# Reconstruction and subgaussian operators in Asymptotic Geometric Analysis

Shahar MENDELSON[1] Alain PAJOR[2]
Nicole TOMCZAK-JAEGERMANN[3]

**Abstract:**

We present a randomized method to approximate any vector $v$ from some set $T \subset \mathbb{R}^n$. The data one is given is the set $T$, vectors $(X_i)_{i=1}^k$ of $\mathbb{R}^n$ and $k$ scalar products $(\langle X_i, v\rangle)_{i=1}^k$, where $(X_i)_{i=1}^k$ are i.i.d. isotropic subgaussian random vectors in $\mathbb{R}^n$, and $k \ll n$. We show that with high probability, any $y \in T$ for which $(\langle X_i, y\rangle)_{i=1}^k$ is close to the data vector $(\langle X_i, v\rangle)_{i=1}^k$ will be a good approximation of $v$, and that the degree of approximation is determined by a natural geometric parameter associated with the set $T$.

We also investigate a random method to identify exactly any vector which has a relatively short support using linear subgaussian measurements as above. It turns out that our analysis, when applied to $\{-1,1\}$-valued vectors with i.i.d, symmetric entries, yields new information on the geometry of faces of random $\{-1,1\}$-polytope; we show that a $k$-dimensional random $\{-1,1\}$-polytope with $n$ vertices is $m$-neighborly for very large $m \le ck/\log(c'n/k)$.

The proofs are based on new estimates on the behavior of the empirical process $\sup_{f\in F}\left|k^{-1}\sum_{i=1}^k f^2(X_i) - \mathbb{E}f^2\right|$ when $F$ is a subset of the $L_2$ sphere. The estimates are given in terms of the $\gamma_2$ functional with respect to the $\psi_2$ metric on $F$, and hold both in exponential probability and in expectation.

[0] *2000 MSC-classification:* 46B07, 41A45, 94B75, 52B05, 62G99

[1] Partially supported by an Australian Research Council Discovery grant.
[2] Partially supported by an Australian Research Council grant.
[3] This author holds the Canada Research Chair in Geometric Analysis.

# 0 Introduction

The aim of this article is to investigate the linear "approximate reconstruction" problem in $\mathbb{R}^n$. In such a problem, one is given a set $T \subset \mathbb{R}^n$ and the goal is to be able to approximate any unknown $v \in T$ using random linear measurements. In other words, one is given the set of values $(\langle X_i, v\rangle)_{i=1}^k$, where $X_1, ..., X_k$ are given independent random vectors in $\mathbb{R}^n$ selected according to some probability measure $\mu$. Using this information (and the fact that the unknown vector $v$ belongs to $T$) one has to produce, with very high probability with respect to $\mu^k$, some $t \in T$, such that the Euclidean norm $|t - v| \leq \varepsilon(k)$ for $\varepsilon(k)$ as small as possible. Of course, the random sampling method has to be "universal" in some sense and not tailored to a specific set $T$; and it is natural to expect that the degree of approximation $\varepsilon(k)$ depends on some geometric parameter associated with $T$.

Questions of a similar flavor have been thoroughly studied in approximation theory for the purpose of computing Gelfand numbers (see in particular [Ka, GG] when $T$ is the unit ball in $\ell_1^n$), in the asymptotic theory of Banach spaces for an analysis of low-codimensional sections (see [Mi, PT1]) and in the form and language presented above in nonparametric statistics and statistical learning theory in [MT] (for more information, see for example, [BBL] and [M] and references therein). This particular problem has been addressed by several authors with view of application to signal reconstruction (see [CT1, CT2, CT3] for the most recent contributions), in the following context: The sets considered were either the unit ball in $\ell_1^n$ or the unit balls in weak $\ell_p^n$ spaces for $0 < p < 1$ - and the proofs of the approximation estimates depended on the choice of those particular sets. The sampling process was done when $X_i$ were distributed according to the Gaussian measure on $\mathbb{R}^n$ or for Bernoulli and Fourier ensembles. The authors developed a unified approach via a spectral property of the underlying matrix determined by the $X_i$'s which they formalized in the concept they called the uniform uncertainty principle (UUP).

In contrast, we present a method which works in a very general geometric setting. Our results hold for *any* set $T \subset \mathbb{R}^n$, and the class of measures that could be used is broad; it contains all probability measures on $\mathbb{R}^n$ which are isotropic and subgaussian, that is, satisfy that for every $y \in \mathbb{R}^n$, $\mathbb{E}| \langle X, y\rangle |^2 = |y|^2$, and the random variables $\langle X, y\rangle$ are subgaussian with constant $\alpha|y|$ for some $\alpha \geq 1$ (see Definition 2.2, below). This class of measures contains, among others, the Gaussian measure on $\mathbb{R}^n$, the uniform measure on the vertices of the combinatorial cube (that is, the Bernoulli case) and the normalized volume measure on various convex, symmetric bodies (e.g. the unit balls of $\ell_p^n$ for $2 \leq p \leq \infty$).

A preliminary form of these results was announced in [MPT]. In particular that article outlined the argument of the Bernoulli case, as presented here. As a consequence, this solved a question left open in [CT1], that the Bernoulli ensemble satisfies the UUP in the strongest form (with constants arbitrarily close to 1, see Definition 1.2 in [CT1] and the comments afterwards).

We will show that the key parameter in the estimate on the degree of approximation $\varepsilon(k)$ is indeed geometric in nature. The analysis of the approximation problem is centered around the way the random operator $\Gamma = k^{-1/2} \sum_{i=1}^{k} \langle X_i, \cdot \rangle \, e_i$ (where $(e_i)_{i=1}^k$ is the standard basis in $\mathbb{R}^k$) acts on subsets of the unit sphere. This topic has been studied extensively in Asymptotic Geometric Analysis, particularly when $(X_i)_{i=1}^k$ are independent, standard Gaussian vectors in $\mathbb{R}^n$. Our results are based on the progress made in the last 20 years in that field, especially in the direction of the so-called low $M^*$ estimates, which were introduced by V. Milman (see, for example the books [MiSc, P] and the survey [GM] for more information on the subject). These estimates enable one to study the diameter of the intersection of the set $T$ with the kernel of random operator $\Gamma$, and this diameter plays a central role in the solution of the reconstruction problem we investigate here.

## 0.1 The Approximate Reconstruction Problem

Let us start by explaining the geometric idea which is at the heart of our analysis of the approximate reconstruction problem. Suppose that $T \subset \mathbb{R}^n$ is the set from which the unknown $v$ was selected, and consider the set of differences $T - v = \{t - v : t \in T\}$. Thus, our aim is to find a point close to 0 in that set and to accomplish that we can only use the empirical data $(\langle X_i, v \rangle)_{i=1}^k$, where $X_1, \ldots, X_k$ are selected according to the probability measure $\mu$.

The way we suggest to approximate $0 \in T - v$ is by showing that the diameter of the intersection of the kernel of $\Gamma = k^{-1/2} \sum_{i=1}^{k} \langle X_i, \cdot \rangle \, e_i$ with $T - v$ is small. This would solve the problem because one can identify points $t - v \in T - v$ which are in the kernel of $\Gamma$ using the given empirical data, and for such points the Euclidean norm $|t - v|$ is bounded by the diameter of the intersection. Of course, the obvious problem in bounding the diameter of the intersection is that the set $T - v$ is not known. However, under mild assumptions on the set $T$ (e.g. that it is quasi-convex and centrally symmetric, see Section 3), it follows that $T-v \subset aT$ for some known constant $a$. Thus, it suffices to prove that $\mathrm{diam}\,(\ker(\Gamma)\cap T)$ is small to ensure that the same holds for the set $T-v$ for any $v \in T$.

