + 1 + q)^{1+q}}$$

and if $a \in [1, \infty)$ the sum is bounded above by

$$\frac{C (\ln n)^{1+q}}{(a - 1) \ln n + 1 + q} + (\ln n)^q$$

The corresponding lower bounds hold by replacing C with c. When $q = 0$ and $i = n$ in the sum, we consider $0^0 = 1$.

Proof. We focus on the upper bounds; the lower bounds follow the same steps. Integrals are estimated using Lemma 8, and we use the fact that $\min\{x, y\}$ is the same order of magnitude as $xy/(x + y)$. First, let $a \in (0, \infty)$. Peeling off the first term, comparing the remaining sum to an integral using monotonicity, and setting $e^{s/a} = n/x$,

$$n^{-a} \sum_{i=1}^n \left(\frac{n}{i} \right)^a \left(\ln \frac{n}{i} \right)^q \leq (\ln n)^q + \frac{n^{1-a}}{a^{1+q}} \int_0^{a \ln n} \exp\left(\left(1 - \frac{1}{a} \right) s \right) s^q ds$$

If $a \in (1, \infty)$ set $w = (1 - 1/a) s$ to get

$$(\ln n)^q + \frac{n^{1-a}}{(a - 1)^{1+q}} \int_0^{(a-1) \ln n} e^w w^q dw$$

If $a = 1$ we get $(\ln n)^q + \int_0^{\ln n} s^q ds$ which can be absorbed into either the case $a \in (0, 1)$ or the case $a \in (1, \infty)$. If $a \in (0, 1)$ then set $w = -(1 - 1/a) s$ to get

$$(\ln n)^q + \frac{n^{1-a}}{(1 - a)^{1+q}} \int_0^{(1-a) \ln n} e^{-w} w^q dw$$

8

If $a = 0$, setting $w = \ln(n/x)$,

$$\sum_{i=1}^{n} i^{-a} \left(\ln \frac{n}{i}\right)^q \le (\ln n)^q + \int_1^n \left(\ln \frac{n}{x}\right)^q dx \le (\ln n)^q + n \int_0^{\ln n} e^{-w} w^q dw$$

For $a \in [0,1]$ the factor $(\ln n)^q$ gets absorbed into the remaining term since

$$\frac{C^{1+q} n^{1-a} (1+q)^{1+q} (\ln n)^{1+q}}{((1-a)\ln n + 1 + q)^{1+q}} = \frac{C^{1+q} n^{1-a} (\ln n)^{1+q}}{\left(1 + \frac{(1-a)\ln n}{1+q}\right)^{1+q}}$$

∎

The following lemma interpolates between the case $a = 1$ and $a \ne 1$.

Lemma 10 *For all $(a, T) \in \mathbb{R} \times [1, \infty)$,*

$$c\frac{1 + T^{1-a}}{1 + |1-a| \ln T} \ln T \le \int_1^T x^{-a} dx \le C\frac{1 + T^{1-a}}{1 + |1-a| \ln T} \ln T$$

Proof. First assume $a \ne 1$ and $T \ne 1$ and write

$$\int_1^T x^{-a} dx = \frac{\exp((1-a)\ln T) - 1}{(1-a)\ln T} \ln T$$

Then interpret $s^{-1}(\exp(s) - 1)$ as the slope of a secant line and bound it above and below by $C(1 + e^s)/(1 + |s|)$ in the cases $s \le -1$, $s \in (-1, 1) \setminus \{0\}$ and $1 \le s$. Then notice that the estimate also holds when $a = 1$ and/or $T = 1$. ∎

Define $\xi_1 : [0, 1] \to [0, 1]$ by $\xi_1(t) = e^t(1 - t)$, from which it follows, see [4], that

$$\xi_1^{-1}(t) \le \min\left\{\sqrt{2(1-t)}, 1 - e^{-1}t\right\} : 0 \le t \le 1$$

The following lemma is taken from [4], which is based on basic estimates for the binomial distribution and the Rényi representation of order statistics from the exponential distribution. Recall that the order statistics of a vector $x \in \mathbb{R}^n$ are denoted $(x_{(i)})_1^n$ (the non-decreasing rearrangement of its coordinates), and the non-increasing rearrangement of the absolute values of the coordinates of x are denoted $(x_{[i]})_1^n$. So if all coordinates of x are non-negative, then $x_{[i]} = x_{(n-i+1)}$.

Lemma 11 *Let $(\gamma_i)_1^n$ be an i.i.d. sample from $(0,1)$ with corresponding order statistics $(\gamma_{(i)})_1^n$ and let $t > 0$. With probability at least $1 - 3^{-1}\pi^2 \exp(-t^2/2)$, the following event occurs: for all $1 \le i \le n$,*

$$\gamma_{(i)} \le 1 - \frac{n - i + 1}{n + 1}\left(1 - \xi_1^{-1}\left(\exp\left(\frac{-t^2 - 4\ln(n - i + 1)}{2(n - i + 1)}\right)\right)\right) \tag{2}$$

and with probability at least $1 - C\exp(-t^2/2)$ the following event occurs: for all $1 \le i \le n$,

$$\gamma_{(i)} \le 1 - \frac{n - i}{n}\exp\left(-c\max\left\{\frac{\left(t + \sqrt{\ln i}\right)\sqrt{i}}{\sqrt{n(n - i + 1)}}, \frac{t^2 + \ln i}{n - i + 1}\right\}\right) \tag{3}$$

9

In Lemma 11, (2) is preferable for $i > n/2$ while (3) is preferable for $i \leq n/2$.

Lemma 12 *Let $n \geq 3$, $t \geq 0$, and let X and Y be independent random vectors in $\mathbb{R}^n$, each with the standard normal distribution. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function that arranges the coordinates of a vector in non-decreasing order. Then with probability at least $1 - C \exp(-t^2)$, the following event occurs: for all $1 \leq i \leq (n+1)/2$,*

$$X_{[i]} \leq C \left(\ln \frac{n}{i} + \frac{t^2}{i} \right)^{1/2} \tag{4}$$

and

$$|TX - TY|_\infty \leq C \min \left\{ \frac{t^2}{\sqrt{\ln n}}, t \right\} \tag{5}$$

and with probability at least 0.51 the following event occurs: for all $1 \leq i \leq (n+1)/2$,

$$c\sqrt{\ln \frac{n}{i}} \leq X_{[i]} \leq C\sqrt{\ln \frac{n}{i}} \tag{6}$$

Proof. We will not repeat 'with probability...' as it is clear that there are probabilities associated to the events in question, and what those probabilities are. To prove estimates for a general distribution based on estimates for the uniform distribution, transform under the action of the inverse cumulative distribution, which is increasing and preserves the operation of arranging in non-increasing order. So we may write $X_{[i]} = \Phi^{-1}\left(\frac{1+\gamma_{(n-i+1)}}{2} \right)$, where $(\gamma_i)_1^n$ is an i.i.d. sample from the uniform distribution on $(0,1)$, and we use the bound

$$c\sqrt{\ln \frac{1}{1-x}} \leq \Phi^{-1}\left(\frac{1+x}{2} \right) \leq C\sqrt{\ln \frac{1}{1-x}} \tag{7}$$

valid for all $x \in (1/3, 1)$, say. (4) now follows from (2) and (7), and includes the upper bound in (6) as a special case. To get a lower bound on an order statistic, we apply Lemma 11 to the i.i.d. random variables $(1 - \gamma_i)_1^n$, which are also uniformly distributed in $(0,1)$ and whose vector of order statistics is $\left(1 - \gamma_{(n-i+1)}\right)_1^n$. An upper bound on $1 - \gamma_{(n-i+1)}$ translates to a lower bound on $\gamma_{(n-i+1)}$. So, from (3) and (7),

$$X_{[i]} \geq c\sqrt{\ln\left(1 - \frac{n-i}{n} \exp\left(-\frac{C\left(1 + \sqrt{\ln i}\right)\sqrt{i}}{n} \right) \right)^{-1}} \geq c\sqrt{\ln \frac{n}{i}}$$

which is seen to hold when $n > n_0$ (using $e^{-z} \geq 1 - z$), and when $n \leq n_0$ this is bounded below by $c_2 \geq c\sqrt{\ln(n/i)}$.

We now consider (5). Its proof, which occupies the next two pages, may later be removed and placed in another paper. It follows from the bounds relating Φ and $\phi = \Phi'$ that

$$\frac{d}{dx}\Phi^{-1}(x) = \frac{1}{\phi(\Phi^{-1}(x))} \leq \frac{C}{\min\{x, 1-x\}\sqrt{\ln \min\{x, 1-x\}^{-1}}} \tag{8}$$

A difference between our current calculations for (5) and what has been done for (4) and (6) above, is that there are no absolute values involved in the definition of $X_{(i)}$, and we write $X_{(i)} = \Phi^{-1}\left(\gamma_{(i)}\right)$, where $(\gamma_i)_1^n$ is an i.i.d. uniform sample. Here we are re-using notation, and this $(\gamma_i)_1^n$ is not the same as the previous $(\gamma_i)_1^n$, which is inconsequential since we are now doing a new calculation. So, by (8) and (2), for all $n/2 < i \leq n$, $\Phi^{-1}\left(\gamma_{(i)}\right) - \Phi^{-1}\left(\frac{i}{n+1}\right)$ is bounded above by

$$C \int_{i/(n+1)}^{1-\frac{n-i+1}{n+1}\left(1-\xi_1^{-1}\exp\left(\frac{-t^2-4\ln(n-i+1)}{2(n-i+1)}\right)\right)} (1-x)^{-1}\left(\ln(1-x)^{-1}\right)^{-1/2} dx$$

$$= C \int_{\ln\frac{n+1}{n-i+1}}^{\ln\frac{n+1}{n-i+1}-\ln\left(1-\xi_1^{-1}\exp\left(\frac{-t^2-4\ln(n-i+1)}{2(n-i+1)}\right)\right)} s^{-1/2} ds$$

$$\leq C \left(\ln\frac{n+1}{n-i+1}\right)^{-1/2} \ln\left(1-\xi_1^{-1}\exp\left(\frac{-t^2-4\ln(n-i+1)}{2(n-i+1)}\right)\right)^{-1}$$

