A SUBGAUSSIAN EMBEDDING THEOREM

BY

Shahar Mendelson*

*Centre for Mathematics and its Applications, The Australian National University,
Canberra, ACT 0200, Australia
and Department of Mathematics, Technion,
I.I.T. Haifa 32000, Israel
e-mail: shahar.mendelson@anu.edu.au*

AND

Nicole Tomczak-Jaegermann**

*Department of Mathematical and Statistical Sciences, University of Alberta,
Edmonton, Alberta, Canada T6G 2G1
e-mail: nicole@ellpspace.math.ualberta.ca*

ABSTRACT

We prove a subgaussian extension of a Gaussian result on embedding sub-
sets of a Euclidean space into normed spaces. Using the concentration of
a random subgaussian vector around its mean we obtain an isomorphic
(rather than almost isometric) result, under an additional cotype assump-
tion on the normed space considered.

1. Introduction

The motivation behind this paper is the following Gaussian result on embedding
subsets of a Euclidean space into normed spaces. Let (g_i) and (g_{ij}) be indepen-
dent standard Gaussian random variables and set $(e_i)_{i=1}^k$ to be the standard

unit vector basis in $\mathbb{R}^k$. We denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^k$ and by S^{n-1} the Euclidean sphere in $\mathbb{R}^n$.

THEOREM 1.1: *There exists an absolute constant c for which the following holds. Set $n, k \geq 1$, consider $T \subset S^{n-1}$ and let $E = (\mathbb{R}^k, \| \ \|_E)$ be a normed space such that for every $x \in \mathbb{R}^k$, $\|x\|_E \leq |x|$. Fix $\varepsilon > 0$ and assume that*

$$(1.1) \qquad \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} g_i t_i \right| \leq c \varepsilon \, \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E.$$

If the random operator $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ is defined by $\Gamma t = \sum_{i=1}^{k} \sum_{j=1}^{n} g_{ij} t_j e_i$, then there is a realization of Γ such that for every $t \in T$,

$$(1 - \varepsilon) \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E \leq \|\Gamma t\|_E \leq (1 + \varepsilon) \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E.$$

Theorem 1.1 is a simple application of the Gaussian min-max theorem proved originally by Y. Gordon in [G1] and is an easy modification of the method of the proof of Dvoretzky's theorem and other applications in [G1] and [G2] (see also [G3]). The theorem has been recently rediscovered by G. Schechtman [S] who has given it a new proof using the majorizing measures approach. Another proof is an unpublished argument by G. Pisier [Pi4], based on his Gaussian concentration measure theorem (cf. e.g., [Pi3]). The formulation of Theorem 1.1 is taken from [S]. One should note that all three approaches are limited to the case where Γ is a Gaussian operator.

In this article we show that one can use the concentration of a random vector $\|\Gamma t\|_E$ around its mean to prove an isomorphic (rather than almost isometric) analog of Theorem 1.1, where the Gaussian operator is replaced by an arbitrary **subgaussian** operator, under a cotype assumption on the space E.

To explain the notion of a subgaussian operator we use here, recall that the ψ_2 norm of a random variable X is defined by

$$\|X\|_{\psi_2} = \inf \left\{ u > 0 : \mathbb{E} \exp \left(|X|^2 / u^2 \right) \leq 2 \right\}.$$

Let μ be a symmetric measure on $\mathbb{R}^n$ which satisfies that for every $t \in \mathbb{R}^n$, $\mathbb{E} \langle X, t \rangle^2 = |t|^2$ and $\| \langle X, t \rangle \|_{\psi_2} \leq L|t|$. In other words, μ is a symmetric isotropic measure on $\mathbb{R}^n$ and linear functionals exhibit a subgaussian decay of the order of $\exp(-cu^2/L^2)$ where c is an absolute constant and $u \geq 1$.

The operator $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ is said to be subgaussian with a constant L if $\Gamma = \sum_{i=1}^{k} \langle X_i, \cdot \rangle e_i$, where $(X_i)_{i=1}^{k}$ are independent random vectors, distributed according to a L subgaussian measure on $\mathbb{R}^n$.

Let us recall the concept of cotype of a Banach space which plays a key role in our approach. A Banach space E has cotype $q \geq 2$ with a constant β_q if for all finite sequences (z_i) in E,

$$\left(\sum_i \|z_i\|_E^q \right)^{1/q} \leq \beta_q \, \mathbb{E} \left\| \sum_i \varepsilon_i z_i \right\|_E,$$

where $(\varepsilon_i)_i$ is a sequence of independent Bernoulli random variables. Now let us formulate our main result.

THEOREM 1.2: *There exist absolute constants $c_1, c_2, c_3 > 0$ for which the following holds. Set $n, k \geq 1$, consider $T \subset S^{n-1}$ and let $E = (\mathbb{R}^k, \| \ \|_E)$ be a normed space satisfying $\|x\|_E \leq |x|$, for $x \in \mathbb{R}^k$. Fix $L > 0$, assume that E has cotype q with a constant β_q and that*

$$(1.2) \qquad \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} g_i t_i \right| \leq c(L, q, \beta_q) \, \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E / \sqrt{\log k},$$

where $c(L, q, \beta_q) := c_3 / L^2 \beta_q^2 \sqrt{q}$.