Therefore, a solution to the approximate reconstruction problem could be obtained by establishing a bound for $\mathrm{diam}\,(\ker(\Gamma)\cap T)$ which holds with high probability. This is a question that has been studied thoroughly in asymptotic geometric analysis (more details on that will follow), and the philosophy behind its solution is surprisingly simple.

Consider the set $T \subset \mathbb{R}^n$ and for every $\rho > 0$ let $T_\rho = T \cap \rho S^{n-1}$, where $S^{n-1}$ is the Euclidean sphere in $\mathbb{R}^n$. Thus, each level set $T_\rho$ divides $T$ into two sets $T_\rho^+ = \{t \in T : |t| \geq \rho\}$ and $T_\rho^- = \{t \in T : |t| < \rho\}$, that is, the parts of $T$ that are "above" and "below" the level set $T_\rho$. The aim is to find the smallest level $\rho_0$ for which $\Gamma$ acts on $T_{\rho_0}^+$ in an almost norm preserving way. In other words, to show that for every $t \in T_{\rho_0}^+$,

$$(1-\theta)|t| \leq |\Gamma t| \leq (1+\theta)|t|$$

for some fixed $0 < \theta < 1$. Thus, points in $T$ can either be mapped in a norm preserving way, or they could have a small norm. This implies that $T \cap \ker(\Gamma) \subset T_{\rho_0}^-$ which is contained in a ball of radius $\rho_0$, and thus, the intersection of the kernel of $\Gamma$ with $T$ has a diameter smaller than $\rho_0$. Actually, a norm preserving estimate yields an even stronger result, namely, that for every $t \in T$,

$$|t| \leq \max\left\{\frac{|\Gamma t|}{1-\theta}, \rho_0\right\}.$$

In particular, when applied to the set $aT$, it follows that $|t-v|$ is small as long as $k^{-1/2}\left(\sum_{i=1}^k |\langle X_i, t-v\rangle|^2\right)^{1/2}$ is small. In the context of the reconstruction problem, this observation gives us a useful method to approximate $v$, namely, to select any $t \in T$ for which $k^{-1/2}\left(\sum_{i=1}^k |\langle X_i, t-v\rangle|^2\right)^{1/2}$ is small.

The above discussion indicates that the key step in the analysis of the reconstruction problem is to find the critical level $\rho_0$ up to which $T_{\rho_0}^+$ is mapped in a $\theta$-norm preserving way with high probability. This parameter is determined by the geometry of the set $T$. To see that, let us assume that the set $T$ is star-shaped around 0, that is, for every $t \in T$ and $0 \leq \lambda \leq 1$, $\lambda t \in T$. Hence, the level sets $T_\rho$ become increasingly rich as $\rho$ decreases, up to a point at which $\Gamma$ can no longer $\theta$-preserve the norm on $T_\rho$. To identify this critical level one has to find a correct notion of complexity that captures the way $\Gamma$ acts on a subset of the sphere. We will show that correct measure of complexity is the $\ell_*$ functional, defined for any set $A$ by

$$\ell_*(A) = \mathbb{E} \sup_{t \in A} \left| \sum_{i=1}^n g_i t_i \right|,$$

where $(g_i)_{i=1}^n$ are independent standard Gaussian random variables. Roughly speaking, $\ell_*(A)$ measures the way points in $A$ are correlated with "random noise" which is modeled by a Gaussian vector, and, as we explain later, it plays a central role in asymptotic geometric analysis. Moreover, the critical level $\rho_0$ can be expressed by the $\ell_*$ functional as the following fixed point

$$r_k^*(\theta, T) := \inf\left\{\rho > 0 : \rho \geq \frac{\ell_*(T \cap \rho S^{n-1})}{\theta\sqrt{k}}\right\}, \tag{0.1}$$

where $\theta$ is determined by the degree of norm preservation one requires. These two parameters will be discussed in more details in section 2, and in particular, we will explain their geometric meaning.

Although the fixed point $\rho_0$ seems hard to compute, there are many methods of finding it. In fact, this parameter is well understood when $T$ is the unit ball of many classical spaces. Even for a general set $T$ it could be upper bounded with a mild looseness using the covering numbers of $T$. To keep this note from becoming too long, we will present explicit bounds only for the sets $T$ previously studied in the context of the reconstruction problem, namely, the unit ball in $\ell_1^n$ and in weak $\ell_p^n$ for $0 < p < 1$. To handle more general cases (sometimes, at the risk of a slightly sub-optimal bound), we refer the reader to standard texts in asymptotic geometric analysis [MiSc, LT, P].

To conclude this part of the introduction, note that we improve the known results concerning the approximate reconstruction problem in three aspects. First, the method we suggest here works for a very general family of sets $T$ (for example, it is applicable for any quasi-convex, centrally symmetric set in $\mathbb{R}^n$, see Section 3) rather than for specific sets as the previous methods of analysis yielded. Moreover, our method gives the optimal estimate for $\varepsilon(k)$ in cases previously studied - which is simply the critical level $\rho_0(T)$. Second, the family of measures we consider here is extensive - any isotropic, subgaussian measure on $\mathbb{R}^n$ would do. Finally, our method gives the correct probabilistic estimate for the success, namely, $1 - \exp(-ck)$ for some absolute constant $c$.

This article is divided into three main parts. The first one is devoted to showing that the critical level $\rho_0$ for an arbitrary set $T$ and a general subgaussian operator $\Gamma$ can be expressed using the geometric parameter defined above. The proof of this fact is the main theoretical ingredient in the solution of the approximate reconstruction problem. In fact, we establish this claim by solving a more general problem in empirical processes theory, which, in the setup we consider here, yields the bound for $\rho_0$.

In the second part we consider specific examples of sets $T$ and show how to bound $\rho_0$ in those cases. Finally, we study another notion of reconstruction, called exact reconstruction, which will be described in the next section.

## 0.2 The main results

For the moment, let us present a simple version of our main result. Let $T \subset \mathbb{R}^n$ be a star-shaped set (i.e. for every $t \in T$ and $0 \le \lambda \le 1$, $\lambda t \in T$).

**Theorem A** *There exist absolute constants $c_1, c_2 > 0$ for which the following holds. Let $T$ be a star-shaped subset of $\mathbb{R}^n$. Let $\mu$ be an isotropic, subgaussian measure with constant $\alpha$ on $\mathbb{R}^n$ and set $X_1, ..., X_k$ be independent, selected according to $\mu$. Then, with probability at least $1-\exp(-c_1 k/\alpha^4)$, every $y, v \in T$ satisfies that*

$$|y-v| \le \sqrt{2}\left(\frac{1}{k}\sum_{i=1}^{k}\left(\langle X_i, v\rangle - \langle X_i, y\rangle\right)^2\right)^{1/2} + r_k^*(c_2/\alpha^2, T-T),$$

*where $r_k^*$ was defined in* (0.1).

If $T$ is a convex, centrally symmetric set, and $v \in T$ is the point we wish to approximate, Theorem A implies that if we select $y \in T$ for which

$$\left(\frac{1}{k}\sum_{i=1}^{k}\left(\langle X_i, v\rangle - \langle X_i, y\rangle\right)^2\right)^{1/2} \le \varepsilon/\sqrt{2},$$

then

$$|y-v| \le \varepsilon + r_k^*(c_2/\alpha^2, T-T) \le \varepsilon + r_k^*(c_2/\alpha^2, 2T) \le \varepsilon + c_3\alpha^2\frac{\ell_*(T)}{\sqrt{k}},$$

where $c_3$ is an absolute constant. Observe that the last inequality is general in the sense that $r_k^*(T) \le c(\theta)\ell_*(T)/\sqrt{k}$ for any $T$. This is evident from the definition of $r_k^*$ and the fact that $\ell_*$ is monotone with respect to inclusion. In many case, even this trivial bound on the critical level is optimal.