The function $s \mapsto -\ln\left(1-\xi_1^{-1}\exp(-s)\right)$ behaves like $\sqrt{s}$ near 0 and like $s+1$ when s is large. We then consider two cases, depending on whether

$$\frac{t^2+4\ln(n-i+1)}{2(n-i+1)} \tag{9}$$

lies in $(0,1)$ or $[1,\infty)$, and in either case the estimate is bounded above by $Ct^2/\sqrt{\ln n}$. Here we have used the fact that for all $a, b \geq 1$,

$$\ln(1+ab) \leq C \ln(1+a)\ln(1+b)$$

which is true since $1+ab \leq (1+a)(1+b)$ and for positive numbers uniformly bounded away from 0 a product dominates a sum up to a constant. Therefore

$$\ln\left(1+\frac{n}{n-i+1}\right)\ln(1+n-i+1) \geq c\ln n$$

which can be modified be deleting the leftmost $1+$ and changing the factor $\ln(1+n-i+1)$ to something larger such as $n-i+1$ or $(n-i+1)/\ln(n-i+1)$. A zero in the denominator doesn't hurt since we are in practice considering the reciprocals. And assuming as we may that $t \geq 1$, $1+t \leq Ct^2$. So this is where the upper bound $Ct^2/\sqrt{\ln n}$ comes from. Obviously this bound can be improved significantly for individual order statistics; we haven't bothered to write out such bounds since for our purposes we need a uniform estimate over all i. A similar calculation with a lower bound for $\gamma_{(i)}$ follows from (3) applied to $(1-\gamma_{(i)})_1^n$: $\Phi^{-1}\left(\frac{i-1}{n}\right) - \Phi^{-1}\left(\gamma_{(i)}\right)$ is bounded above by

$$C \int_{\frac{i-1}{n}\exp\left(-c\max\left\{\frac{\left(t+\sqrt{\ln(n-i+1)}\right)\sqrt{n-i+1}}{n},\frac{t^2}{n}\right\}\right)}^{(i-1)/n} (1-x)^{-1}\left(\ln(1-x)^{-1}\right)^{-1/2} dx$$

$$= C \int_{-\ln\left[1-\frac{i-1}{n}\exp\left(-c\max\left\{\frac{\left(t+\sqrt{\ln(n-i+1)}\right)\sqrt{n-i+1}}{n},\frac{t^2}{n}\right\}\right)\right]}^{\ln\frac{n}{n-i+1}} s^{-1/2} ds$$

11

Using $\int_a^b s^{-1/2}ds = 2(b-a)/(\sqrt{b}+\sqrt{a}) \leq C(b-a)/\sqrt{b}$ valid for $0 < a < b$, applied to this last quantity,

$$\sqrt{b} = \left(\ln \frac{n}{n-i+1}\right)^{1/2}$$

and $b - a$ is equal to

$$C \quad \ln\left[1 + \frac{i-1}{n-i+1}\left(1 - \exp\left(-c\max\left\{\frac{\left(t + \sqrt{\ln(n-i+1)}\right)\sqrt{n-i+1}}{n}, \frac{t^2}{n}\right\}\right)\right)\right]$$

$$\leq \quad C\frac{i-1}{n-i+1}\left(1 - \exp\left(-c\max\left\{\frac{\left(t + \sqrt{\ln(n-i+1)}\right)\sqrt{n-i+1}}{n}, \frac{t^2}{n}\right\}\right)\right)$$

The expression $1 - e^{-s}$ behaves like s for $0 \leq s \leq 1$ and like 1 when $s > 1$. We now consider two cases depending on whether

$$c\max\left\{\frac{\left(t + \sqrt{\ln(n-i+1)}\right)\sqrt{n-i+1}}{n}, \frac{t^2}{n}\right\}$$

lies in $[0, 1]$ or $(1, \infty)$. In the first case, for the entire expression $C(b-a)/\sqrt{b}$, we get the same bound $Ct^2/\sqrt{\ln n}$ as before, using similar simplifications. In the second case we get

$$\frac{Cn}{(n-i+1)\sqrt{\ln \frac{n}{n-i+1}}}$$

However for the expression defining this case to be > 1,

$$t \geq c\min\left\{\sqrt{n}, \frac{n}{\sqrt{\ln(n-i+1)}}\right\}$$

and regardless of which term defines this minimum we end up with the same bound $Ct^2/\sqrt{\ln n}$. We must now handle the discrepancy between $i/(n+1)$ and $(i-1)/n$ in the computations above. From (8),

$$\left|\Phi^{-1}\left(\frac{i}{n+1}\right) - \Phi^{-1}\left(\frac{i-1}{n}\right)\right| \leq \frac{C}{n}\left(1 - \frac{i}{n+1}\right)^{-1}\left(\ln\left(1 - \frac{i}{n+1}\right)^{-1}\right)^{-1/2} \leq \frac{C}{\sqrt{\ln n}}$$

All of this implies that

$$\left|X_{(i)} - \Phi^{-1}\left(\frac{i}{n+1}\right)\right| \leq \frac{Ct^2}{\sqrt{\ln n}}$$

Similar bounds in the case $1 \leq i \leq n/2$ now follow by symmetry, and they also hold for Y since Y has the same distribution as X, and a bound on $|TX - TY|_\infty$ follows by the triangle inequality. For large values of t this can be improved as follows. Let $x^* \in \mathbb{R}^n$ be defined as

$$x_i^* = \Phi^{-1}\left(\frac{i}{n+1}\right)$$

12

Since T acts as an isometry on each of the $n!$ overlapping regions of $\mathbb{R}^n$ determined by the order of the coordinates of a vector, T is 1-Lipschitz on $\mathbb{R}^n$. So $x \mapsto Tx \mapsto Tx - x^* \mapsto |Tx - x^*|_\infty$ is the composition of 1-Lipschitz functions and by Gaussian concentration,

$$|TX - x^*|_\infty \leq \mathbb{M}\,|TX - x^*|_\infty + Ct$$

which also applies to TY, and the result follows again by the triangle inequality. ∎

Lemma 13 *Let $0 \leq r < \infty$, $1 \leq p < \infty$, $n \geq 2$, and let X be a random vector in $\mathbb{R}^n$ with the standard normal distribution. If $r \in [0,1]$ then*

$$\mathbb{M}\sum_{i=1}^n i^{-r} X_{[i]}^p \leq \frac{C^p p^{p/2} n^{1-r} (\ln n)^{1+p/2}}{[p + (1-r)\ln n]^{1+p/2}}$$

and if $r \in [1,\infty)$ then

$$\mathbb{M}\sum_{i=1}^n i^{-r} X_{[i]}^p \leq \frac{C^p (\ln n)^{1+p/2}}{1 + (r-1)\ln n} + C^p (\ln n)^{p/2}$$

with the reverse inequalities holding with C replaced by c.

Proof. The result follows from Eq. (6) of Lemma 12, together with Lemma 9. ∎

Lemma 14 *Let $0 \leq r < \infty$ and $1 \leq p < \infty$. Then*

$$\sup\left\{\left(\sum_{i=1}^n i^{-r} \theta_{[i]}^p\right)^{1/p} : \theta \in S^{n-1}\right\} = \begin{cases} \left(\sum_1^n i^{-2r/(2-p)}\right)^{(2-p)/2p} & : \quad p \in [1,2) \\ 1 & : \quad p \in [2,\infty) \end{cases}$$

For $p \in [1,2)$ this can be bounded above by

$$1 + C\left(\frac{\ln n}{1 + |2 - 2r - p|\ln n}\right)^{\frac{2-p}{2p}}\left(1 + n^{\frac{2-2r-p}{2p}}\right)$$

and below by the same quantity with C replaced with c. For $r < r_0$ (for any universal constant $r_0 > 1$) the leftmost '1+' can be deleted.

Proof. For $p \in [1,2)$ an upper bound follows by Hölder's inequality for $\ell_{2/(2-p)}^n - \ell_{2/p}^n$ duality, with equality when $\theta_i = i^{-r/(2-p)}\left(\sum_{j=1}^n j^{-2r/(2-p)}\right)^{-1/2}$. Now

$$\left(\sum_1^n i^{-2r/(2-p)}\right)^{(2-p)/2p} \leq \left(1 + \int_1^n x^{-2r/(2-p)} dx\right)^{(2-p)/2p}$$

which is bounded using Lemma 10 and noting that $0 < (2-p)/(2p) \leq 1/2$ and that $c < (2-p)^{(2-p)} < C$. For $p \in [2,\infty)$, $\left(\sum_{i=1}^n i^{-r}\theta_{[i]}^p\right)^{1/p} \leq \left(\sum_{i=1}^n \theta_{[i]}^p\right)^{1/p} \leq 1$ with equality when $\theta = e_1$. ∎

We shall use the fact that for all $b \in [1, n]$, not necessarily an integer,

$$\sum_{i=1}^{n} i^{-r} x_{[i]}^{p} \leq \frac{2n}{b} \sum_{i=1}^{b} i^{-r} x_{[i]}^{p}$$

where the sum on the right is over all $i \in \mathbb{N}$ such that $1 \leq i \leq b$.

Lemma 15 *Let $\psi : \mathbb{R}^n \mapsto \mathbb{R}$ and let $A \subset \mathbb{R}^n$ be convex set with nonempty interior. Let E denote the collection of all $x \in \mathbb{R}^n$ such that the coordinates of x are distinct and non-zero. Assume that ψ is continuous on A and differentiable on $A \cap E$. Then, as an element of $[0, \infty]$,*

$$\mathrm{Lip}\,(\psi|_A) = \sup_{x \in A \cap E} |\nabla \psi(x)|$$

Proof. (Sketch) We focus on proving that $LHS \leq RHS$ when $RHS < \infty$; the rest comes down to approximating a supremum. Note that E is dense in $\mathbb{R}^n$, and by considering sets of the form $\mathrm{conv}\{z, B\}$ where $z \in A$ and B is any ball of non-zero radius contained in A, we see that $\mathrm{int}\,(A)$ is dense in A. So, suppose $|\nabla \psi(z)| \leq L$ for all $z \in A \cap E$ and some $L < \infty$, and consider any $x_1, y_1 \in A$ with $x_1 \neq y_1$, and any $\varepsilon > 0$. We consider sequences x_2, x_3, x_4 and y_2, y_3, y_4 such that each x_{i+1} is sufficiently close to its predecessor x_i and by continuity each $\psi(x_{i+1})$ is sufficiently close to $\psi(x_i)$. Similarly so with the y's. This can be done so that $x_2, y_2 \in \mathrm{int}(A)$, $x_3, y_3 \in \mathrm{int}(A)$ and that all coordinates of x_3 and y_3 are distinct ($2n$ distinct coordinates in total), and that $x_4, y_4 \in A \cap E$ such that x_4 and y_4 also have completely distinct coordinates between the two of them. We bound $|\psi(x_1) - \psi(y_1)|$ up to a term involving ε by integrating over the line segment joining x_4 and y_4, which contains at most $n + n(n-1)/2$ points not in E. Details are left to the reader. ∎

7   Order statistics, norms and quasi-norms

As part of the proof of Theorem 17, towards estimating the distribution of a gradient in order to derive concentration inequalities, we need estimates for the distribution of certain functionals of order statistics. However, in order to avoid the blowup of a quasi-norm constant as $p \to 1$ we need more. We need:

- A deterministic bound of the form

$$\sum_{i=1}^{n} i^{-2r} x_{[i]}^{2(p-1)} \leq \varphi(\|x\|)$$

  where $\|\cdot\|$ is a norm, $\varphi : [0, \infty) \to [0, \infty)$, and the inequality is valid for all $x \in \mathbb{R}^n$,

- a bound on the distribution of $\varphi(\|X\|)$, where X is a random vector in $\mathbb{R}^n$ with the standard normal distribution.