If Γ is a subgaussian operator with a constant L then, with probability close to 1, for every $t \in T$

$$(1.3) \qquad c'(L, q, \beta_q) \, \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E \leq \|\Gamma t\|_E \leq c_2 L \, \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E,$$

where $c'(L, q, \beta_q) := c_1 / L\beta_q \sqrt{q}$. Moreover, if Γ is an operator with independent Bernoulli random entries then the assertion of the theorem remains true with (1.2) replaced by

$$(1.4) \qquad \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} g_i t_i \right| \leq c(L, q, \beta_q) \, \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E.$$

To illustrate the main theorem we shall formulate just one result for random ± 1 embeddings of the Euclidean space.

COROLLARY 1.3: *Let $E = (\mathbb{R}^k, \| \ \|_E)$ be a normed space of cotype q with a constant β_q such that for every $x \in \mathbb{R}^k$, $\|x\|_E \leq |x|$. Assume that*

$$n \leq (c_1/\beta_q^2 q)\left(\mathbb{E}\bigg\| \sum_{i=1}^{k} g_i e_i \bigg\|_E\right)^2,$$

where $c_1 > 0$ is an absolute constant. Let $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ be a random operator with independent Bernoulli entries. Then, with probability close to 1, the random subspace $F := \Gamma(\mathbb{R}^n) \subset \mathbb{R}^k$ spanned by ± 1 vectors satisfies $c'A|z| \leq \|z\|_E \leq c''A|z|$ for all $z \in F$, where $A = \mathbb{E}\|\sum g_i e_i\|_E/\sqrt{k}$, and $c', c'' > 0$ depend on L, q, β_q.

In particular, if the Euclidean unit ball on $\mathbb{R}^k$ is the maximal volume ellipsoid for the unit ball B_E of E then the assertion holds once $n \leq (c_1/\beta_q^2 q)k^{2/q}$.

We end this introduction with some notation and definitions we will use throughout this note. For a finite set A, denote by $|A|$ its cardinality. With a minor abuse of notation, we also denote by $|\cdot|$ the fixed Euclidean structure we consider, and let (e_i) be the standard unit vector basis with respect to that structure. If $A \subset \mathbb{R}^m$, let $\ell_*(A) = \mathbb{E}\sup_{a \in A} |\sum_{i=1}^{m} g_i a_i|$, where $(g_i)_{i=1}^{m}$ are independent standard Gaussian variables. Note that if $A = B_{E^*}$, the unit ball of the dual space to E, then $\ell_*(A) = \mathbb{E}\|\sum_{i=1}^{k} g_i e_i\|_E$. Finally, all absolute constants are denoted by $c, c_1, \ldots$. Their values may change from line to line.

2. Preliminaries - Subgaussian vectors

We will be interested in independent, symmetric random variables that exhibit a subgaussian tail behaviour. We begin this section by recalling some of the basic facts concerning random variables with certain decay properties. For basic properties of such random variables we refer the reader to [LTa, DG].

Definition 2.1: For $\alpha \geq 1$ the ψ_α norm, of a random variable X is defined by

$$\|X\|_{\psi_\alpha} = \inf \left\{ \lambda > 0 : \mathbb{E} \exp(|X|^\alpha / \lambda^\alpha) \leq 2 \right\}.$$

It is standard to verify that if X has a bounded ψ_α norm, then for $t \geq 1$

$$\mathbb{P}\left(|X| \geq t\right) \leq 2 \exp(-ct^\alpha / \|X\|_{\psi_\alpha}^\alpha)$$

where c is an absolute constant. The converse is also true, and if X has a tail bounded by $\exp(-t^\alpha / K^\alpha)$ then $\|X\|_{\psi_\alpha} \leq c_1 K$.

Another basic property of the ψ_2 norm is that if $X_1, \ldots, X_n$ are independent and centered then

$$(2.1) \qquad \left\| \sum_{i=1}^n a_i X_i \right\|_{\psi_2} \leq c \left(\sum_{i=1}^n a_i^2 \|X_i\|_{\psi_2}^2 \right)^{1/2},$$

where c is an absolute constant. If $(X_i)_{i=1}^n$ are symmetric, (2.1) follows from an easy calculation using the moment generating function of the sum and the fact that for every $t > 0$, $\mathbb{E} \exp(tX) \leq \exp(t^2 b^2)$, where $b = c_1 \|X\|_{\psi_2}$. The extension of (2.1) from the symmetric case to the centered one is evident from a standard symmetrization argument.

We require two additional preliminary results. First, a bound due to Pisier [Pi1] (cf. also [VW] Section 2.2) on the ψ_α norm of a maximum of k random variables, which we formulate only for $\alpha = 1$.

LEMMA 2.2: *There exists an absolute constant c for which the following holds. Let $X_1, \ldots, X_k$ be random variables. Then,*

$$\left\| \max_{1 \leq i \leq k} |X_i| \right\|_{\psi_1} \leq c \max_{1 \leq i \leq k} \|X_i\|_{\psi_1} \log k.$$

The second preliminary result deals with the ψ_1 behaviour of a sum of independent ψ_1 random variables, and is due to Talagrand (see, e.g. [LTa], Theorem 6.21).