In view of applications to concrete examples, a power of our method lies in the fact that the parameter $r_k^*(\theta, T)$ can be estimated for unit balls of classical normed or quasi-normed spaces (see Section 3). For example, if $T = B_1^n$ then with the same hypothesis and probability as above, one has

$$|y - v| \le \sqrt{2} \left( \frac{1}{k} \sum_{i=1}^{k} \left( \langle X_i, v \rangle - \langle X_i, y \rangle \right)^2 \right)^{1/2} + c\alpha^2 \sqrt{\frac{\log(c\alpha^4 n/k)}{k}}$$

where $c > 0$ is an absolute constant; this leads to the optimal estimate for $\varepsilon(k)$ for that set.

The second question we tackle here is of a similar nature, and focuses on "exact reconstruction" of vectors in $\mathbb{R}^n$ that have a short support. This question was extensively studied in recent years, most notably, by Donoho (see, e.g. [CDS, DE, D1, DET]).

Suppose that $z \in \mathbb{R}^n$ is in the unit Euclidean ball, and has a relatively short support $m \ll n$. The aim is to use a random sampling procedure to identify $z$ exactly, rather than just to approximate it. The motivation for this problem comes from *error correcting codes*, in which one has to overcome random noise that corrupts a part of a transmitted signal. The noise is modelled by adding to the transmitted vector the noise vector $z$. The assumption that the noise does not change many bits in the original signal implies that $z$ has a short support, and thus, in order to correct the code, one has to identify the noise vector $z$ exactly. Since error correcting codes are not the main focus of this article, we will not explore this topic further, but rather refer the interested reader to [MS, CT2, RV] and references therein for more information.

In the geometric context we are interested in, the problem has been studied in [CT2, RV], where it was shown that if $z$ has a short support relative to the dimension $n$ and the size of the sample $k$, and if $\Gamma$ is a $k \times n$ matrix whose entries are independent, standard Gaussian variables, then with probability at least $1 - \exp(-ck)$, the minimization problem

$$(P) \qquad \min \|v\|_{\ell_1^n} \quad \text{for} \quad \Gamma v = \Gamma z,$$

has a unique solution, which is $z$. Thus, a solution to this minimization problem will pin-point the "noise vector" $z$. The idea of using such a minimization problem to identify the noise vector was first suggested in [CDS].

Let us mention that the method of proof in [CT2, RV] is purely Gaussian in nature, as it is based on spectral and concentration properties of Gaussian matrices which are not known in other cases.

We extend this result (using a different argument) to a general random matrix whose rows are $X_1, ..., X_k$, sampled independently according to an isotropic, subgaussian measure.

**Theorem B** *Let $\Gamma$ be as above. With probability at least $1 - \exp(-c_1\, k/\alpha^4)$, any vector $z$ whose support has cardinality less than $c_2 k/\log(c_3 n/k)$ is the unique minimizer of the problem $(P)$, where $c_1, c_2, c_3$ are absolute constants.*

The ability to extend the result to a general subgaussian measure yields instant dividends as the same analysis gives new information on the geometry of $\{-1,1\}$ random polytopes. Indeed, consider the $k \times n$ matrix $\Gamma$ whose entries are independent symmetric $\{-1,1\}$ valued random variables. Thus, $\Gamma$ is a random operator selected according to the uniform measure on the combinatorial cube $\{-1,1\}^n$, which is an isotropic, subgaussian measure with constant $\alpha = 1$. The columns of $\Gamma$, denoted by $v_1, ..., v_n$ are vectors in $\{-1,1\}^k$ and let $K^+ = \mathrm{conv}(v_1, ..., v_n)$ be the convex polytope generated by $\Gamma$; $K^+$ is thus called a random $\{-1,1\}$-polytope.

A convex polytope is called $m$-neighborly if any set of less than $m$ of its vertices is the vertex set of a face (see [Z]). The following result yields the surprising fact that a random $\{-1,1\}$-polytope is $m$-neighborly for a relatively large $m$. In particular, it will have the maximal number of $r$-dimensional faces for $r \le m$.

**Theorem C** *There exist absolute constants $c_1, c_2, c_3$ for which the following holds. For $1 \le k \le n$, with probability larger than $1 - \exp(-c_1 k)$, a $k$-dimensional random $\{-1,1\}$ convex polytope with $n$ vertices is $m$-neighborly provided that*

$$m \le \frac{c_2 k}{\log(c_3 n/k)}.$$

All the results we present here, concerning both notions of reconstruction are based on understanding the extent to which the random operator $\Gamma = k^{-1/2} \sum_{i=1}^k \langle X_i, \cdot \rangle\, e_i$ acts in an almost norm preserving way on subsets of the sphere $S^{n-1}$. In other words, to find the correct notion of complexity of a set

$T \subset S^{n-1}$ which governs

$$\sup_{t\in T} \left| \frac{1}{k} \sum_{i=1}^{k} \langle X_i, t\rangle^2 - 1 \right| = \sup_{t\in T} \left| |\Gamma t|^2 - 1 \right| .$$

In our analysis we consider a more general question, in which the class of linear functionals $F = \{\langle t, \cdot\rangle : t \in T\}$ is replaced by a set $F$ of real-valued functions, defined on an arbitrary probability space $(\Omega, \mu)$ and is contained in the sphere of $L_2(\mu)$. Hence, the analogous problem is to study

$$\sup_{f\in F} \left| \frac{1}{k} \sum_{i=1}^{k} f^2(X_i) - 1 \right| ,$$

and in particular, to find the appropriate complexity measure of the class $F$ which can be used to bound this empirical process.

In this general setup, and just like in the linear case where we assumed that the functions $\langle t, \cdot\rangle$ exhibit a subgaussian behavior with a fixed constant, we assume that each $f \in F$ is subgaussian. To be more accurate, $F$ is a bounded set with respect to the $\psi_2$ norm, which is defined for a random variable $Y$ by

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

and measures the extent in which $Y$ is subgaussian [LT, VW].

It turns out that the correct notion of complexity for the empirical process above is the so-called $\gamma_2$ functional [Ta3]. To define it, consider a metric space $(T, d)$, and for a finite set $A \subset T$ denote by $|A|$ the cardinality of $A$. An *admissible sequence* of $T$ is a collection of subsets of $T$, $\{T_s : s \ge 0\}$, such that for every $s \ge 1$, $|T_s| = 2^{2^s}$ and $|T_0| = 1$. For $p = 1, 2$, we define the $\gamma_p$ functional by

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

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 $T$.

**Theorem D** *There exist absolute constants $c_1, c_2, c_3$ for which the following holds. Let $(\Omega, \mu)$ be a probability space, set $F$ be a subset of the unit sphere of*

*$L_2(\mu)$ and assume that* $\mathrm{diam}(F, \| \|_{\psi_2}) = \alpha$. *Then, for any* $\theta > 0$ *and* $k \geq 1$ *satisfying*

$$c_1 \alpha \gamma_2(F, \| \|_{\psi_2}) \leq \theta \sqrt{k},$$

*with probability at least* $1 - \exp(-c_2 \theta^2 k / \alpha^4)$,

$$\sup_{f \in F} \left| \frac{1}{k} \sum_{i=1}^{k} f^2(X_i) - \mathbb{E} f^2 \right| \leq \theta.$$

*Moreover, if $F$ is symmetric, then*

$$\mathbb{E} \sup_{f \in F} \left| \frac{1}{k} \sum_{i=1}^{k} f^2(X_i) - \mathbb{E} f^2 \right| \leq c_3 \alpha^2 \frac{\gamma_2(F, \| \|_{\psi_2})}{\sqrt{k}}.$$

As will be shown later, if $T \subset S^{n-1}$, $F = \{\langle t, \cdot \rangle : t \in T\}$ and $\mu$ is an isotropic subgaussian measure on $\mathbb{R}^n$ then $\gamma_2(F, \| \|_{\psi_2}) \sim \ell_*(T)$, which is a deep result due to Talagrand. Thus, Theorem A follows from the more general formulation of Theorem D.