Since these estimates will affect the bounds we end up with in Theorem 17, the problem is not purely existential; we want good estimates. Of greatest interest is the case $p \in$

$(1, 3/2)$ where the quantity $\sum_{i=1}^{n} i^{-2r} x_{[i]}^{2(p-1)}$ is not already a function of a norm of x. Geometrically, the problem is related to finding a convex subset of a given non-convex set whose complement has comparable Gaussian measure to the complement of the given non-convex set.
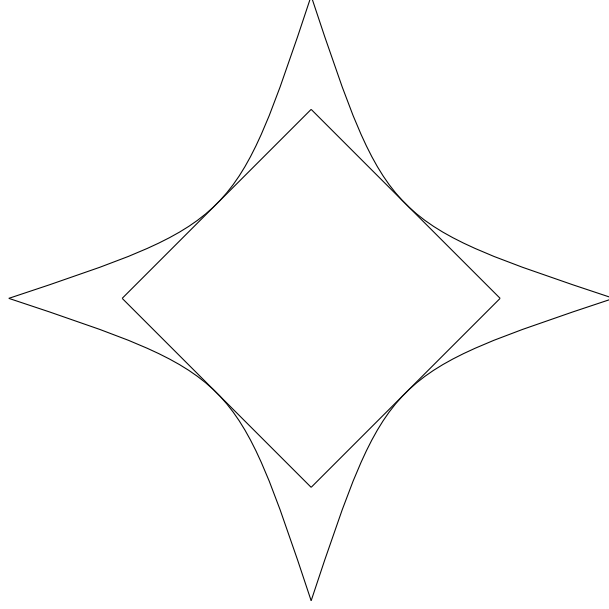

Figure 2

Theorem 16 *Let X be a random vector in $\mathbb{R}^n$, $n \geq 3$, with the standard normal distribution, $0 \leq r < \infty$, $1 \leq p < \infty$ and $t > 0$. In each of the following four cases, definitions are given for R, S are $|\cdot|_\sharp$, and in each case,*

$$\mathbb{P}\left\{|X|_\sharp \leq S\right\} \geq 1 - C \exp\left(-t^2/2\right) \qquad and \qquad |X|_\sharp \leq S \Rightarrow \sum_{i=1}^{n} i^{-2r} X_{[i]}^{2(p-1)} \leq R$$

Case I: *If $p \in [3/2, \infty)$ then for all $x \in \mathbb{R}^n$ set*

$$|x|_\sharp = \left(\sum_{i=1}^{n} i^{-2r} x_{[i]}^{2(p-1)}\right)^{\frac{1}{2(p-1)}}$$

and $R = S^{2(p-1)} = A + Bt^{2(p-1)}$ where

$$A = \frac{C^p p^p n^{1-2r} (\ln n)^p}{[p + (1-2r)\ln n]^p} 1_{\{0 \leq r \leq 1/2\}} + \left(\frac{C^p (\ln n)^p}{1 + (2r-1)\ln n} + C^p (\ln n)^{p-1}\right) 1_{\{1/2 < r < \infty\}}$$

and

$$B = C\left[1 + \left(\frac{\ln n}{1 + |2 - 2r - p|\ln n}\right)^{2-p} \left(1 + n^{2-2r-p}\right)\right] 1_{\{3/2 \leq p < 2\}} + C^p 1_{\{2 \leq p < \infty\}}$$

15

In this case $|\cdot|_\sharp$ is a norm.

Case II: If $p \in [1, 3/2)$ then for all $x \in \mathbb{R}^n$,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \leq C \left(\sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} + \frac{t^2}{i} \right)^{p-1} \right)^{3-2p} |x|_\sharp^{2(p-1)}$$

where

$$|x|_\sharp = \sum_{i=1}^{n/e} \frac{i^{-2r} x_{[i]}}{\left(\ln \frac{n}{i} + \frac{t^2}{i} \right)^{\frac{3-2p}{2}}} \tag{10}$$

and

$$
\begin{aligned}
R &= S = C \sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} + \frac{t^2}{i} \right)^{p-1} \\
&\leq \frac{Cn^{1-2r} (\ln n)^p}{[1 + (1 - 2r) \ln n]^p} 1_{\{0 \leq r \leq 1/2\}} + C \left(\frac{(\ln n)^p}{1 + (2r - 1) \ln n} + (\ln n)^{p-1} \right) 1_{\{1/2 < r < \infty\}} \\
&+ C \left(1 + \frac{1 + n^{2-2r-p}}{1 + |2 - 2r - p| \ln n} \ln n \right) t^{2(p-1)}
\end{aligned}
$$

Under the added assumption that $p \geq 3/2 - 2r$, and because the sums have been restricted to $1 \leq i \leq n/e$, $|\cdot|_\sharp$ is a norm.

Case III: If $p < 3/2 - 2r$ (so necessarily $0 \leq r < 1/4$ and $1 \leq p < 3/2$), then for all $x \in \mathbb{R}^n$,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \leq C n^{(1-2r)(3-2p)} |x|_\sharp^{2(p-1)}$$

where

$$|x|_\sharp = \sum_{i=1}^n i^{-2r} x_{[i]}$$

and

$$
\begin{aligned}
S &= C n^{1-2r} + C n^{\frac{1-4r}{2}} \left(\frac{\ln n}{1 + (1 - 4r) \ln n} \right)^{\frac{1}{2}} t \\
R &= C n^{(1-2r)(3-2p)} S^{2(p-1)} \leq C n^{1-2r} + C n^{2-2r-p} \left(\frac{\ln n}{1 + (1 - 4r) \ln n} \right)^{p-1} t^{2(p-1)}
\end{aligned}
$$

Case IV: When $r \in (1/4, 1/2]$ and $p = 2(1 - r)$, in which case $p \in [1, 3/2)$, the result in Case II can be improved as follows:

<u>Case IVa:</u> If $(1 - 2r) \ln n \geq e$ (which excludes the case $p = 1$), then for all $x \in \mathbb{R}^n$,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \leq C (\ln n)^{3-2p} |x|_\sharp^{2(p-1)}$$

16

where

$$|x|_\sharp = \sum_{i=1}^{n/e} \beta_i^{\frac{-(3-2p)}{2(p-1)}} i^{\frac{-r}{p-1}} x_{[i]},$$

and

$$\beta_i = \frac{(1-2r)^p \ln n}{n^{1-2r}} i^{-2r} \left(\ln \frac{n}{i}\right)^{p-1} + i^{-1}$$

The coefficient $\beta_i^{\frac{-(3-2p)}{2(p-1)}} i^{\frac{-r}{p-1}}$ is non-increasing in i for $1 \le i \le n/e$ (so that $|\cdot|_\sharp$ is a norm). In this case

$$S = C^{\frac{1}{p-1}}(1-2r)^{\frac{-p}{2(p-1)}} (\ln n)^{\frac{-(3-2p)}{2(p-1)}} n^{\frac{1}{2}} + C^{\frac{1}{p-1}} (\ln n)^{\frac{1}{2}} t$$

and

$$R = C (\ln n)^{3-2p} S^{2(p-1)} \le C(1-2r)^{-1} n^{1-2r} + C (\ln n)^{2-p} t^{2(p-1)}$$

<u>*Case IVb:*</u> *If $(1-2r)\ln n < e$, then for all $x \in \mathbb{R}^n$,*

$$\sum_{i=1}^{n} i^{-2r} x_{[i]}^{2(p-1)} \le C (\ln n) |x|_\sharp^{2(p-1)}$$

where $|\cdot|_\sharp = |\cdot|$ is the standard Euclidean norm, $S = Cn^{\frac{1}{2}} + t$, and

$$R = C (\ln n) S^{2(p-1)} \le C (\ln n) t^{2(p-1)}$$

Proof. Case I: $p \in [3/2, \infty)$. In this case $|\cdot|_{2r,2(p-1)}$ is a norm and it follows by classical Gaussian concentration that with probability at least $1 - C \exp\left(-t^2/2\right)$,

$$\sum_{i=1}^{n} i^{-2r} X_{[i]}^{2(p-1)} = |X|_{2r,2(p-1)}^{2(p-1)} \le \left[\mathbb{M} |X|_{2r,2(p-1)} + t\mathrm{Lip}\, |\cdot|_{2r,2(p-1)}\right]^{2(p-1)}$$

Estimates for $\mathbb{M} |X|_{2r,2(p-1)}$ and $\mathrm{Lip}\, |\cdot|_{2r,2(p-1)}$ follow from Lemmas 13 and 14, and we leave the computation to the reader (with the reminder that we are applying these results with $2r$ and $2(p-1)$ instead of r and p). Certain numerical simplifications can be made based on the values of p and r and the existance of the factor C^p.

Proof. Case II: We consider $p \in (1, 3/2)$ and reclaim the case $p = 1$ by taking a limit. For all $x \in \mathbb{R}^n$, by Hölder's inequality, $\sum_1^n i^{-2r} x_{[i]}^{2(p-1)}$ is bounded above by

$$C \sum_{i=1}^{n/e} \frac{i^{-4r(p-1)} x_{[i]}^{2(p-1)}}{(\ln (n/i) + t^2/i)^{(p-1)(3-2p)}} i^{-2r(3-2p)} \left(\ln (n/i) + t^2/i\right)^{(p-1)(3-2p)}$$

$$\le C \left(\sum_{i=1}^{n/e} \frac{i^{-2r} x_{[i]}}{(\ln (n/i) + t^2/i)^{(3-2p)/2}}\right)^{2(p-1)} \left(\sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} + \frac{t^2}{i}\right)^{p-1}\right)^{3-2p}$$

$$= C |x|_\sharp^{2(p-1)} \left(\sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} + \frac{t^2}{i}\right)^{p-1}\right)^{3-2p}$$

17

Eq. (4) from Lemma 12 then implies that

$$|X|_\sharp = \sum_{i=1}^{n/e} \frac{i^{-2r} X_{[i]}}{(\ln(n/i) + t^2/i)^{(3-2p)/2}} \le C' \sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} + \frac{t^2}{i} \right)^{p-1} = \frac{S}{C''}$$

Assuming this event occurs, the above calculation involving Hölder's inequality implies that $\sum_1^n i^{-2r} X_{[i]}^{2(p-1)}$ is bounded above by S. Lastly,

$$S \le C \sum_{i=1}^{n/e} i^{-2r} \left(\ln \frac{n}{i} \right)^{p-1} + C t^{2(p-1)} \sum_{i=1}^{n/e} i^{-2r-p+1}$$

which is bounded above using Lemmas 9 and 10.