THEOREM 2.3: *There exists an absolute constant c such that if $X_1, \ldots, X_k$ are independent, centered random variables then*

$$\left\| \sum_{i=1}^k X_i \right\|_{\psi_1} \leq c \left(\left\| \sum_{i=1}^k X_i \right\|_{L_1} + \left\| \max_{1 \leq i \leq k} |X_i| \right\|_{\psi_1} \right).$$

Consider a random process $\{Z_t : t \in T\}$ indexed by a metric space (T, d). We say that the process is subgaussian with respect to the metric d if it satisfies

the increment condition that for every $x, y \in T$ and every $u \geq 1$,

$$\mathbb{P}\left(|Z_x - Z_y| \geq ud(x,y)\right) \leq 2\exp(-u^2/2)$$

(the constants 2 and $1/2$ can be replaced by any other fixed constants, as can be the restriction $u \geq 1$).

The **generic chaining** mechanism allows one to obtain an exponential deviation inequality for $\sup_{t \in T}|Z_t - Z_{t_0}|$ where t_0 is an arbitrary point in T using an increment condition. To that end we require the definition of the γ_2 functional [Ta3].

Definition 2.4: Let (T, d) be a metric space. An *admissible sequence* of T is a collection of subsets of T, $\{T_s : s \geq 0\}$, such that for every $s \geq 1$, $|T_s| = 2^{2^s}$ and $|T_0| = 1$. Define the γ_2 functional by

$$\gamma_2(T, d) = \inf \sup_{t \in T} \sum_{s=0}^{\infty} 2^{s/2} d(t, T_s),$$

where $d(t, T_s)$ is the distance between the set T_s and t, and the infimum is taken with respect to all admissible sequences of T.

It is well-known that the $\gamma_2(T)$ functional, with respect to a Euclidean metric, is connected to the behaviour of a Gaussian process indexed by T. Indeed, let $\{G_t : t \in T\}$ be a centered Gaussian process indexed by a set T, and for every $s, t \in T$, let $d^2(s, t) = \mathbb{E}|G_s - G_t|^2$. Then

$$c_1 \gamma_2(T, d) \leq \mathbb{E} \sup_{t \in T} |G_t| \leq c_2 \gamma_2(T, d),$$

where c_1 and c_2 are absolute constants. The upper bound is due to Fernique [F] and the lower bound is Talagrand's majorizing measure theorem [Ta1]. The proof of both parts and the most recent survey on the topic can be found in [Ta3].

Let us state a typical deviation estimate that follows from a generic chaining argument. Proofs of a similar flavor can also be found in [Ta3, Chapter 1].

THEOREM 2.5: *There exists an absolute constant c for which the following holds. Let (T, d) be a metric space and assume the process $\{Z_t : t \in T\}$ is subgaussian with respect to the metric d. Then, for any $u \geq 4$ and any $t_0 \in T$,*

$$\mathbb{P}\left(\sup_{t \in T}|Z_t - Z_{t_0}| \geq u\gamma_2(T, d)\right) \leq 2\exp\left(-cu^2\right).$$

Let $(\xi_i)_{i=1}^k$ be i.i.d. symmetric random variables which satisfy the subgaussian condition $\|\xi\|_{\psi_2} \leq L$. We are interested in the moments of the random variable $\|\sum_{i=1}^k \xi_i e_i\|_E$, where $\|\ \|_E$ is a norm on $\mathbb{R}^k$. An important part in our argument is the fact that under a cotype assumption on E, these moments are equivalent to the analogous Gaussian moments.

The first observation in this direction compares the $L_p(E)$ norm of a subgaussian vector with the $L_1(E)$ norm of a Gaussian one. The proof we present here is due to Jain and Marcus [JM]. It was shown to us by G. Pisier and we present it for the sake of completeness.

LEMMA 2.6: *There exists an absolute constant c for which the following holds. Let $p \geq 1$ and set $(\xi_i)_{i=1}^k$ to be independent symmetric random variables which satisfy that $\mathbb{E}\xi_i^2 = 1$ and $\|\xi_i\|_{\psi_2} \leq L$. Then*

$$\left\| \sum_{i=1}^k \xi_i e_i \right\|_{L_p(E)} \leq cL\sqrt{p} \left\| \sum_{i=1}^k g_i e_i \right\|_{L_1(E)}.$$

Proof. Let ξ be a symmetric L subgaussian random variable. Then there is a centered Gaussian variable denoted by Y with variance $\sigma_Y^2 = cL^2$, such that for every $t \geq 1$,

$$\mathbb{P}\left(|\xi| \geq t\right) \leq \mathbb{P}\left(|Y| \geq t\right).$$

Let ξ^* and Y^* be the decreasing rearrangements of ξ and Y, respectively, defined on $[0,1]$. That is, $\xi^*(x) = \inf\{t : \mathbb{P}(|\xi| > t) \leq x\}$ for all $x \in [0,1]$, and Y^* is defined similarly. Note that Y^* is continuous on $[0,1]$ and for every $x \in [0,1]$, $\xi^*(x) \leq 1 + Y^*(x)$.