Theorem D improves a result of a similar flavor from [KM] in two ways. First of all, the bound on the probability is exponential in the sample size $k$ which was not the case in [KM]. Second, we were able to bound the expectation of the supremum of the empirical process using only a $\gamma_2$ term. This fact is surprising by itself because the expectation of the supremum of empirical processes is usually controlled by two terms; the first one bounds the subgaussian part of the process and is captured by the $\gamma_2$ functional with respect to the underlying metric. The other term is needed to control subexponential part of the empirical process, and is bounded by the $\gamma_1$ functional with respect to an appropriate metric (see [Ta3] for more information on the connections between the $\gamma_p$ functionals and empirical processes). Theorem D shows that the expectation of the supremum of $|Z_f|$ behaves as if $\{Z_f : f \in F\}$ were a subgaussian process with respect to the $\psi_2$ metric (although it is not), and this is due to the key fact that all the functions in $F$ have the same second moment.

Let us mention that in the Gaussian case, the estimate on the expectation of the supremum of $|Z_f|$ follows from Gordon's min-max Theorem [Go]. Also, the way Gaussian operators act on the subsets of the sphere has been recently used by Schechtman [S] to construct embeddings of subsets of $\ell_2$ in an arbitrary normed space.

We end this introduction with the organization of the article. In Section 1 we present the proof of Theorem D and some of its corollaries we require. In Section 2 we illustrate these results in the case of linear processes which corresponds to linear measurements. In Section 3 we investigate the approximate reconstruction problem for a general set and provide sharp estimates when $T$ is the unit ball of some classical spaces. In Section 4 we present a proof for the exact reconstruction scheme, with its application to the geometry of random $\{-1,1\}$-polytopes.

Throughout this article we will use letters such as $c, c_1, ..$ to denote absolute constants which may change depending on the context. We denote by $(e_i)_{i=1}^n$ the canonical basis of $\mathbb{R}^n$, by $|x|$ the Euclidean norm of a vector $x$ and by $B_2^n$ the Euclidean unit ball. We also denote by $|I|$ the cardinality of a finite set $I$. Let $p > 0$, we denote by $\ell_p^n$ the space $\mathbb{R}^n$ equipped with the norm (or quasi-norm when $p < 1$) defined for any $t = (t_i)_{i=1}^n \in \mathbb{R}^n$ by $\|t\|_{\ell_p} = (\sum_{i=1}^n |t_i|^p)^{1/p}$.

*Acknowledgement:* A part of this work was done when the second and the third authors were visiting the Australian National University, Canberra; and when the first and the third authors were visiting Université de Marne-la-Vallée. They wish to thank these institutions for their hospitality.

# 1 Empirical Processes

In this section we present some results in empirical processes that will be central to our analysis. All the results focus on understanding the process $Z_f = \frac{1}{k}\sum_{i=1}^k f^2(X_i) - \mathbb{E}f^2$, where $k \geq 1$ and $X_1, ..., X_k$ are independent random variables distributed according to a probability measure $\mu$. In particular, we investigate the behavior of $\sup_{f\in F} |Z_f|$ in terms of various metric structures on $F$, and under the key assumption that every $f \in F$ has the same second moment. The parameters involved are standard in *generic chaining* type estimates (see [Ta3] for a comprehensive study of this topic). The philosophy and general approach of the proof come from [KM].

Recall that the $\psi_p$ norm of a random variable $X$ is defined as

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

It is standard to verify (see for example [VW]) that if $X$ has a bounded $\psi_2$ norm, then it is subgaussian with parameter $c\|X\|_{\psi_2}$ for some absolute constant $c$. More generally, a bounded $\psi_p$ norm implies that $X$ has a tail behavior, $\mathbb{P}(|X| > u)$, of the type $\exp(-cu^p/\|X\|_{\psi_p})$.

Our first fundamental ingredient is the well known Bernstein's inequality which we shall use in the form of a $\psi_1$ estimate ([VW]).

**Lemma 1.1** *Let $Y_1, ..., Y_k$ be independent random variables with zero mean such that for some $b > 0$ and every $i$, $\|Y_i\|_{\psi_1} \le b$. Then, for any $u > 0$,*

$$\mathbb{P}\Big(\Big|\frac{1}{k}\sum_{i=i}^{k} Y_i\Big| > u\Big) \le 2\exp\left(-c\,k\min\Big(\frac{u}{b}, \frac{u^2}{b^2}\Big)\right), \tag{1.1}$$

*where $c > 0$ is an absolute constant.*

We will be interested in classes of functions $F \subset L_2(\mu)$ bounded in the $\psi_2$-norm; we assume without loss of generality that $F$ is symmetric and we let $\operatorname{diam}(F, \|\ \|_{\psi_2}) := 2\sup_{f\in F}\|f\|_{\psi_2}$. Additionally, in many technical arguments we shall often consider classes $F \subset S_{L_2}$, where $S_{L_2} = \{f : \|f\|_{L_2} = 1\}$ is the unit sphere in $L_2(\mu)$.

Let $X_1, X_2, \ldots$ be independent random variables distributed according to $\mu$. Fix $k \ge 1$ and for $f \in F$ define the random variables $Z_f$ and $W_f$ by

$$Z_f = \frac{1}{k}\sum_{i=1}^{k} f^2(X_i) - \mathbb{E}f^2 \qquad \text{and} \qquad W_f = \left(\frac{1}{k}\sum_{i=1}^{k} f^2(X_i)\right)^{1/2}.$$

The first lemma follows easily from Bernstein's inequality. We state it in the form convenient for further use, and give a proof of one part, for completeness.

**Lemma 1.2** *There exists an absolute constant $c_1 > 0$ for which the following holds. Let $F \subset S_{L_2}$, $\alpha = \operatorname{diam}(F, \|\ \|_{\psi_2})$ and set $k \ge 1$. For every $f, g \in F$ and every $u \ge 2$ we have*

$$\mathbb{P}\big(W_{f-g} \ge u\|f-g\|_{\psi_2}\big) \le 2\exp(-c_1 k u^2).$$

*Also, for every $u > 0$,*

$$\mathbb{P}\left(|Z_f - Z_g| \ge u\,\alpha\,\|f-g\|_{\psi_2}\right) \le 2\exp\big(-c_1 k\min(u, u^2)\big),$$

*and*

$$\mathbb{P}\left(|Z_f| \ge u\,\alpha^2\right) \le 2\exp\big(-c_1 k\min(u, u^2)\big).$$

**Proof.** We show the standard proof of the first estimate. Other estimates are proved similarly (see e.g., [KM], Lemma 3.2).