Proof. Case III: $p < 3/2 - 2r$ (so necessarily $0 \le r < 1/4$ and $1 \le p < 3/2$). For $p \ne 1$, by Hölder's inequality,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \le \left(\sum_{i=1}^n i^{-2r} x_{[i]} \right)^{2(p-1)} \left(\sum_{i=1}^n i^{-2r} \right)^{3-2p} \le C n^{(1-2r)(3-2p)} |x|_\sharp^{2(p-1)}$$

and for $p = 1$ the same bound is seen to hold. An upper bound on the quantiles of $|X|_\sharp$ follows from Gaussian concentration (making use of Lemmas 10 and 13).

Proof. Case IV: $p = 2(1 - r)$ and $r \in (1/4, 1/2]$, in which case $p \in [1, 3/2)$. The sub-case $(1 - 2r) \ln n < e$ is clear enough by Hölder's inequality for $p \ne 1$,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \le \left(\sum_{i=1}^n (i^{-2r})^{\frac{1}{2-p}} \right)^{2-p} \left(\sum_{i=1}^n \left(x_{[i]}^{2(p-1)} \right)^{\frac{1}{p-1}} \right)^{p-1}$$

and then noting that the exponent $-2r/(2 - p) = -1$ and applying classical Gaussian concentration to $|\cdot|$. For $p = 1$ there is nothing to show. In this sub-case,

$$\frac{(\ln n)^{2-p}}{\ln n} = \exp\left(-(1 - 2r) \ln \ln n\right) \in \left[e^{-1}, 1\right] \qquad n^{1-2r} = \exp\left((1 - 2r) \ln n\right) \in [1, e^e)$$

which is how we simplify the exponents of $\ln n$.

The rest of the proof deals with the other sub-case $(1 - 2r) \ln n \ge e$. Throughout, we make use of the relation $p = 2(1 - r)$ which is not always explicitly re-stated, and the reader should make a mental note of this. The case $p = 1$ is automatically excluded from this sub-case. By taking the constant C in the probability bound to be at least $\sqrt{e}$ we may assume that $t \ge 1$. For any sequence $(\alpha_i)_1^{\lfloor n/e \rfloor}$ with $\alpha_i > 0$, by Hölder's inequality,

$$\sum_{i=1}^n i^{-2r} x_{[i]}^{2(p-1)} \le C \left(\sum_{i=1}^{n/e} \alpha_i^{\frac{1}{3-2p}} \right)^{3-2p} \left(\sum_{i=1}^{n/e} i^{\frac{-r}{p-1}} \alpha_i^{\frac{-1}{2(p-1)}} x_{[i]} \right)^{2(p-1)}$$

$$= C \left(\sum_{i=1}^{n/e} \beta_i \right)^{3-2p} \left(\sum_{i=1}^{n/e} i^{\frac{-r}{p-1}} \beta_i^{\frac{-(3-2p)}{2(p-1)}} x_{[i]} \right)^{2(p-1)}$$

18

where $\beta_i = \alpha_i^{\frac{1}{3-2p}}$. This vector $\beta \in \mathbb{R}^{\lfloor n/e \rfloor}$ is considered a variable for now, and its value will later be fixed to match the value quoted in the statement of the result. Summing only up to n/e will ensure that $i^{\frac{2r}{3-2p}}\beta_i$ is non-decreasing in i, which then implies that

$$|x|_\sharp = \sum_{i=1}^{n/e} i^{\frac{-r}{p-1}} \beta_i^{\frac{-(3-2p)}{2(p-1)}} x_{[i]} \tag{11}$$

is a norm. By classical Gaussian concentration applied to $|\cdot|_\sharp$, with probability at least $1 - C\exp\left(-t^2/2\right)$,

$$\sum_{i=1}^{n} i^{-2r} X_{[i]}^{2(p-1)} \leq \left(\sum_{i=1}^{n/e} \beta_i\right)^{3-2p} \left[\mathbb{M}|X|_\sharp + t\mathrm{Lip}\left(|\cdot|_\sharp\right)\right]^{2(p-1)} \tag{12}$$

The median can be estimated using (4) and the Lipschitz constant computed as the Euclidean norm of the gradient, which gives

$$\mathbb{M}|X|_\sharp \leq C\sum_{i=1}^{n/e} i^{\frac{-r}{p-1}} \beta_i^{\frac{-(3-2p)}{2(p-1)}} \left(\ln\frac{n}{i}\right)^{1/2}$$

$$\mathrm{Lip}\left(|\cdot|_\sharp\right) = \left(\sum_{i=1}^{n/e} i^{\frac{-2r}{p-1}} \beta_i^{\frac{-(3-2p)}{p-1}}\right)^{1/2}$$

We temporarily assume that $\sum \beta_i = 1$, which we may do by homogeneity, although this condition will later be relaxed. We wish to minimize the function

$$\psi(\beta) = \sum_{i=1}^{n/e} i^{\frac{-r}{p-1}} \beta_i^{\frac{-(3-2p)}{2(p-1)}} \left(\ln\frac{n}{i}\right)^{1/2} + t\left(\sum_{i=1}^{n/e} i^{\frac{-2r}{p-1}} \beta_i^{\frac{-(3-2p)}{p-1}}\right)^{1/2}$$

over the collection of all $\beta \in \mathbb{R}^{\lfloor n/e \rfloor}$ such that $i^{\frac{2r}{3-2p}}\beta_i$ is positive and non-decreasing in i and such that $\sum \beta_i = 1$. The method of Lagrange multipliers leads us to solve the equations

$$\frac{\partial \psi(\beta)}{\partial \beta_i} = -\lambda$$

which can be written as

$$B_1 i^{\frac{-r}{p-1}} \left(\ln\frac{n}{i}\right)^{1/2} \beta_i^{\frac{-1}{2(p-1)}} + B_2 i^{\frac{-2r}{p-1}} \beta_i^{\frac{-(2-p)}{p-1}} = 1$$

where B_1 and B_2 are positive values that do not depend on i. This implies that

$$1/2 \leq \max\left\{B_1 i^{\frac{-r}{p-1}} \left(\ln\frac{n}{i}\right)^{1/2} \beta_i^{\frac{-1}{2(p-1)}}, B_2 i^{\frac{-2r}{p-1}} \beta_i^{\frac{-(2-p)}{p-1}}\right\} \leq 1$$

19

and therefore

$$\beta_i \leq \max\left\{ 2^{2(p-1)}B_1^{2(p-1)}i^{-2r}\left(\ln\frac{n}{i}\right)^{p-1}, 2^{\frac{p-1}{2-p}}B_2^{\frac{p-1}{2-p}}i^{\frac{-2r}{2-p}}\right\}$$

with the reverse inequality holding when $2^{2(p-1)}$ and $2^{\frac{p-1}{2-p}}$ are deleted. At this point, and by homogeneity, we remove the condition $\sum \beta_i = 1$ and are led to the definition

$$\beta_i = Ai^{-2r}\left(\ln\frac{n}{i}\right)^{p-1} + i^{-1}$$

for some $A > 0$. B_1, B_2 and the powers of 2 dissapear since they do not depend on i and we have re-scaled β, and we have used the equation $p = 2(1-r)$ to simplify the exponent $-2r/(2-p)$. We now minimize over A. It follows from Lemma 9 that

$$\sum_{i=1}^{n/e} \beta_i \ \leq \ CA(1-2r)^{-p}n^{1-2r} + C\ln n$$

With an eye on (12), it is clear that the bounds for $\mathbb{M}\,|X|_\sharp$ and $\mathrm{Lip}\left(|\cdot|_\sharp\right)$ are decreasing in A. It therefore does not help to let A slip below the point where

$$CA(1-2r)^{-p}n^{1-2r} = C\ln n$$

because as A continues to decrease beyond this point $\sum \beta_i$ stays the same order of magnitude while $\mathbb{M}\,|X|_\sharp + t\mathrm{Lip}\left(|\cdot|_\sharp\right)$ increases. We may therefore assume that

$$A \geq \frac{c(1-2r)^p\ln n}{n^{1-2r}} \qquad\qquad \sum_{i=1}^{n/e}\beta_i \leq CA(1-2r)^{-p}n^{1-2r}$$

If we look back at (12) with our new bound for $\sum \beta_i$ and our definition of β_i, and we take A out of the expression for $\sum \beta_i$ and move it into the powers of β_i with corresponding exponents $\frac{-(3-2p)}{2(p-1)}$ and $\frac{-(3-2p)}{p-1}$ in the expression $\mathbb{M}\,\|X\| + t\mathrm{Lip}\,(\|\cdot\|)$, we see that these powers of β_i become

$$\left(i^{-2r}\left(\ln\frac{n}{i}\right)^{p-1} + A^{-1}i^{-1}\right)^{\frac{-(3-2p)}{2(p-1)}} \qquad\qquad \left(i^{-2r}\left(\ln\frac{n}{i}\right)^{p-1} + A^{-1}i^{-1}\right)^{\frac{-(3-2p)}{p-1}}$$

So, in our current range for A, the expression to be minimized (or at least the bound that we have for it) is increasing. This leads us to take

$$A = \frac{(1-2r)^p\ln n}{n^{1-2r}} \tag{13}$$

Recall that for $|\cdot|_\sharp$ to be a norm, see (11), it is sufficient for $i^{2r/(3-2p)}\beta_i$ to be non-decreasing in i, equivalently for $i^{-r/(p-1)}\beta_i^{-(3-2p)/(2p-2)}$ to be non-increasing. Writing

$$\omega_i = i^{\frac{-r}{p-1}}\beta_i^{\frac{-(3-2p)}{2(p-1)}} = \left[An^{\frac{4r(p-1)}{3-2p}}\left(\frac{n}{i}\left(\ln\frac{n}{i}\right)^{\frac{-(3-2p)}{4r}}\right)^{\frac{-4r(p-1)}{3-2p}} + i^{\frac{p-1}{3-2p}}\right]^{\frac{-(3-2p)}{2(p-1)}}$$