Indeed, observe that if f^* is a decreasing rearrangement of f then for every t, $f^*(\mathbb{P}(|f| \geq t)) \leq t$, and the equality holds if f^* is continuous at the point $x = \mathbb{P}(|f| \geq t)$.

Now, for $t \geq 1$ set $x_t := \mathbb{P}(|\xi| \geq t)$ and $y_t := \mathbb{P}(|Y| \geq t)$. Then $y_t \geq x_t$ and $\xi^*(x_t) \leq Y^*(y_t) \leq Y^*(x_t)$. This implies that $\xi^*(x) \leq Y^*(x)$ for all $x \in [0, x_1]$, where $x_1 = \mathbb{P}(|\xi| \geq 1)$. For $x \geq x_1$, $\xi^*(x) \leq 1$, and Y^* is non-negative. The combination of the two observations gives the desired inequality.

Let $(\xi_i)_{i=1}^k$ be as in the lemma and consider $(\xi_i^*)_{i=1}^k$. If $(\varepsilon_i)_{i=1}^k$ are independent Bernoulli variables then $\varepsilon_i \xi_i^*$ has the same distribution as ξ_i, for $i = 1, \ldots, k$. A similar fact holds for $(Y_i^*)_{i=1}^k$ and $(Y_i)_{i=1}^k$.

By the Kahane–Khintchine inequality and the contraction principle for Bernoulli processes (see, e.g. [LTa]), for every $p \geq 1$,

$$\mathbb{E}\Big\|\sum_{i=1}^{k}\xi_i e_i\Big\|_E^p = \int_0^1 \mathbb{E}_\varepsilon\Big\|\sum_{i=1}^{k}\varepsilon_i\xi_i^*(x)e_i\Big\|_E^p \leq cp^{p/2}\int_0^1 \mathbb{E}_\varepsilon\Big\|\sum_{i=1}^{k}\varepsilon_i(Y_i^*(x)+1)e_i\Big\|_E.$$

Hence,

$$\Big\|\sum_{i=1}^{k}\xi_i e_i\Big\|_{L_p(E)} \leq c\sqrt{p}\Big(\Big\|\sum_{i=1}^{k}\varepsilon_i e_i\Big\|_{L_1(E)} + \Big\|\sum_{i=1}^{k}Y_i e_i\Big\|_{L_1(E)}\Big).$$

The claim now follows from the definition of Y and the fact that Rademacher averages are dominated by Gaussian ones (see, e.g. [LTa] or [MS], Appendix II). ∎

Lemma 2.6 provides the "general" direction in the equivalence between the $L_p(E)$ norms of subgaussian and Gaussian vectors—that holds without any structural assumptions on E. The reverse direction is not true for an arbitrary space E, but rather, requires cotype assumptions. The following theorem was proved in [MaPi] (see also [Pi2]). The proof for the Rademacher case can be found in Appendix II of [MS]. It is straightforward to check that only minimal modifications to that proof are needed to show that it actually holds for an arbitrary subgaussian symmetric variable, and thus we omit the proof.

THEOREM 2.7: *For every $p \geq 1$ there exists a constant $c(p)$ for which the following holds. Let E be a Banach space of cotype q with a constant β_q. Let $(\xi_i)_{i=1}^{k}$ be independent symmetric random variables such that $\mathbb{E}\xi_i^2 = 1$ and $\|\xi_i\|_{\psi_2} \leq L$. Then, for every $x_1, \ldots, x_k \in E$,*

$$\Big\|\sum_{i=1}^{k}g_i x_i\Big\|_{L_p(E)} \leq c(p)L\sqrt{q}\beta_q\Big\|\sum_{i=1}^{k}\xi_i x_i\Big\|_{L_p(E)}.$$

Moreover, $c(p)$ is an absolute constant for $p \leq q$.

For our needs it suffices to obtain an equivalence result for $p = 1$.

COROLLARY 2.8: *There exist absolute constants c_1 and c_2 for which the following holds. Consider $(\xi_i)_{i=1}^{k}$ as above and assume that E has cotype q with a constant β_q. Then,*

$$(c_1/L)\Big\|\sum_{i=1}^{k}\xi_i e_i\Big\|_{L_1(E)} \leq \Big\|\sum_{i=1}^{k}g_i e_i\Big\|_{L_1(E)} \leq c_2 L\sqrt{q}\beta_q\Big\|\sum_{i=1}^{k}\xi_i e_i\Big\|_{L_1(E)}.$$

3. The main result

The proof of Theorem 1.2 is based on three ingredients. The first one is a concentration result for each one of the random variables

$$Z_t = \left\| \sum_{i=1}^{k} \langle X_i, t \rangle e_i \right\|_E$$

around its mean.