Clearly,

$$\mathbb{E}\, W_{f-g}^2 = \frac{1}{k}\mathbb{E}\sum_{i=1}^{k}(f-g)^2(X_i) = \mathbb{E}(f-g)^2(X_1) = \|f-g\|_{L_2}^2.$$

Applying Bernstein's inequality it follows that for $t>0$,

$$\begin{aligned}&\mathbb{P}\left(\left|W_{f-g}^2 - \|f-g\|_{L_2}^2\right| \geq t\right)\\ &\leq \ 2\,\exp\left(-ck\min\left(\frac{t}{\|(f-g)^2\|_{\psi_1}}, \left(\frac{t}{\|(f-g)^2\|_{\psi_1}}\right)^2\right)\right).\end{aligned}$$

Since $\|h^2\|_{\psi_1} = \|h\|_{\psi_2}^2$ for every function $h$, then letting $t = (u^2-1)\|f-g\|_{\psi_2}^2$,

$$\begin{aligned}\mathbb{P}\Big(W_{f-g}^2 \ &\geq \ u^2\|f-g\|_{\psi_2}^2\Big)\\ &\leq \ \mathbb{P}\Big(W_{f-g}^2 - \|f-g\|_{L_2}^2 \geq (u^2-1)\|f-g\|_{\psi_2}^2\Big)\\ &\leq \ 2\,\exp\left(-ck\min(u^2/2, u^4/4)\right),\end{aligned}$$

as promised. ■

Now we return to one of the basic notions used in this paper, that of the $\gamma_2$-functional. Let $(T,d)$ be a metric space. Recall that 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 $|T_0| = 1$.

**Definition 1.3** *For a metric space $(T,d)$ and $p = 1,2$, define*

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

*where the infimum is taken with respect to all admissible sequences of $T$. In cases where the metric is clear from the context, we will denote the $\gamma_p$ functional by $\gamma_p(T)$.*

Set $\pi_s : T \to T_s$ to be a metric projection function onto $T_s$, that is, for every $t \in T$, $\pi_s(t)$ is a nearest element to $t$ in $T_s$ with respect to the metric $d$. It is easy to verify by the triangle inequality that for every admissible sequence and every $t \in T$, $\sum_{s=0}^{\infty} 2^{s/2} d(\pi_{s+1}(t), \pi_s(t)) \leq (1 + 1/\sqrt{2}) \sum_{s=0}^{\infty} 2^{s/2} d(t, T_s)$ and that $\mathrm{diam}(T, d) \leq 2\gamma_2(T, d)$.

We say that a set $F$ is *star-shaped* if the fact that $f \in F$ implies that $\lambda f \in F$ for every $0 \leq \lambda \leq 1$.

The next Theorem shows that excluding a set with exponentially small probability, $W_f$ is close to being an isometry in the $L_2(\mu)$ sense for functions in $F$ that have a relatively large norm.

**Theorem 1.4** *There exist absolute constants $c, \bar{c} > 0$ for which the following holds. Let $F \subset L_2(\mu)$ be star-shaped, $\alpha = \mathrm{diam}(F, \| \ \|_{\psi_2})$ and $k \geq 1$. For any $0 < \theta < 1$, with probability at least $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$, then for all $f \in F$ satisfying $\mathbb{E} f^2 \geq \rho_k(\theta/c\alpha)^2$, we have*

$$(1 - \theta)\mathbb{E} f^2 \leq W_f^2 \leq (1 + \theta)\mathbb{E} f^2, \tag{1.2}$$

*where*

$$\rho_k(\theta) = \rho_k(\theta, F) := \inf\Big\{\rho > 0 : \rho \geq \frac{\gamma_2(F \cap \rho S_{L_2}, \| \ \|_{\psi_2})}{\theta\sqrt{k}}\Big\}. \tag{1.3}$$

In particular, Theorem 1.4 implies the following empirical error bound, namely that for every $f \in F$,

$$\mathbb{E} f^2 \leq \max\left\{\frac{k^{-1}\sum_{i=1}^{k} f^2(X_i)}{1 - \theta}, \rho_k^2(\theta)\right\}.$$

The two-sided inequality (1.2) is intimately related to an estimate on $\sup_{f \in F} |Z_f|$, which, in turn, is based on two ingredients. The first one shows, in the language of the standard chaining approach, that one can control the "end parts" of all chains. Its proof is essentially the same as Lemma 2.3 from [KM].

**Lemma 1.5** *There exists an absolute constant $C$ for which the following holds. Let $F \subset S_{L_2}$, $\alpha = \mathrm{diam}(F, \| \ \|_{\psi_2})$ and $k \geq 1$. There is $F' \subset F$ such that $|F'| \leq 4^k$ and with probability at least $1 - \exp(-k)$, we have, for every $f \in F$,*

$$W_{f - \pi_{F'}(f)} \leq C\gamma_2(F, \| \ \|_{\psi_2})/\sqrt{k}, \tag{1.4}$$

*where $\pi_{F'}(f)$ is a nearest point to $f$ in $F'$ with respect to the $\psi_2$ metric.*

**Proof.** Let $\{F_s : s \geq 0\}$ be an "almost optimal" admissible sequence of $F$. Then for every $f \in F$,

$$\sum_{s=0}^{\infty} 2^{s/2} \|\pi_{s+1}(f) - \pi_s(f)\|_{\psi_2} \leq 2\gamma_2(F, \| \ \|_{\psi_2}).$$

Let $s_0$ be the minimal integer such that $2^{s_0} > k$, and let $F' = F_{s_0}$. Then $|F'| \leq 2^{2k} = 4^k$. Write

$$f - \pi_{s_0}(f) = \sum_{s=s_0}^{\infty} \left(\pi_{s+1}(f) - \pi_s(f)\right).$$

Since $W$ is sub-additive then

$$W_{f-\pi_{s_0}}(f) \leq \sum_{s=s_0}^{\infty} W_{\pi_{s+1}(f)-\pi_s(f)}.$$

For any $f \in F$, $s \geq s_0$ and $\xi \geq 2$, noting that $2^s > k$, it follows by Lemma 1.2 that

$$\mathbb{P}\left(W_{\pi_{s+1}(f)-\pi_s(f)} \geq \xi \frac{2^{s/2}}{\sqrt{k}} \|\pi_{s+1}(f) - \pi_s(f)\|_{\psi_2}\right) \leq 2\exp(-c_1\xi^2 2^s). \quad (1.5)$$

Since $|F_s| \leq 2^{2^s}$, there are at most $2^{2^{s+2}}$ pairs of $\pi_{s+1}(f)$ and $\pi_s(f)$. Thus, for every $s \geq s_0$, the probability of the event from (1.5) holding for some $f \in F$ is less than or equal to $2^{2^{s+2}} \cdot 2\exp(-c_1\xi^2 2^s) \leq \exp(2^{s+3} - c\xi^2 2^s)$, which, for $\xi \geq \xi_0 := \max(4/\sqrt{c_1}, 2)$, does not exceed $\exp(-c_1\xi^2 2^{s-1})$.

Combining these estimates together it follows that

$$W_{f-\pi_{s_0}(f)} \leq \xi \sum_{s=s_0}^{\infty} \frac{2^{s/2}}{\sqrt{k}} \|\pi_{s+1}(f) - \pi_s(f)\|_{\psi_2} \leq 2\xi \, \frac{\gamma_2\big(F, \| \ \|_{\psi_2}\big)}{\sqrt{k}},$$

outside a set of probability

$$\sum_{s=s_0}^{\infty} \exp(-c_1\xi^2 2^{s-1}) \leq \exp(-c_1\xi^2 2^{s_0}/4).$$

We complete the proof setting, for example $\xi = \xi_0$. ∎

**Remark 1.6** The proof of the lemma shows that there exist absolute constants $c', c'' > 0$ such that for every $\xi \geq c'$,

$$\mathbb{P}\left( \sup_{f \in F} W_{f - \pi_{F'}(f)} \geq \xi \frac{\gamma_2\big(F, \| \|_{\psi_2}\big)}{\sqrt{k}} \right) \leq \exp(-c'' \xi^2 k),$$

a fact which will be used later.