20

and noting that $z \left(\ln z \right)^{-(3-2p)/(4r)}$ is increasing for $z \geq \exp\left((3-2p)/(4r)\right)$, we see that ω_i is decreasing. We now bound $\mathbb{M}\left|X\right|_\sharp$ and $\mathrm{Lip}\left(\left|\cdot\right|_\sharp\right)$. From the definition of β_i,

$$\beta_i^{-1} \leq \min\left\{A^{-1}i^{2r}\left(\ln\frac{n}{i}\right)^{-(p-1)}, i\right\} \tag{14}$$

which leads us to solve,

$$A^{-1}i^{2r}\left(\ln\frac{n}{i}\right)^{-(p-1)} = i$$

Keeping in mind that $1 - 2r = p - 1$, the above equation holds precisely when

$$\frac{n}{i}\left(\ln\frac{n}{i}\right)^{-1} = A^{\frac{1}{p-1}}n \tag{15}$$

The function $z \mapsto z/\ln z$ is increasing on $[e, \infty)$ and we will show that provided $n > n_0$ (for a universal constant $n_0 > 1$),

$$\frac{1}{2}e^2 \leq A^{\frac{1}{p-1}}n \leq \frac{2e}{3\ln n}n^{\frac{3}{2e}} \tag{16}$$

so that (15) has exactly one solution for $i \in [n^{1-\frac{3}{2e}}, n/e^2]$, denoted A_0 (not necessarily an integer) which satisfies

$$n^{1-\frac{3}{2e}} \leq A_0 \leq e^{-2}n \tag{17}$$

$$A_0 \ln\frac{n}{A_0} = A^{\frac{-1}{p-1}} \tag{18}$$

$$\ln\left(A^{1/(p-1)}n\right) = \ln\frac{n}{A_0} - \ln\ln\frac{n}{A_0} \tag{19}$$

$$\ln\left(A^{1/(p-1)}n\right) \leq \ln\left(\frac{n}{A_0}\right) \leq \left(1 - \frac{1}{e}\right)\ln\left(A^{1/(p-1)}n\right) \tag{20}$$

The assumption $n > n_0$ does not limit our generality since the result is directly seen to hold when $n \leq n_0$ in which case many of the coefficients involved are bounded by constants. From (20) and the defining inequality of the current sub-case, it follows that

$$\frac{1}{2}\ln\left((1-2r)\ln n\right) \leq (1-2r)\ln\left(\frac{n}{A_0}\right) \leq \left(1 - \frac{1}{e}\right)\ln\left((1-2r)\ln n\right)$$

For the left inequality we used the fact that $p - 1 \in (0, 1/2)$ and $z\ln z \geq -1/e$ for $z \in (0, 1/2)$. The right inequality is more straightforward. We now verify (16). Recalling (13) and the fact that $1 - 2r = p - 1$, which we are using constantly, the lower bound in (16) holds provided

$$\ln n \geq \frac{e^{(2-\ln 2)(p-1)}}{(p-1)^p}$$

From the definition of the current sub-case, $\ln n \geq e/(p-1)$, so a sufficient condition for the above inequality to hold is

$$2 - \ln 2 \leq \frac{1}{p-1} + \ln(p-1)$$

21

which is true by considering $1/z + \ln z$ for $z \in (0, 1/2)$. The upper bound in (16) holds provided

$$(p-1)\ln n \le \exp\left(\frac{p-1}{p}\ln\left(\frac{2e}{3}n^{\frac{3}{2e}}\right)\right)$$

which holds by applying $e^z \ge ez$. This completes the task of verifying (16). From (14) it follows that

$$\beta_i^{-1} \le \begin{cases} Ci & : \ i \le A_0 \\ CA^{-1}i^{2r}\left(\ln\frac{n}{i}\right)^{-(p-1)} & : \ i > A_0 \end{cases} \tag{21}$$

In an integral where the integrand grows or decays at a controlled rate, one can change an upper bound of $A_0 + 1$ to A_0 at the expense of a constant. Using

$$\int_a^b e^{-\omega}\omega^{1/2}d\omega \le \frac{C(b-a)e^{-a}a^{1/2}}{1+b-a}$$

valid as long as $1 \le a \le b$, and

$$\int_0^b e^{-\omega}\omega^{p-1}d\omega \le C$$

which gives the correct order of magnitude for (say) $b \ge 1/2$, $\mathbb{M}\,|X|_\sharp$ is bounded above by

$$C^{\frac{1}{p-1}}\sum_{i=1}^{A_0}i^{-\frac{1}{2}}\left(\ln\frac{n}{i}\right)^{\frac{1}{2}} + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}\sum_{i=A_0}^{n/e}i^{-2r}\left(\ln\frac{n}{i}\right)^{p-1}$$

$$\le \ C^{\frac{1}{p-1}}n^{-\frac{1}{2}}\int_1^{A_0}\left(\frac{n}{x}\right)^{1/2}\left(\ln\frac{n}{x}\right)^{1/2}dx + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}n^{-2r}\int_{A_0}^{n/e}\left(\frac{n}{x}\right)^{2r}\left(\ln\frac{n}{x}\right)^{p-1}dx$$

$$\le \ C^{\frac{1}{p-1}}n^{\frac{1}{2}}\int_{\frac{1}{2}\ln\frac{n}{A_0}}^{\frac{1}{2}\ln n}e^{-\omega}\omega^{1/2}d\omega + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}(1-2r)^{-p}n^{1-2r}\int_{1-2r}^{(1-2r)\ln\frac{n}{A_0}}e^{-\omega}\omega^{p-1}d\omega$$

$$\le \ C^{\frac{1}{p-1}}A_0^{\frac{1}{2}}\left(\ln\frac{n}{A_0}\right)^{\frac{1}{2}} + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}(1-2r)^{-p}n^{1-2r}$$

$$\le \ C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}(1-2r)^{-p}n^{1-2r}$$

We claim that

$$\int_0^b e^{\omega}\omega^{-(3-2p)}d\omega \le \begin{cases} C(p-1)^{-1}b^{2(p-1)} & : \ 0 \le b \le 1 \\ C(p-1)^{-1} + Ce^b b^{-(3-2p)} & : \ b \ge 1 \end{cases}$$

For $0 \le b \le 1$ this is clear. For $b \ge 3$ this follows because on $[2, \infty)$ the local exponential growth rate of the integrand is

$$\frac{d}{d\omega}[\omega - (3-2p)\ln\omega] = 1 - \frac{3-2p}{\omega} \in [0.5, 1]$$

and for $1 < b < 3$ the bound follows by monotonicity in b. Using the claim just proved, $\text{Lip}\left(|\cdot|_\sharp\right)$ is bounded above by

$$\left[C^{\frac{1}{p-1}}\ln A_0 + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{p-1}}n^{-4r}\int_{A_0}^{n/e}\left(\frac{n}{x}\right)^{4r}\left(\ln\frac{n}{x}\right)^{-(3-2p)}dx\right]^{\frac{1}{2}}$$

$$\leq \ C^{\frac{1}{p-1}}\left(\ln A_0\right)^{\frac{1}{2}} + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}(4r-1)^{1-p}n^{\frac{1-4r}{2}}\left(\int_{4r-1}^{(4r-1)\ln\frac{n}{A_0}}e^\omega\omega^{-(3-2p)}d\omega\right)^{\frac{1}{2}}$$

If $(4r-1)\ln\left(n/A_0\right) < 1$ then this is bounded by

$$C^{\frac{1}{p-1}}\left(\ln n\right)^{\frac{1}{2}} + C^{\frac{1}{p-1}}A^{\frac{-(3-2p)}{2(p-1)}}n^{\frac{1-4r}{2}}\left(\ln\frac{n}{A_0}\right)^{p-1} \leq C^{\frac{1}{p-1}}\left(\ln n\right)^{\frac{1}{2}}$$

To see why the last inequality is true, note that the inequality

$$(\ln n)^{\frac{1}{2}} \geq A^{\frac{-(3-2p)}{2(p-1)}}n^{\frac{1-4r}{2}}\left(\ln\frac{n}{A_0}\right)^{p-1}$$

reduces to

$$A_0^{\frac{3-2p}{2}}\left(\ln\frac{n}{A_0}\right)^{\frac{1}{2}} \leq n^{\frac{4r-1}{2}}\left(\ln n\right)^{\frac{1}{2}}$$

which in turn follows since $1 \leq A_0 \leq n$ and $3 - 2p = 4r - 1$. If $(4r-1)\ln\left(n/A_0\right) \geq 1$ then using $4r - 1 = 3 - 2p$ and $A_0\ln\left(n/A_0\right) = A^{-1/(p-1)}$, we get the same bound, i.e.

$$\text{Lip}\left(|\cdot|_\sharp\right) \leq C^{\frac{1}{p-1}}\left(\ln n\right)^{\frac{1}{2}}$$

Going all the way back to (12), regardless of whether $(4r-1)\ln\left(n/A_0\right)$ lies in $[0,1)$ or $[1,\infty)$,

$$\sum_{i=1}^{n}i^{-2r}X_{[i]}^{2(p-1)} \leq C(1-2r)^{-p}n^{1-2r} + Ct^{2(p-1)}\left(\ln n\right)^{2-p}$$

Then note that

$$\frac{(1-2r)^{-p}}{(1-2r)^{-1}} = \exp\left(-(p-1)\ln(p-1)\right) \in (c,1)$$

■

8   Statement and proof of the main result

The following diagram indicates the various cases considered in Theorem 17; it will be useful to refer back to it when reading the proof.

Case ia: $\frac{3}{2} \le p < \infty$, $0 \le r \le \frac{1}{2}$

Case ib*: $\frac{3}{2} \le p < \infty$, $\frac{1}{2} < r \le 1$

Case ib**: $\frac{3}{2} \le p < \infty$, $1 < r \le 2$

Case iia: $1 \le p < \frac{3}{2}$, $\frac{3-2p}{4} \le r \le \frac{1}{2}$, $p \ne 2 - 2r$