Definition 3.1: Let $k \geq 1$ and set $(\xi_i)_{i=1}^{k}$ to be independent random variables with distribution ξ. A space $E = (\mathbb{R}^k, \| \ \|_E)$ is α concentrated with respect to ξ if

$$\mathbb{P}\left(\left| \left\| \sum_{i=1}^{k} \xi_i e_i \right\|_E - \mathbb{E}\left\| \sum_{i=1}^{k} \xi_i e_i \right\|_E \right| \geq \frac{1}{2}\mathbb{E}\left\| \sum_{i=1}^{k} \xi_i e_i \right\|_E \right)$$
$$\leq 2\exp\left(-\frac{(\mathbb{E}\| \sum_{i=1}^{k} \xi_i e_i \|_E)^2}{\alpha^2} \right).$$

The concentration result for Gaussian or Rademacher vectors (see e.g. [LTa]) shows that if for every $x \in E$, $\|x\|_E \leq |x|$, then E satisfies Definition 3.1 for α which is an absolute constant. We will show later that a slightly weaker estimate $\alpha \sim L\sqrt{\log k}$ holds for an arbitrary k-dimensional space and any L subgaussian random variable. The concentration assumption, combined with an approximation argument, leads to the following technical embedding result.

THEOREM 3.2: *There exist absolute constants c_1 and c_2 for which the following holds. Let μ be an isotropic ψ_2 probability measure on $\mathbb{R}^n$ with a constant L and set X to be distributed according to μ. Let Γ be the corresponding operator, i.e., $\Gamma t = \sum_{i=1}^{k} \langle X_i, t \rangle e_i$ for $t \in \mathbb{R}^n$. Consider $T \subset S^{n-1}$ and for every $t \in T$ denote $\xi_t = \langle X, t \rangle$.*

Let E be a k dimensional space which is α concentrated with respect to ξ_t for every $t \in T$. Fix $\varepsilon > 0$ and set

$$(3.1) \qquad H := L\left(\ell_*(T) + \varepsilon\mathbb{E}\left\| \sum_{i=1}^{k} g_i e_i \right\|_E \right)$$

and $p_0 := \exp\left(c_1 \ell_*^2(T)/\varepsilon^2 - \inf_{t \in T} (\mathbb{E}\| \sum_{i=1}^k \langle X_i, t \rangle e_i \|_E)^2/\alpha^2\right)$. *Then, for every* $s \geq 1$, *with probability at least* $1 - p_0 - 2\exp(-s^2)$, *we have, for every* $t \in T$,

$$(3.2) \qquad \frac{1}{2}\mathbb{E}\|\Gamma t\|_E - c_2 s H \leq \|\Gamma t\|_E \leq \frac{3}{2}\mathbb{E}\|\Gamma t\|_E + c_2 s H.$$

Proof. Let Λ be an ε-net of T with respect to the Euclidean norm $|\cdot|$. By Sudakov's inequality (cf. e.g., [Pi3]), $\log|\Lambda| \leq c_1 \ell_*^2(T)/\varepsilon^2$. Applying the α concentration assumption,

$$\mathbb{P}\left(\exists\, v \in \Lambda : \left\| \left\| \sum_{i=1}^k \langle X_i, v \rangle e_i \right\|_E - \mathbb{E}\left\| \sum_{i=1}^k \langle X_i, v \rangle e_i \right\|_E \right\| \geq \frac{1}{2}\mathbb{E}\left\| \sum_{i=1}^k \langle X_i, v \rangle e_i \right\|_E \right)$$

$$\leq \sum_{v \in \Lambda} \exp\left(-\frac{(\mathbb{E}\| \sum_{i=1}^k \langle X_i, v \rangle e_i \|_E)^2}{\alpha^2} \right)$$

$$\leq \exp\left(c_1 \frac{\ell_*^2(T)}{\varepsilon^2} - \frac{\inf_{t \in T}(\mathbb{E}\| \sum_{i=1}^k \langle X_i, t \rangle e_i \|_E)^2}{\alpha^2} \right) = p_0.$$

Set $U := (T - T) \cap \varepsilon B_2^k$ and observe that $T \subset \Lambda + U$. Thus by an easy approximation argument the theorem will follow from the above estimate, the lemma below and the fact that $\ell_*(T - T) \leq 2\ell_*(T)$. ∎

The proof of the following lemma is based on a standard generic chaining argument [Ta3].

LEMMA 3.3: *Let* $\tilde{T} \subset \mathbb{R}^n$, $\varepsilon > 0$, $\tilde{U} := \tilde{T} \cap \varepsilon B_2^n$ *and consider* E *and* Γ *to be as in Theorem 3.2. Then, for any* $s \geq 1$,

$$\mathbb{P}\left(\sup_{u \in \tilde{U}} \|\Gamma u\|_E \leq c_3 s \tilde{H} \right) \geq 1 - 2\exp(-s^2),$$

where $\tilde{H}$ *is given by* (3.1), *by replacing* T *by* $\tilde{T}$, *and* $c_3 > 0$ *is an absolute constant.*

Proof. Let B_{E^*} denote the unit ball of E^* and consider the process

$$(u, y) \rightarrow Z_{u,y} := \sum_{i=1}^k \langle X_i, u \rangle \langle e_i, y \rangle.$$

indexed by $\tilde{U} \times B_{E^*}$. Clearly, $\sup_{\{u \in \tilde{U},\, y \in B_{E^*}\}} Z_{u,y} = \sup_{u \in \tilde{U}} \|\Gamma u\|_E$. Note that Z is centered and subgaussian with respect to the appropriate metric. Indeed,

the increment associated with the elements (u, y) and (u_1, y_1) satisfies

$$|Z_{u,y} - Z_{u_1,y_1}| \le |Z_{u,y} - Z_{u,y_1}| + |Z_{u,y_1} - Z_{u_1,y_1}|$$
$$= \left| \sum_{i=1}^{k} \langle X_i, u \rangle \langle y - y_1, e_i \rangle \right| + \left| \sum_{i=1}^{k} \langle X_i, u - u_1 \rangle \langle y_1, e_i \rangle \right|.$$