The next lemma estimates the supremum $\sup_{f \in F'} |Z_f|$, where the supremum is taken over a subset $F'$ of $F$ of a relatively small cardinality, or in other words, over the "beginning part" of a chain. In the proof of the next lemma, in order to get an estimate for probability exponential in $k$, we split the beginning part of the chain into two pieces, estimated by two slightly different arguments: we use a separate (generic chaining) argument for the "middle part" of a chain while for the "very beginning" we use a standard concentration estimate.

**Lemma 1.7** *There exist absolute constants $C$ and $c''' > 0$ for which the following holds. Let $F \subset S_{L_2}$ and $\alpha = \mathrm{diam}(F, \| \|_{\psi_2})$. Let $k \geq 1$ and $F' \subset F$ such that $|F'| \leq 4^k$. Then for every $w > 0$,*

$$\sup_{f \in F'} |Z_f| \leq C\alpha \frac{\gamma_2(F, \| \|_{\psi_2})}{\sqrt{k}} + \alpha^2 w, \tag{1.6}$$

*with probability larger than or equal to $1 - 3\exp(-c''' \, k \, \min(w, w^2))$.*

**Remark 1.8** A similar, but more general, statement appears in [Ta3], Theorem 1.2.9. Indeed, one could deduce from this theorem a weaker form of (1.6) in which the term $\alpha^2 w$ is replaced by a term of a form $\alpha^2 w \log k$.

**Proof.** Let $(F_s)_{s=0}^{\infty}$ be an almost optimal admissible sequence of $F'$, set $s_0$ to be the minimal integer such that $2^{s_0} > 2k$ and fix $s_1 \leq s_0$ to be determined later. Since $|F'| \leq 4^k$, it follows that $F_s = F'$ for every $s \geq s_0$, and that

$$Z_f - Z_{\pi_{s_1}(f)} = \sum_{s=s_1+1}^{s_0} \left( Z_{\pi_s(f)} - Z_{\pi_{s-1}(f)} \right).$$

By Lemma 1.2, for every $f \in F'$, $1 \le s \le s_0$ and $u > 0$,

$$\begin{aligned}\mathbb{P}\Big(\big|Z_{\pi_s(f)} - Z_{\pi_{s-1}(f)}\big| \ge u\alpha\sqrt{\frac{2^s}{k}}\,\|\pi_{s+1}(f) - \pi_s(f)\|_{\psi_2}\Big) \\ &\le 2\exp\big(-c_1 k \min\big((u\sqrt{2^s/k}), (u\sqrt{2^s/k})^2\big)\big) \\ &\le 2\exp(-c_1 \min(u, u^2) 2^{s-2}).\end{aligned}$$

(For the latter inequality observe that if $s \le s_0$ then $2^s/k \le 4$, and thus $\min\big((u\sqrt{2^s/k}), (u\sqrt{2^s/k})^2\big) \ge \min(u, u^2)\, 2^s/(4\,k)$.)

Taking $u$ large enough (for example, $u = 2^5/c_1$ will suffice) we may ensure that

$$\sum_{s=s_1+1}^{s_0} 2^{2^{s+2}} \exp(-c_1 u 2^{s-2}) \le \sum_{s=s_1+1}^{s_0} \exp(-2^{s+3}) \le \exp(-2^{s_1}).$$

Therefore, since there are at most $2^{2^{s+2}}$ possible pairs of $\pi_{s+1}(f)$ and $\pi_s(f)$, there is a set of probability at most $\exp(-2^{s_1})$ such that outside this set we have

$$\sup_{f \in F'} |Z_f - Z_{\pi_{s_1}(f)}| \le \frac{\alpha\, u}{\sqrt{k}} \sum_{s=s_1}^{s_0} 2^{s/2} \|\pi_{s+1}(f) - \pi_s(f)\|_{\psi_2} \le c'\alpha \frac{\gamma_2(F)}{\sqrt{k}}.$$

Denote $F_{s_1}$ by $F''$ and observe that $|F''| \le 2^{2^{s_1}}$. Thus the latter estimate implies

$$\sup_{f \in F'} |Z_f| \le c'\alpha \frac{\gamma_2(F)}{\sqrt{k}} + \sup_{g \in F''} |Z_g|.$$

Applying Lemma 1.2, for every $w > 0$ we get

$$\mathbb{P}\left(|Z_g| \ge \alpha^2 w\right) \le 2\exp(-c_1 k \min(w, w^2)).$$

Thus, given $w > 0$, choose $s_1 \le s_0$ to be the largest integer such that $2^{s_1} < c_1 k \min(w, w^2)/2$. Therefore, outside a set of probability less than or equal to $|F''|\, 2\exp(-c_1 k \min(w, w^2)) \le \exp(-c_1 k \min(w, w^2)/2)$ we have $|Z_g| \le \alpha^2 w$ for all $g \in F''$. To conclude, outside a set of probability $3\exp(-c_1 k \min(w, w^2)/2)$,

$$\sup_{f \in F'} |Z_f| \le c'\alpha \frac{\gamma_2(F)}{\sqrt{k}} + \alpha^2 w,$$

as required. ∎

**Proof of Theorem 1.4.** Fix an arbitrary $\rho > 0$, and for the purpose of this proof we let $F(\rho) = F/\rho \cap S_{L_2}$, where $F/\rho = \{f/\rho : f \in F\}$.

Our first and main aim is to estimate $\sup_{f\in F(\rho)} |Z_f|$ on a set of probability close to 1.

Fix $u, w > 0$ to be determined later. Let $F' \subset F(\rho)$ be as Lemma 1.5, with $|F'| \leq 4^k$. For every $f \in F(\rho)$ denote by $\pi(f) = \pi_{F'}(f)$ a closest point to $f$ with respect to the $\psi_2$ metric on $F(\rho)$. By writing $f = (f - \pi(f)) + \pi(f)$, it is evident that

$$|Z_f| \leq W^2_{f-\pi(f)} + 2W_{f-\pi(f)}W_{\pi(f)} + |Z_{\pi(f)}|,$$

and thus,

$$\sup_{f\in F(\rho)} |Z_f| \leq \sup_{f\in F(\rho)} W^2_{f-\pi(f)} + 2 \sup_{f\in F(\rho)} W_{f-\pi(f)} \sup_{g\in F'} W_g + \sup_{g\in F'} |Z_g|. \tag{1.7}$$

Applying Lemma 1.5, the first term in (1.7) is estimated using the fact that

$$\sup_{f\in F(\rho)} W_{f-\pi(f)} \leq C \frac{\gamma_2(F(\rho), \| \, \|_{\psi_2})}{\sqrt{k}},$$

with probability at least $1 - \exp(-k)$.

The factor $W_g$ in (1.7) is upper bounded by $\sqrt{\sup_{f\in F(\rho)} |Z_f| + 1}$ and by Lemma 1.7, $\sup_{f\in F(\rho)} |Z_f|$ is bounded by $C''\alpha\frac{\gamma_2(F(\rho),\| \, \|_{\psi_2})}{\sqrt{k}} + \alpha^2 w$ on a set of probability larger than $1 - 3e^{-ck \min(w,w^2)}$, for some numerical constants $C$" and $c > 0$.