Case iib*: $1 \le p < \frac{3}{2}$, $\frac{1}{2} < r \le 1$

Case iib**: $1 \le p < \frac{3}{2}$, $1 < r \le 2$

Case iii: $1 \le p < \frac{3}{2}$, $p < \frac{3}{2} - 2r$

Case iv: $1 \le p < \frac{3}{2}$, $p = 2 - 2r$

------ : boundary of a region   ------ : other relevant line
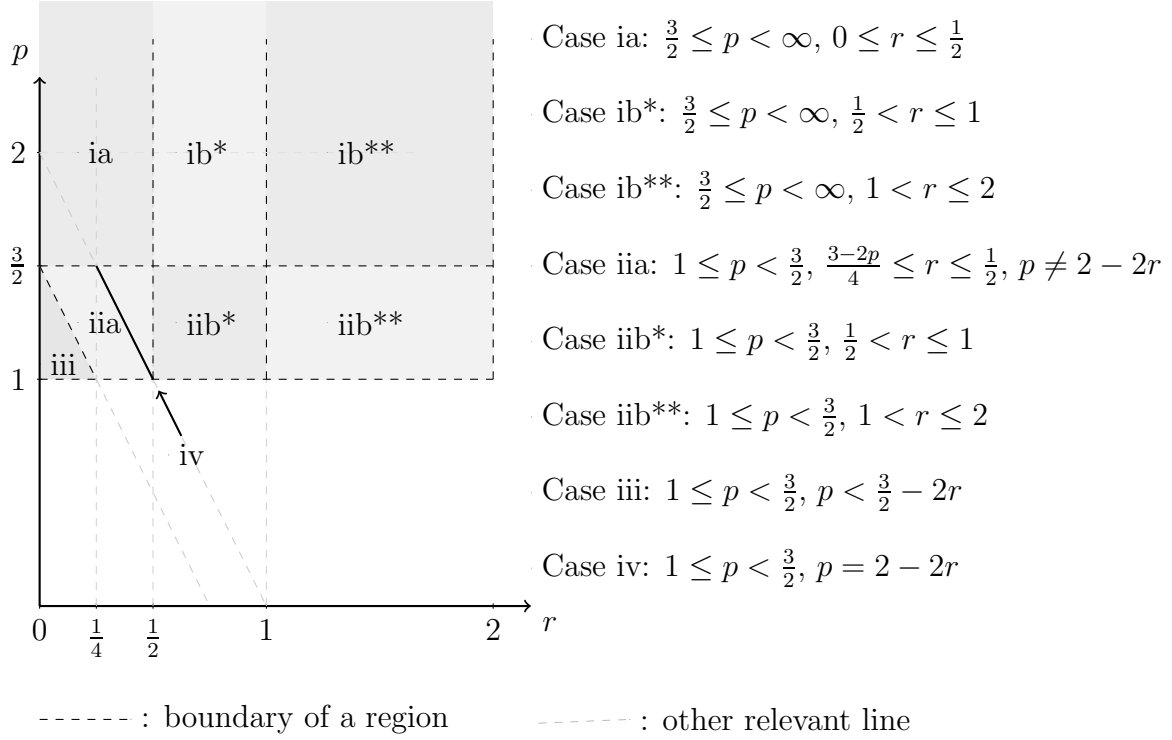
Figure 1

Theorem 17 *There exist universal constants $C, c > 0$ and a function $(r, p) \mapsto c_{r,p}$ from $[0, 2] \times [1, \infty)$ to $(0, \infty)$ such that the following is true. Let*

$$(n, k, \varepsilon, r, p) \in \mathbb{N} \times \mathbb{N} \times (0, 1/2) \times [0, 2] \times [1, \infty)$$

and let G be a random $n \times k$ matrix with i.i.d. standard normal random variables as entries. Cases i-iv will be defined as in Figure 1 above. In each case variables E and F will be defined, and as long as $k \le \min\{E, F\}$, with probability at least $1 - C \exp(-\min\{E, F\})$ the following event occurs: for all $x \in \mathbb{R}^k$,

$$(1 - \varepsilon) M_{r,p} |x| \le |Gx|_{r,p} \le (1 + \varepsilon) M_{r,p} |x|$$

where $M_{r,p}$ denotes the median of $|Ge_1|_{r,p}$. E and F are defined as follows:
In __Case ia__: $\frac{3}{2} \le p < \infty$, $0 \le r \le \frac{1}{2}$ and

$$E = \frac{c^p n (\ln n)^2 [p + (1 - 2r) \ln n]^p \varepsilon^2}{(p + \ln n)^{2+p}} \ge \begin{cases} c_{r,p} n \varepsilon^2 & : \ r \ne 1/2 \\ c_{r,p} n (\ln n)^{-p} \varepsilon^2 & : \ r = 1/2 \end{cases}$$

$$F = \frac{c p n^{\frac{2(1-r)}{p}} (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}}}{\left(1 + n^{\frac{2-2r-p}{p}}\right) (p + \ln n)^{1+\frac{2}{p}}} \left(\frac{1 + |2 - 2r - p| \ln n}{\ln n}\right)^{\max\left\{\frac{2-p}{p}, 0\right\}}$$

$$\ge \begin{cases} c_{r,p} n \varepsilon^{\frac{2}{p}} & : \ p < 2 - 2r \\ c_{r,p} n (\ln n)^{-\left(\frac{2}{p} - 1\right)} \varepsilon^{\frac{2}{p}} & : \ p = 2 - 2r \\ c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} & : \ p > 2 - 2r \end{cases}$$

24

In *Case ib**: $\frac{3}{2} \leq p < \infty$, $\frac{1}{2} < r \leq 1$ and

$$E = \frac{c^p p^p n^{2(1-r)} (\ln n)^2 [1 + (2r-1)\ln n] \varepsilon^2}{[p + (1-r)\ln n]^{2+p}} \geq \begin{cases} c_{r,p} n^{2(1-r)} (\ln n)^{-(p-1)} \varepsilon^2 & : \ r \neq 1 \\ c_{r,p} n^{2(1-r)} (\ln n)^3 \varepsilon^2 & : \ r = 1 \end{cases}$$

$$F = \frac{cpn^{\frac{2(1-r)}{p}} (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}}}{[p + (1-r)\ln n]^{1+\frac{2}{p}}} \geq \begin{cases} c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} & : \ r \neq 1 \\ c_{r,p} n^{\frac{2(1-r)}{p}} (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}} & : \ r = 1 \end{cases}$$

In *Case ib***: $\frac{3}{2} \leq p < \infty$, $1 < r \leq 2$ and

$$E = \frac{c^p (\ln n)^3 \varepsilon^2}{[1 + (r-1)\ln n]^2} \geq \begin{cases} c_{r,p} (\ln n) \varepsilon^2 & : \ r \neq 1 \\ c_{r,p} (\ln n)^3 \varepsilon^2 & : \ r = 1 \end{cases}$$

$$F = \frac{c (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}}}{[1 + (r-1)\ln n]^{\frac{2}{p}}} \geq \begin{cases} c_{r,p} (\ln n) \varepsilon^{\frac{2}{p}} & : \ r \neq 1 \\ c_{r,p} (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}} & : \ r = 1 \end{cases}$$

In *Case iia*: $1 \leq p < \frac{3}{2}$, $\frac{3-2p}{4} \leq r \leq \frac{1}{2}$, $p \neq 2 - 2r$ and

$$E = \frac{cn [1 + (1-2r)\ln n]^p \varepsilon^2}{(\ln n)^p} \geq \begin{cases} c_{r,p} n \varepsilon^2 & : \ r \neq 1/2 \\ c_{r,p} n (\ln n)^{-p} \varepsilon^2 & : \ r = 1/2 \end{cases}$$

$$F = \frac{cn^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}}}{1 + n^{\frac{2-2r-p}{p}}} \left(\frac{1 + |2 - 2r - p| \ln n}{\ln n} \right)^{1/p} \geq \begin{cases} c_{r,p} n \varepsilon^{\frac{2}{p}} & : \ p < 2 - 2r \\ c_{r,p} n (\ln n)^{-\frac{1}{p}} \varepsilon^{\frac{2}{p}} & : \ p = 2 - 2r \\ c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} & : \ p > 2 - 2r \end{cases}$$

In *Case iib**: $1 \leq p < \frac{3}{2}$, $\frac{1}{2} < r \leq 1$ and

$$E = \frac{cn^{2(1-r)} (\ln n)^2 [1 + (2r-1)\ln n] \varepsilon^2}{[1 + (1-r)\ln n]^{2+p}} \geq \begin{cases} c_{r,p} n^{2(1-r)} (\ln n)^{-(p-1)} \varepsilon^2 & : \ r \neq 1 \\ c_{r,p} n^{2(1-r)} (\ln n)^3 \varepsilon^2 & : \ r = 1 \end{cases}$$

$$F = \frac{cn^{\frac{2(1-r)}{p}} (\ln n)^{1+\frac{1}{p}} [1 + |2 - 2r - p| \ln n]^{\frac{1}{p}} \varepsilon^{\frac{2}{p}}}{[1 + (1-r)\ln n]^{1+\frac{2}{p}}} \geq \begin{cases} c_{r,p} n^{\frac{2(1-r)}{p}} \varepsilon^{\frac{2}{p}} & : \ r \neq 1 \\ c_{r,p} n^{\frac{2(1-r)}{p}} (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}} & : \ r = 1 \end{cases}$$

In *Case iib***: $1 \leq p < \frac{3}{2}$, $1 < r \leq 2$ and

$$E = \frac{c (\ln n)^3 \varepsilon^2}{[1 + (r-1)\ln n]^2} \geq c_{r,p} (\ln n) \varepsilon^2$$

$$F = \frac{c (\ln n)^{1+\frac{2}{p}} \varepsilon^{\frac{2}{p}}}{[1 + (r-1)\ln n]^{\frac{2}{p}}} \geq c_{r,p} (\ln n) \varepsilon^{\frac{2}{p}}$$

In *Case iii*: $1 \leq p < \frac{3}{2}$, $p < \frac{3}{2} - 2r$ and

$$E = cn \varepsilon^2$$

$$F = \frac{cn [1 + (1-4r)\ln n]^{1-\frac{1}{p}} \varepsilon^{\frac{2}{p}}}{(\ln n)^{1-\frac{1}{p}}} \geq c_{r,p} n \varepsilon^{\frac{2}{p}}$$

In *Case iv*: $1 \le p < \frac{3}{2}$, $p = 2 - 2r$ *and*

$$E \;=\; \frac{cn\left[1 + (1-2r)\ln n\right]\varepsilon^2}{\ln n} \;\ge\; \begin{cases} c_{r,p}n\varepsilon^2 & : \quad r \ne 1/2 \\ c_{r,p}n\left(\ln n\right)^{-1}\varepsilon^2 & : \quad r = 1/2 \end{cases}$$

$$F \;=\; \frac{cn\varepsilon^{\frac{2}{p}}}{\left(\ln n\right)^{\frac{2-p}{p}}}$$

The parameters E and F in Theorem 17 can be bounded as follows:

$$E \ge \begin{cases} c_{r,p}n\varepsilon^2 & : \quad 0 \le r < 1/2 \\ c_{r,p}n\left(\ln n\right)^{-p}\varepsilon^2 & : \quad r = 1/2 \\ c_{r,p}n^{2(1-r)}\left(\ln n\right)^{-(p-1)}\varepsilon^2 & : \quad 1/2 < r < 1 \\ c_{r,p}\left(\ln n\right)^3\varepsilon^2 & : \quad r = 1 \end{cases}$$

and

$$F \ge \begin{cases} c_{r,p}n\varepsilon^{\frac{2}{p}} & : \quad 0 \le r \le 1/2, p < 2 - 2r \\ c_{r,p}n\left(\ln n\right)^{1-\frac{2}{p}}\varepsilon^{\frac{2}{p}} & : \quad 0 \le r \le 1/2, p = 2 - 2r \\ c_{r,p}n^{\frac{2(1-r)}{p}}\varepsilon^{\frac{2}{p}} & : \quad 0 \le r \le 1/2, p > 2 - 2r \\ c_{r,p}n^{\frac{2(1-r)}{p}}\varepsilon^{\frac{2}{p}} & : \quad 1/2 < r < 1 \\ c_{r,p}\left(\ln n\right)^{1+\frac{2}{p}}\varepsilon^{\frac{2}{p}} & : \quad r = 1 \end{cases}$$

For $n \ge n_0(r,p)$, the coefficient $c_{r,p}$ can be written explicitly in terms of r and p, as can $n_0(r,p)$.