Using (2.1) we can bound the ψ_2 norm of the first term by

$$\| \langle X, u \rangle \|_{\psi_2} \left(\sum_{i=1}^{k} \langle y - y_1, e_i \rangle^2 \right)^{1/2} \le L|u| \cdot |y - y_1| \le L\varepsilon |y - y_1|,$$

where the last inequality follows from $\tilde{U} \subset \varepsilon B_2^k$.

By a similar argument, the second term has a ψ_2 norm at most $L|u-u_1||y_1| \le L|u-u_1|$, because $\|y_1\|_{E^*} \le 1$ and since the condition $\|x\|_E \le |x|$ for every x implies $|y| \le \|y\|_{E^*}$ for every y. Hence, adding the two probabilities corresponding to the two terms above, it follows that for every $s \ge 1$,

$$\mathbb{P}\left(|Z_{u,y} - Z_{u_1,y_1}| \ge c_4 s L \left(\varepsilon |y - y_1| + |u - u_1| \right) \right) \le 2 \exp(-s^2).$$

Applying a generic chaining argument combined with the majorizing measures Theorem,

$$\mathbb{E} \sup_{u,y} |Z_{u,y}| \le cL \left(\gamma_2(\tilde{U}, |\cdot|) + \varepsilon \gamma_2(B_{E^*}, |\cdot|) \right)$$
$$\le c_3 L \left(\ell_*(\tilde{U}) + \varepsilon \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E \right) \le c_3 \tilde{H},$$

and moreover, for every $s \ge 1$

$$\mathbb{P}\left(\sup_{u,y} |Z_{u,y}| \ge c_3 s \tilde{H} \right) \le 2 \exp(-s^2),$$

from which the claim follows. This also completes the proof of Theorem 3.2.   ∎

Remark 3.4: Let us note that in the Gaussian case, a version of Lemma 3.3 (with probability larger than or equal to $1 - 1/s$) is a consequence of Chevet's inequality (based on Slepian's lemma) followed by Chebyshev's inequality.

In order to apply Theorem 3.2, we shall need to control probability p_0 and the expressions appearing in (3.2). This will follow from the third and the final

fact we require, namely, that if E has cotype q with a constant β_q and $T \subset S^{n-1}$ then

$$\frac{c_5}{L} \sup_{t \in T} \mathbb{E} \left\| \sum_{i=1}^{k} \langle X_i, t \rangle e_i \right\|_E \leq \mathbb{E} \left\| \sum_{i=1}^{k} g_i e_i \right\|_E$$

(3.3)

$$\leq c_4 L \sqrt{q} \beta_q \inf_{t \in T} \mathbb{E} \left\| \sum_{i=1}^{k} \langle X_i, t \rangle e_i \right\|_E,$$

for suitable absolute constants c_4, c_5. This is evident by applying Corollary 2.8 for $\xi_i = \langle X_i, t \rangle$.

Thus, all that is needed for the proof of Theorem 1.2 is to show an appropriate concentration result for an arbitrary subgaussian measure.

Let us mention once again that for the Gaussian of the Rademacher measures such a concentration result is well-known, and yields the estimate of α being an absolute constant. This will show the "moreover" part of the theorem. For the general case, consider the following

LEMMA 3.5: *There exists an absolute constant $c_6 > 0$ for which the following holds. Let $(\xi_i)_{i=1}^{k}$ be i.i.d. symmetric subgaussian random variables with $\|\xi\|_{\psi_2} \leq L$. Let $E = (\mathbb{R}^k, \| \ \|_E)$, such that for every $x \in E$, $\|x\|_E \leq |x|$. Then for every $t \geq 2$,*

$$\mathbb{P} \left(\left| \left\| \sum_{i=1}^{k} \xi_i e_i \right\|_E - \mathbb{E} \left\| \sum_{i=1}^{k} \xi_i e_i \right\|_E \right| \geq t \right) \leq 2 \exp(-c_6 t^2 / L^2 \log^2 k).$$

The idea of the proof is to truncate the variables ξ_i at the right level, that is, to split each of the ξ_i to $\xi^- = \xi 1_{\{|\xi| \leq \rho\}}$ and $\xi^+ = \xi 1_{\{|\xi| > \rho\}}$ for some fixed ρ. We will show that the truncated part $\| \sum_{i=1}^{k} \xi_i^- e_i \|_E$ is concentrated around its mean using Talagrand's convex distance inequality [Ta2] (see also [L], Corollary 4.10). The second part is bounded with high probability because the variables ξ_i decay rapidly.