Assuming that $\frac{\gamma_2(F(\rho),\| \, \|_{\psi_2})}{\sqrt{k}} + \alpha^2 w \leq 1$, the factor $W_g$ is bounded by $\sqrt{2}$. Combining these estimates and substituting into (1.7), $\sup_{f\in F(\rho)} |Z_f|$ is upper bounded by

$$\begin{aligned} C^2 \frac{\gamma_2(F(\rho), \| \, \|_{\psi_2})^2}{k} + 2C \frac{\gamma_2(F(\rho), \| \, \|_{\psi_2})}{\sqrt{k}} \sqrt{2} \\ + C''\alpha \frac{\gamma_2(F(\rho), \| \, \|_{\psi_2})}{\sqrt{k}} + \alpha^2 w, \end{aligned} \tag{1.8}$$

with probability at least $1 - e^{-k} - 3e^{-ck \min(w,w^2)}$.

Given $0 < \theta < 1$ we want the condition $\sup_{f\in F(\rho)} |Z_f| \leq \theta$ to be satisfied with probability close to 1. This can be achieved by imposing suitable conditions on the parameters involved. Namely, if $\rho > 0$ and $w > 0$ satisfy

$$\tilde{C}\,\alpha\,\frac{\gamma_2(F(\rho),\|\ \|_{\psi_2})}{\sqrt{k}} \le \theta/4, \qquad \alpha^2 w \le \theta/4, \tag{1.9}$$

where $\tilde{C} = \max(2\sqrt{2}\,C, C'')$, then

$$\frac{\gamma_2(F(\rho),\|\ \|_{\psi_2})}{\sqrt{k}} + \alpha^2 w \le 1$$

as required for (1.8). Moreover, each of the last three terms in (1.8) is upper bounded by $\theta/4$, and the first term is upper bounded by $(\theta/4)^2$.

To ensure that (1.9) holds, let $w = \min\big(1, \theta/(4\alpha^2)\big)$. The above discussion reveals that as long as $\rho$ satisfies

$$4\,\tilde{C}\,\alpha\,\frac{\gamma_2(F(\rho),\|\ \|_{\psi_2})}{\theta\sqrt{k}} \le 1, \tag{1.10}$$

then $\sup_{f\in F(\rho)} |Z_f| \le \theta$ on a set of measure larger than or equal to $1 - 4e^{-ck\theta^2/\alpha^4}$, where $c > 0$ is a new absolute constant. Hence, whenever $\rho$ satisfies (1.10) then (1.2) holds for all $f \in F(\rho)$. Finally, note that $\gamma_2(F(\rho),\|\ \|_{\psi_2}) = (1/\rho)\gamma_2(F \cap \rho S_{L_2},\|\ \|_{\psi_2})$, and thus (1.10) is equivalent to the inequality in the definition of $\rho_k^*(\theta)$.

To conclude the proof, for a fixed $0 < \theta < 1$ set $\rho = \rho_k(\theta/c\alpha)$, with $c = 4\tilde{C}$ being the constant from (1.10). Note that if $X_1, ..., X_k$ satisfy (1.2) for all $f \in F(\rho)$ then, since $F$ is star-shaped, the homogeneity of this condition implies that the same holds for all $f \in F$ with $\mathbb{E}f^2 \ge \rho^2$, as claimed. ■

Let us note a consequence for the supremum of the process $Z_f$, that is of independent interest.

**Corollary 1.9** *There exist absolute constants $C', c' > 0$ for which the following holds. Let $F \subset S_{L_2}$, $\alpha = \operatorname{diam}(F,\|\ \|_{\psi_2})$ and $k \ge 1$. With probability at least $1 - \exp\left(-c' \min\{k, \gamma_2^2(F,\|\ \|_{\psi_2})/\alpha^2\}\right)$ one has*

$$\sup_{f\in F} |Z_f| \le C'\,\alpha\,\max\Big(\frac{\gamma_2(F,\|\ \|_{\psi_2})}{\sqrt{k}}, \frac{\gamma_2^2(F,\|\ \|_{\psi_2})}{k}\Big).$$

*Moreover, if $F$ is symmetric,*

$$\mathbb{E}\sup_{f\in F} |Z_f| \le C'\,\alpha\,\max\Big(\frac{\gamma_2(F,\|\ \|_{\psi_2})}{\sqrt{k}}, \frac{\gamma_2^2(F,\|\ \|_{\psi_2})}{k}\Big).$$

**Proof.** This follows from the proof of Theorem 1.4 with $\rho = 1$. More precisely, the first part is a direct consequence of (1.8), by setting

$$w = \max\left\{\frac{\gamma_2(F, \| \|_{\psi_2})}{\alpha\sqrt{k}}, \frac{\gamma_2^2(F, \| \|_{\psi_2})}{\alpha^2 k}\right\}.$$

For the "moreover part" first use (1.7) for expectations, estimate the middle term by Cauchy-Schwarz inequality and note that $W_g^2 \le 1 + Z_g$ for all $g \in F'$ to yield that in order to estimate $\mathbb{E}\sup_{f\in F}|Z_f|$ it suffices to bound

$$\mathbb{E}\sup_{f\in F'}|Z_f| \quad \text{and} \quad \mathbb{E}\sup_{f\in F} W_{f-\pi(f)}^2.$$

For simplicity denote $\gamma_2(F, \| \|_{\psi_2})$ by $\gamma_2(F)$ and let us begin with the second term. Applying Remark 1.6 and setting $G = \{f - \pi(f) : f \in F\}$ and $u = c'\gamma_2(F)/\sqrt{k}$, where $c'$ is the constant from the remark, we obtain

$$\begin{aligned}\int_0^\infty \mathbb{P}\left(\sup_{g\in G} W_g^2 \ge t\right) dt &\le u^2 + \int_{u^2}^\infty \mathbb{P}\left(\sup_{g\in G} W_g^2 \ge t\right) dt \\ &\le u^2 + u^2\int_1^\infty \exp(-c''vk)dv,\end{aligned}$$

where the last inequality follows by changing the integration variable to $t = u^2 v$. This implies that

$$\mathbb{E}\sup_{f\in F} W_{f-\pi(f)}^2 \le C'\frac{\gamma_2^2(F, \| \|_{\psi_2})}{k},$$

for some absolute constant $C' > 1$.

Next, we have to bound $\mathbb{E}\sup_{f\in F'}|Z_f|$, and to that end we use Lemma 1.7. Setting $u = 2C\alpha\gamma_2(F)/\sqrt{k}$, and then changing the integration variable to $t = u/2 + \alpha^2 w$, it is evident that

$$\begin{aligned}\int_0^\infty \mathbb{P}\left(\sup_{f\in F'}|Z_f| \ge t\right) dt &\le u + \int_u^\infty \mathbb{P}\left(\sup_{f\in F'}|Z_f| \ge t\right) dt \\ &= u + \alpha^2\int_{u/2\alpha^2}^\infty \mathbb{P}\left(\sup_{f\in F'}|Z_f| \ge \frac{u}{2} + \alpha^2 w\right) dw \\ &\le u + 3\,\alpha^2\int_{u/2\alpha^2}^\infty \exp\left(-c'''k\min(w^2, w)\right) dw.\end{aligned}$$

Changing variables in the last integral $w = ru/2\alpha^2$ and using the fact that $\gamma_2(F) \geq 1$ for a symmetric set $F$ in the unit sphere of $L_2(\mu)$, the last expression is bounded above by

$$u + (3/2)u \int_1^\infty \exp\left(-c''' \min\left(\frac{r}{2\alpha^2}, \left(\frac{r}{2\alpha^2}\right)^2\right)\right) dr = C'u,$$

where $C > 0$ is an absolute constant. ∎

# 2 Subgaussian Operators

We now illustrate the general result of Section 1 in the case of linear processes, which was the topic that motivated our study. The processes correspond then to random matrices with rows distributed according to measures on $\mathbb{R}^n$ satisfying some natural geometric conditions. Our results imply concentration estimates for related random subgaussian operators, which eventually lead to the desired reconstruction results for linear measurements for general sets.