Proof. Let G be a random matrix with i.i.d. standard normal random variables as entries, and let $\theta \in S^{n-1}$. $G\theta$ therefore has the standard normal distribution in $\mathbb{R}^n$. Setting $\psi(x) = \sum i^{-r}x_{[i]}^p$,

$$|\nabla\psi(x)| = p\left(\sum_{i=1}^{n} i^{-2r}x_{[i]}^{2(p-1)}\right)^{1/2} \tag{22}$$

which is valid for all x with distinct non-zero coordinates. Fix any $t > 0$. With (22) in mind, for $j \in \{0,1\}$ set

$$A_j \;=\; \left\{x \in \mathbb{R}^n : |x|_\sharp \le \left(\frac{4}{3}\right)^j S\right\}$$

where $|\cdot|_\sharp$ and S (and R below) are as in Theorem 16. The cases in that theorem overlap, which is not a problem as long as you pick a case that applies to the values of p and r in question, and stick with that case. We shall apply the cases as follows:

- If $3/2 \le p < \infty$ use Case I.
- If $1 \le p < 3/2$ and $p \ge 3/2 - 2r$ and $p \ne 2 - 2r$ use Case II.
- If $1 \le p < 3/2$ and $p < 3/2 - 2r$ use Case III.

26

● If $1 \leq p < 3/2$ and $p = 2 - 2r$ use Case IV (either IVa or IVb, whichever applies).

Since $r \leq 2$, the bounds in Theorem 16 simplify slightly, and can be written as follows. In each of the cases below, definitions are given for A and B, and in each case $R \leq A + Bt^{2(p-1)}$, where R is as defined in Theorem 16 (with the same value of t that appears here). Conditions defining the case (i.e. i, ii, iii, iv) come first and are either without brackets or with square brackets [...], the square brackets indicating a redundant condition. Conditions defining the sub-case (i.e. a, b) come last and are in parentheses (...).

Case ia: $3/2 \leq p < \infty$ (and $0 \leq r \leq 1/2$),

$$A = \frac{C^p p^p n^{1-2r} (\ln n)^p}{[p + (1 - 2r) \ln n]^p}$$

$$B = C^p \left(\frac{\ln n}{1 + |2 - 2r - p| \ln n} \right)^{\max\{2-p,0\}} \left(1 + n^{2-2r-p} \right)$$

Case ib: $3/2 \leq p < \infty$ (and $1/2 < r \leq 2$),

$$A = \frac{C^p (\ln n)^p}{1 + (2r - 1) \ln n}$$

$$B = C^p \left(\frac{\ln n}{1 + |2 - 2r - p| \ln n} \right)^{\max\{2-p,0\}} \left(1 + n^{2-2r-p} \right)$$

Case iia: $1 \leq p < 3/2$, $p \geq 3/2 - 2r$, $p \neq 2 - 2r$ (and $0 \leq r \leq 1/2$),

$$A = \frac{C n^{1-2r} (\ln n)^p}{[1 + (1 - 2r) \ln n]^p}$$

$$B = C \frac{1 + n^{2-2r-p}}{1 + |2 - 2r - p| \ln n} \ln n$$

Case iib: $1 \leq p < 3/2$, $[p \geq 3/2 - 2r]$, $[p \neq 2 - 2r]$ (and $1/2 < r \leq 2$),

$$A = C \frac{(\ln n)^p}{1 + (2r - 1) \ln n}$$

$$B = C \frac{1 + n^{2-2r-p}}{1 + |2 - 2r - p| \ln n} \ln n$$

Case iii: $1 \leq p < 3/2$ and $p < 3/2 - 2r$,

$$A = C n^{1-2r}$$

$$B = C n^{2-2r-p} \left(\frac{\ln n}{1 + (1 - 4r) \ln n} \right)^{p-1}$$

Case iv: $1 \leq p < 3/2$ and $p = 2 - 2r$,

$$A = C \min \left\{ (1 - 2r)^{-1}, \ln n \right\} n^{1-2r}$$

$$B = C (\ln n)^{2-p}$$

27

By Theorem 16, $\gamma_n(A_1) \geq \gamma_n(A_0) \geq 1 - C\exp(-t^2/2)$. Since A_1 is convex, a bound on the gradient transfers directly to a bound on the Lipschitz constant (using Lemma 15 to ignore points of non-differentiability), and $Lip(\psi|_{A_1}) \leq C^p R^{1/2}$. We may now extend the restriction $\psi|_{A_1}$ to a Lipschitz function $\psi^* : \mathbb{R}^n \to \mathbb{R}$ such that $Lip(\psi^*) = Lip(\psi|_{A_1})$. By a result of Schechtman (see his comments near the end of the paper), as long as $k \leq ct^2$, with probability at least $1 - C\exp(-ct^2)$, the following event occurs: for all $\theta \in S^{k-1}$,

$$|\psi^*(G\theta) - \mathbb{E}\psi^*(G\theta)| \leq tLip(\psi^*) \tag{23}$$

Here it is essential to have a result that applies to Lipschitz functions besides just norms (the most typical application of Schechtman's result is to norms). We now show that for all $\theta \in S^{k-1}$, $G\theta \in A_1$ and therefore $\psi^*(G\theta) = \psi(G\theta)$. To do this, consider a $1/4$-net $\mathcal{N} \subset S^{k-1}$ with cardinality $|\mathcal{N}| \leq 12^k$. By the union bound, with probability at least $1 - 12^k C\exp(-ct^2) \geq 1 - C\exp(-c_2 t^2)$, for all $\omega \in \mathcal{N}$, $G\omega \in A_0$. Now for any $\theta \in S^{k-1}$ write $\theta = \sum_{i=0}^{\infty} \varepsilon_i \omega_i$, where $0 \leq \varepsilon_i \leq 4^{-i}$ and $\omega_i \in \mathcal{N}$. Then

$$|G\theta|_\sharp \leq \sum_{i=0}^{\infty} 4^{-i} |G\omega_i|_\sharp \leq \frac{4}{3}S$$

which shows that $G\theta \in A_1$. (23) can then be written as

$$|\psi(G\theta) - \mathbb{E}\psi^*(G\theta)| \leq tC^p R^{1/2}$$

valid for all $\theta \in S^{k-1}$. It is an elementary calculation that concentration about any point implies concentration about the median, with slightly modified constant, so

$$\left|\frac{\psi(G\theta)}{\mathbb{M}\psi(G\theta)} - 1\right| \leq \frac{C^p t R^{1/2}}{\mathbb{M}\psi(G\theta)} \tag{24}$$

Choose t so that

$$4p^{-1}\frac{C^p t R^{1/2}}{\mathbb{M}\psi(G\theta)} = \varepsilon \tag{25}$$

and assume that $\varepsilon \in (0, 2/p)$ so that $p\varepsilon/4 < 1/2$ (in order to satisfy $|1 - u| \leq 1/2$ below). The bounds for $\varepsilon \in [2/p, 1/2)$ will follow from the bounds for $\varepsilon \in (0, 2/p)$ by monotonicity and by changing the value of C that appears in C^p. Using $\left|1 - u^{1/p}\right| \leq p^{-1} 2^{1-1/p} |1 - u|$, which holds when $|1 - u| \leq 1/2$, with $u = \psi(G\theta)/\mathbb{M}\psi(G\theta)$, (24) implies

$$\left|\frac{\psi(G\theta)^{1/p}}{\mathbb{M}\psi(G\theta)^{1/p}} - 1\right| \leq 2p^{-1}\frac{C^p t R^{1/2}}{\mathbb{M}\psi(G\theta)} < \varepsilon$$

It then follows from homogeneity that

$$(1 - \varepsilon)|x| \mathbb{M}|Ge_1|_{r,p} \leq |Gx|_{r,p} \leq (1 + \varepsilon)|x| \mathbb{M}|Ge_1|_{r,p}$$

for all $x \in \mathbb{R}^k$. (25) can be used to bound t in terms of ε, since R has been expressed in terms of t in Theorem 16, and in Lemma 13 $\mathbb{M}\psi(G\theta)$ is bounded below in t. The sufficient condition $k \leq ct^2$ and the probability bound $1 - C\exp(-ct^2)$ can then be

written in terms of ε. The inversion can be simplified by converting a sum to a max and using the fact that if $f(x) = \max\{g(x), h(x)\}$ with g, h continuous and increasing, then $f^{-1}(x) = \min\{g^{-1}(x), h^{-1}(x)\}$. The result is that,

$$t^2 \geq \min\left\{c^p A^{-1}\left(\mathbb{M}\sum_{i=1}^n i^{-r} X_{[i]}^p\right)^2 \varepsilon^2, cB^{-1/p}\left(\mathbb{M}\sum_{i=1}^n i^{-r} X_{[i]}^p\right)^{2/p} \varepsilon^{2/p}\right\}$$

where A and B are as defined in cases i-iv above, and $X = Ge_1$ follows the standard normal distribution in $\mathbb{R}^n$. Cases ib and iib split into cases ib*, ib**, iib* and iib**, depending on whether $1/2 < r \leq 1$ or $1 < r \leq 2$. The final bounds can then be written as in the statement of the theorem. ∎

9   The general case

Here we study the norm

$$|x|_{\omega,p} = \left(\sum_{i=1}^n \omega_i x_{[i]}^p\right)^{1/p}$$

where $(\omega_i)_1^n$ is any non-increasing sequence in $[0, 1]$ with $\omega_1 = 1$. Our main interest is in the case $1 \leq p < \infty$, however our proof will force us to consider also $0 < p < 1$. The proof of Lemma 14 generalizes easily and we see that

$$\sup\left\{\left(\sum_{i=1}^n \omega_i \theta_{[i]}^p\right)^{1/p} : \theta \in S^{n-1}\right\} = \left\{\begin{array}{ll} \left(\sum_1^n \omega_i^{\frac{2}{2-p}}\right)^{\frac{2-p}{2p}} & : \quad p \in [1, 2) \\ 1 & : \quad p \in [2, \infty) \end{array}\right.$$

and generalizing Lemma 13, for $0 < p < \infty$,

$$\mathbb{M}|X|_{\omega,p} \leq C\left(\sum_{i=1}^n \omega_i \left(\ln\frac{n}{i}\right)^{p/2}\right)^{1/p}$$

with the reverse inequality with C replaced by c. For $1 \leq p < \infty$, $|\cdot|_{\omega,p}$ is a norm, while for $0 < p < 1$ the triangle inequality is replaced with

$$|x + y|_{\omega,p} \leq 2^{1/p}\left(|x|_{\omega,p} + |y|_{\omega,p}\right)$$

(proof just as in the classical case of ℓ_p^n), and $|\cdot|_{\omega,p}$ is no longer a norm but a quasi-norm. For an infinite sum, using induction one can show that

$$\left|\sum_{i=1}^\infty x_i\right|_{\omega,p} \leq \sum_{i=1}^\infty 2^{i/p} |x_i|_{\omega,p}$$

As sometimes happens, the general result is easier to state and prove than is the special case, since the complexity is hidden. The proof of Theorem 18 below is a variation of the proof of Theorem 17, using the quasi-norm property instead of the triangle inequality.