Proof. Consider the truncated variable $\xi^- = \xi 1_{\{|\xi| \leq \rho\}}$ and $\xi^+ = \xi 1_{\{|\xi| > \rho\}}$, where ρ will be determined later. The first step is to prove a concentration result for $\| \sum_{i=1}^{k} \xi_i^- e_i \|_E$ around its mean. Indeed, as a function on $[-\rho, \rho]^k$, $\| \sum_{i=1}^{k} a_i e_i \|_E$ is convex. Moreover, since $\|x\|_E \leq |x|$, it is 1-Lipschitz on $([-\rho, \rho]^k, | \cdot |)$. Thus, by the convex distance inequality it follows that for every

$t > 0$

$$(3.4) \qquad \mathbb{P}\left(\left|\left\|\sum_{i=1}^{k}\xi_i^- e_i\right\|_E - \mathbb{E}\left\|\sum_{i=1}^{k}\xi_i^- e_i\right\|_E\right| \geq t\right) \leq \exp(-c_1' t^2/\rho^2)$$

for a suitable absolute constant c_1'.

To take care of the "unbounded" part we shall require some properties of Orlicz norms. First, by Theorem 2.3, if $X_1, ..., X_k$ are independent, centered random variables then

$$(3.5) \qquad \left\|\sum_{i=1}^{k}X_i\right\|_{\psi_1} \leq c_2'\left(\left\|\sum_{i=1}^{k}X_i\right\|_{L_1} + |\max|X_i|\|_{\psi_1}\right).$$

Also, for any random variable X, $\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2$, and, by Lemma 2.2, if $(X_i)_{i=1}^{k}$ are identically distributed random variables then $\|\max_{1\leq i \leq k}|X_i|\|_{\psi_1} \leq \|X\|_{\psi_1}\log k$.

Let $X_i = (\xi_i^+)^2 - \mathbb{E}(\xi_i^+)^2$. By the remarks above it is clear that $\|X_i\|_{\psi_1} \leq c_3' L^2$ and a standard calculation yields that

$$\|\sum_{i=1}^{k}X_i\|_{L_1} \leq c_3' k(\rho^2 + 2L^2)\exp(-c_4'\rho^2/L^2)$$

for suitable absolute constants c_3' and c_4'. Hence, by (3.5),

$$\left\|\sum_{i=1}^{k}(\xi_i^+)^2\right\|_{\psi_1} \leq c_5'\left(k(\rho^2 + 2L^2)\exp(-c_4'\rho^2/L^2) + L^2\log k\right)$$

for an absolute constant c_5'.

Since $\|x\|_E \leq |x|$ then $\|\sum_{i=1}^{k}\xi_i^+ e_i\|_E \leq \left(\sum_{i=1}^{k}(\xi_i^+)^2\right)^{1/2}$. Thus, setting $\rho = c_6' L\sqrt{\log k}$,

$$\left\|\sum_{i=1}^{k}\xi_i^+ e_i\right\|_{\psi_2(E)} \leq c_7' L\sqrt{\log k}, \qquad \mathbb{E}\left\|\sum_{i=1}^{k}\xi_i^+ e_i\right\|_E \leq 1.$$

Therefore, if $t \geq 2$,

$$(3.6) \qquad \mathbb{P}\left(\left\|\sum_{i=1}^{k}\xi_i^+ e_i\right\|_E + \mathbb{E}\left\|\sum_{i=1}^{k}\xi_i^+ e_i\right\|_E \geq t\right) \leq 2\exp(-c_8' t^2/L^2\log k).$$

The claim follows by combining (3.4) and (3.6).   ∎

Proof of Theorem 1.2. Let $c_1, c_2, c_3, c_4, c_5, c_6$ denote the constants appearing earlier throughout this section. Let α be the constant from Definition 3.1 (we may clearly assume that $\alpha \geq 1$), and note that by Lemma 3.5, $\alpha \leq L\sqrt{\log k}/c_6$. However, for clarity, we shall make our calculations for a general α. Let $s \geq 1$. We shall show that for every $s \geq 1$, (1.3) holds with probability $1 - \tilde{p}_0 - 2\exp(-s^2)$, where $\tilde{p}_0 = \exp(-\tilde{c}_4 \, (\mathbb{E}\|\sum_{i=1}^{k} g_i e_i\|_E)^2/\alpha^2)$ and $\tilde{c}_4 > 0$ is an absolute constant.

To this end we shall use Theorem 3.2. Fix $\varepsilon > 0$ to be determined later and assume that T and E satisfy

$$(3.7) \qquad \ell_*(T) \leq \left(\varepsilon c'/L\sqrt{q}\,\beta_q\,\alpha\right) \mathbb{E}\left\|\sum_{i=1}^{k} g_i e_i\right\|_E$$

for some constant $0 < c' < 1$. A straightforward calculation shows that if c' is sufficiently small (depending on c_1, c_4) then p_0 defined in the Theorem satisfies $p_0 \leq \tilde{p}_0$, with constant $\tilde{c}_4$ depending only on c_1, c_4 and c'. Furthermore,

$$H \leq (\varepsilon/\sqrt{q}\,\beta_q)\big((c'/\alpha) + L\sqrt{q}\,\beta_q\big)\mathbb{E}\left\|\sum_{i=1}^{k} g_i e_i\right\|_E.$$

Combining this with (3.2) and (3.3) it follows that

$$\|\Gamma t\|_E \geq \left(\frac{1}{2c_4 L\sqrt{q}\,\beta_q} - \frac{c_2 s\varepsilon}{\sqrt{q}\,\beta_q}\Big(\frac{c'}{\alpha} + L\sqrt{q}\,\beta_q\Big)\right)\mathbb{E}\left\|\sum_{i=1}^{k} g_i e_i\right\|_E.$$

Letting $\varepsilon = c''/(sL^2\sqrt{q}\,\beta_q)$ for an appropriate absolute constant $c'' > 0$, and noting that L, q and β_q are greater than 1, we conclude the left hand side of (1.3). The same choice of ε also yields the right hand side. Finally we note that with our choice of $\varepsilon > 0$ condition (3.7) amounts to (1.2).

As already mentioned, the "moreover part" follows from the fact that in this case α is an absolute constant. This completes the proof. ∎

As an example, let us consider the case when $E = \ell_2^k$ for an appropriate dimension k. Since E has cotype 2 with constant 1 our assumptions hold, and we obtain that for any isotropic, subgaussian vector X, the random operator $\Gamma = \sum_{i=1}^{k} \langle X_i, \cdot \rangle e_i$ satisfies

$$c_1\sqrt{k} \leq \|\Gamma t\| \leq c_2\sqrt{k},$$

whenever $\ell_*(T) \leq c_3 \sqrt{k}/\log k$ (and in the Gaussian or Rademacher cases, the logarithmic term can be removed). This can be viewed as a weaker, "isomorphic" version of the main result in [KM] (see also [MPT]). Of course, the proof in [KM] uses the fact that E is Euclidean in a very strong way while the main idea here is that the proof should work for an arbitrary space with cotype.

Another noteworthy point that follows from the proof is that if X is a Gaussian vector, one can obtain a version of Theorem 1.1. In this case, as was noted in Remark 3.4, the use of the majorizing measure theorem in Lemma 3.3 can be replaced by Chevet's inequality (or Slepian's lemma). However, this argument gives worse dependence on $\varepsilon > 0$, namely, the main assumption (1.1) must be satisfied with ε^2 replacing ε.

References

[DG]   V. H. de la Pena and E. Giné, *Decoupling*, Springer-Verlag, New York, 1999.

[F]   X. Fernique, *Régularité des trajectoires des fonctiones aléatoires gaussiennes*, in *Ecole d'Eté deProbabilités de St-Flour 1974*, Lecture Notes in Mathematics 480, Springer-Verlag, 1975, pp. 1–96.

[G1]   Y. Gordon, *Some inequalities for Gaussian processes and applications*, Israel Journal of Mathematics, **50** (1985), 265–289.

[G2]   Y. Gordon, *On Milman's inequality and random spaces which escape through a mesh*, in *Geometric Aspects of Funtional Analysis (1986/87)*, Lectures Notes in Mathematics, 1988, pp. 84-106.

[G3]   Y. Gordon, *A note on an observation of G. Schechtman*, in *Geometric Aspects of Funtional Analysis (1986/87)* Lectures Notes in Mathematics, in print.

[JM]   N. C. Jain and M. B. Marcus, *Continuity of sub-Gaussian processes*, in *Probability on Banach spaces*, Adv. Probab. Related Topics, 4, 1978, pp. 81-196.

[KM]   B. Klartag, and S. Mendelson, *Empirical processes and random projections*, Journal of Functional Analysis, **225** (2005), 229–245.

[L]   M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys an Monographs Vol 89, AMS, 2001.

[LTa]   M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Isoperimetry and processes, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23. Springer-Verlag, Berlin, 1991.

[MaPi] B. Maurey and G. Pisier, *Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach*, Studia Mathematica, **58** (1976), 45–90.

[MPT]   S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, *Reconstruction and subgaussian operators*, in Asymptotic Geometric Analysis, Geometric and Functional Analysis, **17** (2007), 1248–1282.

[MS]   V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces*, Lecture Notes in Mathematics, 1200, Springer 1986.

[Pi1]   G. Pisier, *Some applications of the metric entropy condition to harmonic analysis,* in *Lecture Notes in Mathematics,* 995, Springer, Berlin, 1983, pp. 123–154.

[Pi2]   G. Pisier, *Probabilistic methods in the geometry of Banach spaces,* in *Probability and analysis* (Varenna, 1985), Lecture Notes in Mathematics, 1206, Springer, Berlin, 1986, pp. 167–241.

[Pi3]   G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry,* Cambridge University Press, Cambridge, 1989.

[Pi4]   G. Pisier, *Private communication.*

[S]   G. Schechtman, *Two observations regarding embedding subsets of Euclidean spaces in normed spaces,* Advances in Mathematics **200** (2006), 125–135.

[Ta1]   M. Talagrand, *Regularity of Gaussian processes,* Acta Mathematica **159** (1987), 99–149.

[Ta2]   M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces,* Publications Mathématiques. Institut de Hautes Études Scientifiques, **81** (1995), 73–205.

[Ta3]   M. Talagrand, M. *The Generic Chaining. Upper and Lower Bound of Stochastic Processes,* Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005.

[VW]   A. W. Van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes,* Springer Series in Statistics, Springer Verlag, New York, 1996.