Theorem 18 *There exists a universal constant $c > 0$ such that the following is true. Let*

$$(n, k, \varepsilon, p) \in \mathbb{N} \times \mathbb{N} \times (0, 1/2) \times [1, \infty)$$

and let $(\omega_i)_1^n$ be any non-increasing sequence in $[0, 1]$ with $\omega_1 = 1$. Let $|\cdot|_{\omega, p}$ denote the corresponding Lorentz norm, and let G be a random $n \times k$ matrix with i.i.d. standard normal random variables as entries. If $p \neq 1$ set

$$d = \left(1 + \frac{1}{p-1}\right)^{-1} \min \left\{ \frac{c^p \left(\sum_{i=1}^n \omega_i \left(\ln \frac{n}{i}\right)^{p/2}\right)^2 \varepsilon^2}{\sum_{i=1}^n \omega_i^2 \left(\ln \frac{n}{i}\right)^{p-1}}, cB^{-1/p} \left(\sum_{i=1}^n \omega_i \left(\ln \frac{n}{i}\right)^{p/2}\right)^{2/p} \varepsilon^{2/p} \right\}$$

where

$$B = \begin{cases} \sum_{i=1}^n \omega_i^2 i^{-(p-1)} & : \quad 1 < p < 3/2 \\ \left(\sum_{i=1}^n \omega_i^{\frac{2}{2-p}}\right)^{2-p} & : \quad 3/2 \leq p < 2 \\ 1 & : \quad 2 \leq p < \infty \end{cases}$$

and if $p = 1$ set

$$d = \frac{c \left(\sum_{i=1}^n \omega_i \left(\ln \frac{n}{i}\right)^{1/2}\right)^2 \varepsilon^2}{\sum_{i=1}^n \omega_i^2}$$

Assume that $k \leq d$. With probability at least $1 - 2e^{-d}$ the following event occurs: for all $x \in \mathbb{R}^k$,

$$(1 - \varepsilon) |x| \, \mathbb{M} \, |Ge_1|_{\omega, p} \leq |Gx|_{\omega, p} \leq (1 + \varepsilon) |x| \, \mathbb{M} \, |Ge_1|_{\omega, p}$$

Proof. First, assume that $p \neq 1$ (we leave the case $p = 1$ to the reader). For $x \in \mathbb{R}^n$, set

$$\psi(x) = \sum_{i=1}^n \omega_i x_{[i]}^p \qquad |x|_\sharp = \left(\sum_{i=1}^n \omega_i^2 x_{[i]}^{2(p-1)}\right)^{\frac{1}{2(p-1)}}$$

and so

$$|\nabla \psi(x)| = p \left(\sum_{i=1}^n \omega_i^2 x_{[i]}^{2(p-1)}\right)^{1/2}$$

which is valid for all x with distinct non-zero coordinates. The points of non-differentiability are not a problem, by Lemma 15. Consider any $\theta \in S^{n-1}$. Set $R = A + C^p B t^{2(p-1)}$, where $A = C^p \sum_{i=1}^n \omega_i^2 \left(\ln \frac{n}{i}\right)^{p-1}$ and B is defined in the statement of the theorem. For $j \in \{0, 1\}$, set

$$A_j = \left\{ x \in \mathbb{R}^n : |x|_\sharp \leq \left(\frac{4q}{3}\right)^j R^{\frac{1}{2(p-1)}} \right\}$$

where

$$q = \begin{cases} 2^{-1/(p-1)} & : \quad 1 < p < 3/2 \\ 1 & : \quad 3/2 \leq p < \infty \end{cases}$$

For $1 < p < 3/2$ we now use (4) from Lemma 12 on order statistics from the standard normal distribution, as used in Case II of Theorem 16, and for $3/2 \leq p < \infty$ we use

Gaussian concentration applied to $|\cdot|_\sharp$. The result of this is that $\gamma_n (A_1) \geq \gamma_n (A_0) \geq 1 - C \exp\left(-ct^2\right)$. It follows by definition of A_1 and by its convexity, and the expression for $|\nabla \psi|$ that $\mathrm{Lip}\left(\psi|_{A_1}\right) \leq C^p R^{1/2}$. By the extension property of real valued Lipschitz functions on metric spaces, $\psi|_{A_1}$ can be extended to a function $\psi^* : \mathbb{R}^n \mapsto \mathbb{R}$ such that $\mathrm{Lip}\left(\psi^*\right) = \mathrm{Lip}\left(\psi|_{A_1}\right)$. By Schechtman's result (Theorem 7 here), as long as $k \leq ct^2$, the following event occurs with probability at least $1 - C \exp\left(-ct^2\right)$: for all $\theta \in S^{k-1}$,

$$\left|\psi^* (G\theta) - \mathbb{M}\psi^* (G\theta)\right| \leq Ct\mathrm{Lip}\left(\psi^*\right)$$

We now show that for all $\theta \in S^{k-1}$, $G\theta \in A_1$ and therefore $\psi (G\theta) = \psi^* (G\theta)$. Let $\mathcal{N} \subset S^{k-1}$ be a $q/4$-net in S^{k-1}. The standard volumetric bound shows that $\mathcal{N}$ can be chosen so that $|\mathcal{N}| \leq (12/q)^k$. By the bound on $\gamma_n (A_0)$ and the union bound, with probability at least $1 - C (12/q)^k \exp\left(-ct^2\right) \geq 1 - C \exp\left(-c't^2\right)$, the following event occurs: for all $\theta \in \mathcal{N}$, $G\theta \in A_0$, i.e.

$$|G\theta|_\sharp \leq R^{\frac{1}{2(p-1)}}$$

Now for any $\theta \in S^{k-1}$, write $\theta = \sum_1^\infty \varepsilon_i \theta_i$, where $0 \leq \varepsilon_i \leq (q/4)^i$ and $\theta_i \in \mathcal{N}$, to conclude (using the quasi-norm property),

$$|G\theta|_\sharp \leq R^{\frac{1}{2(p-1)}} \sum_{i=0}^\infty (q)^{-(i-1)} (q/4)^i \leq \frac{4q}{3} R^{\frac{1}{2(p-1)}}$$

So $G\theta \in A_1$ as desired, and $\psi^* (G\theta) = \psi (G\theta)$. So, conditioning on the events dealt with above, for all $\theta \in S^{k-1}$,

$$|\psi (G\theta) - \mathbb{M}\psi^* (G\theta)| \leq Ct\mathrm{Lip}\left(\psi\right)$$

The rest of the proof is identical to the proof of Theorem 17, and we may change the coefficient of e^{-d} from C to 2 by changing the value of c. ∎

The bounds in Theorem 18 are non-optimal when:

• the coefficient sequence $(\omega_i)_1^n$ approximates $(i^{-r})_1^n$, for $p = 2 - 2r$ $(1/4 < r < 1/2)$, and in this case we refer the reader to Theorem 17 and Corollary 2,
• when $r > 1$ in which case we refer the reader to Theorem 3, and when
• $p \to 1$ $(p \neq 1)$, which is due to the presence of the factor $1/(1 + 1/(p - 1))$, but this is not an issue in the asymptotic case $n \to \infty$ while ω and p remain fixed. We refer the reader to Theorem 17 and Corollary 4 for estimates that do not include this extra factor.

References

[1] Bobkov, S. G., Nayar, P., Tetali, P.: Concentration properties of restricted measures with applications to non-Lipschitz functions. Geometric aspects of functional analysis, 25-53, Lecture Notes in Math. 2169, Springer, Cham, 2017.

[2] Fresen, D. J.: Explicit Euclidean embeddings in permutation invariant normed spaces. Adv. Math. 266, 1–16 (2014).

[3] Fresen, D. J.: Variations and extensions of the Gaussian concentration inequality, Part I. To appear in Quaest. Math. (pending revision). An older version of Part I and Part II combined available at arXiv:1812.10938.

[4] Fresen, D. J.: Variations and extensions of the Gaussian concentration inequality, Part II. To appear in Electron. J. Probab. (pending revision). An older version of Part I and Part II combined available at arXiv:1812.10938.

[5] Gordon, Y.: Some inequalities for Gaussian processes and applications. Israel J. Math. 50, 265-289, 1985.

[6] Huang, H., Wei, F.: Upper bound for the Dvoretzky dimension in Milman-Schechtman theorem. Geometric aspects of functional analysis, 181–186, Lecture Notes in Math., 2169, Springer, Cham, 2017.

[7] Lindenstrauss, Joram; Tzafriri, Lior Classical Banach spaces. I. Sequence spaces. Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92. Springer-Verlag, Berlin-New York, 1977.

[8] Milman, V. D.: A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. (Russian) Funkcional. Anal. i Priložen. 5 (4), 28-37 (1971)

[9] Milman, V. D., Schechtman, G., Asymptotic Theory of Finite-dimensional Normed Spaces. Lecture Notes in Math., 1200, Springer-Verlag, Berlin, 1986.

[10] Milman, V. D., Schechtman, G.: Global versus local asymptotic theories of finite-dimensional normed spaces. Duke Math J. 90, 73-93 (1997)

[11] Paouris, G., Valettas, P., Zinn, J.: Random Version of Dvoretzky's theorem in ℓ_p^n. Stochastic Process. Appl. 127 (10), 3187-3227 (2017)

[12] Paouris, G., Valettas, P.: Dichotomies, structure, and concentration in normed spaces. Adv. Math. 332, 438–464 (2018)

[13] Pisier, G.: Probabilistic methods in the geometry of Banach spaces. CIME, Varenna, 1985. Lecture Notes in Mathematics 1206, 167-241 (1986)

[14] Schechtman, G.: A remark concerning the dependence on ϵ in Dvoretzky's theorem. Geometric aspects of functional analysis (1987-88), 274-277, Lecture Notes in Math., 1376, Springer, Berlin, 1989.

[15] Schechtman, G.: Euclidean sections of convex bodies. Asymptotic geometric analysis, 271-288, Fields Inst. Commun., 68, Springer, New York, 2013.

[16] Schechtman, G.: Two observations regarding embedding subsets of Euclidean spaces in normed spaces. Adv. Math. 200, 125-135 (2006)

[17] Tikhomirov, K.: Almost Euclidean sections in symmetric spaces and concentration of order statistics. J. Funct. Anal. 265 (9), 2074–2088 (2013)