The fundamental fact that allows us to pass from the purely metric statement of the previous section to the geometric result we present below follows from Talagrand's lower bound on the expectation of the supremum of a Gaussian process in terms of $\gamma_2$ of the indexing set. To present our result, the starting point is the fundamental definition of the $\ell_*$-functional (which is in fact the so-called $\ell$-functional of a polar set).

**Definition 2.1** *Let $T \subset \mathbb{R}^n$ and let $g_1, ..., g_n$ be independent standard Gaussian random variables. Denote by $\ell_*(T) = \mathbb{E}\sup_{t\in T} |\sum_{i=1}^n g_i t_i|$, where $t = (t_i)_{i=1}^n \in \mathbb{R}^n$.*

There is a close connection between the $\ell_*$- and $\gamma_2$- functionals given by the majorizing measure Theorem. 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

$$c_2\gamma_2(T,d) \leq \mathbb{E}\sup_{t\in T} |G_t| \leq c_3\gamma_2(T,d),$$

where $c_2, c_3 > 0$ are absolute constants. The upper bound is due to Fernique [F] and the lower bound was established by Talagrand [Ta1]. The proof of both parts and the most recent survey on the topic can be found in [Ta3]

In particular, if $T \subset \mathbb{R}^n$ and $G_t = \sum g_i t_i$, then $d(s,t) = |s - t|$, and thus

$$c_2\gamma_2(T, |\ |) \leq \ell_*(T) \leq c_3\gamma_2(T, |\ |). \tag{2.1}$$

**Definition 2.2** *A probability measure $\mu$ on $\mathbb{R}^n$ is called isotropic if for every $y \in \mathbb{R}^n$, $\mathbb{E}| \langle X, y\rangle |^2 = |y|^2$, where $X$ is distributed according to $\mu$.*

*Recall that a measure $\mu$ satisfies a $\psi_2$ condition with a constant $\alpha$ if for every $y \in \mathbb{R}^n$, $\| \langle X, y\rangle \|_{\psi_2} \leq \alpha|y|$.*

*A subgaussian or $\psi_2$ operator is a random operator of the form*

$$\Gamma = \sum_{i=1}^{k} \langle X_i, \, . \rangle \, e_i \tag{2.2}$$

*where the $X_i$ are independent, distributed according to an isotropic $\psi_2$ measure.*

The reader should note that, compared with the Introduction, the present definition of $\Gamma$ does not include the normalizing factor $k^{-1/2}$.

Perhaps the most important example of an isotropic $\psi_2$ probability measure on $\mathbb{R}^n$ with a bounded constant other than the Gaussian measure is the uniform measure on $\{-1, 1\}^n$. The latter corresponds to a matrix $\Gamma$ whose entries are independent, symmetric Bernoulli variables. Naturally, if $X$ is distributed according to a general isotropic $\psi_2$ measure then the coordinates of $X$ need no longer be independent. For example, the normalized Lebesgue measure on an appropriate multiple of the unit ball in $\ell_p^n$ for $2 \leq p \leq \infty$ is an isotropic $\psi_2$ measure with a constant independent of $n$ and $p$. For more details on such measures see [MP].

For a set $T \subset \mathbb{R}^n$ and $\rho > 0$ recall that

$$T_\rho = T \cap \rho S^{n-1}. \tag{2.3}$$

The next result shows that given $T \subset \mathbb{R}^n$, subgaussian operators are very close to being an isometry on the subset of elements of $T$ which have a "large enough" norm.

**Theorem 2.3** *There exist absolute constants $c, \bar{c} > 0$ for which the following holds. Let $T \subset \mathbb{R}^n$ be a star-shaped set and put $\mu$ to be an isotropic $\psi_2$ probability measure with constant $\alpha \geq 1$. For $k \geq 1$ let $X_1, \ldots, X_k$ be independent, distributed according to $\mu$ and define $\Gamma$ by (2.2). If $0 < \theta < 1$, then with probability at least $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$, for all $x \in T$ such that $|x| \geq r_k^*(\theta/c\alpha^2)$, we have*

$$(1-\theta)|x|^2 \leq \frac{|\Gamma x|^2}{k} \leq (1+\theta)|x|^2, \tag{2.4}$$

*where*

$$r_k^*(\theta) = r_k^*(\theta, T) := \inf\Big\{\rho > 0 : \rho \geq \ell_*(T_\rho)/\big(\theta\sqrt{k}\big)\Big\}. \tag{2.5}$$

*In particular, with the same probability, every $x \in T$ satisfies*

$$|x|^2 \leq \max\Big\{(1-\theta)^{-1}|\Gamma x|^2/k,\ r_k^*(\theta/c\alpha^2)^2\Big\}. \tag{2.6}$$

**Remark 2.4** Theorem A in the Introduction follows by considering the set $T - T$ and $\theta = 1/2$.

Before proving Theorem 2.3, and to put it in perspective, one should understand the geometric meaning of the two parameters we use, and in particular, their connection to identifying sections of low codimension and small diameter of $T$. Note that when $T$ is symmetric, $\ell_*$ is, up to a multiplicative factor of the order of $\sqrt{n}$, the mean width of the body $T$, which is the average on the unit sphere of the width $2\sup_{t\in T}|\sum_{i=1}^n x_i t_i|$ in the direction of the vector $x$. The role played by this parameter in the analysis of the diameter of low-codimensional sections is essential and was discovered in [Mi], and an almost optimal quantitative result was obtained in [PT1]. Note that $\ell_*$ is the same for $T$ and its convex-hull. Thus, it does not distinguish the unit balls in weak $\ell_p^n$ spaces for $0 < p < 1$ from the one of $\ell_1^n$. On the other hand, $r_k^*$ is obtained as the Gaussian mean width $\ell_*(T)$ after rounding $T$, that is, after truncating its peaky part. The reason this truncation gives better results when trying to identify sections of low codimension of small diameter is simple. Observe that for a star-shaped body $T$ the Euclidean diameter is strictly smaller than some $\rho$ if and only if the same holds for the rounded body $T_\rho = \{x \in T\,;\, |x| \leq \rho\}$, (where with some abuse of notation, by the diameter of $T$ we mean $\max_{x\in T}|x|$). Thus, it is possible to get rid of the peaky part of $T$ at some right level $\rho$ and decrease the mean width without influencing sections of a small diameter. This is useful because now a *random* section will have a small diameter. This idea already appeared in [Gl] in the particular case of $\ell_p^n$ for $1 < p < 2$, in [GG] for $p = 1$, and in a more general setting in [MiPi] and [PT2]. A parameter of the same flavor as $r_k^*$ first appeared implicitly in [MiPi] in the context of bounding the Gelfand numbers of an operator, and was formulated geometrically in a similar way to the formulation above in [PT2]. This was done in order to get the Gaussian version of Corollary 2.6, below. In 1990, following Gordon's min-max Theorem for Gaussian processes [Go, Go1], V. Milman realized the importance of the parameter $\rho$ solving the equation $\rho = \frac{\ell_*(T \cap \rho S^{n-1})}{\theta\sqrt{k}}$. In this direction, the article [Mi1] was the first of many applications in which this parameter played a central role (see, e.g., [GMT] and references therein for other, more recent applications). In this article we do not continue in this direction, but are more interested in estimates in the spirit of [PT1] and [PT2] (see Corollary 2.6 and the comments following it).

**Proof of Theorem 2.3.** We use Theorem 1.4 for the set of functions $F$ consisting of linear functionals of the form $f = f_x = \langle \cdot, x \rangle$, for $x \in T$. By the isotropicity of $\mu$, $\|f\|_{L_2} = |x|$ for $f = f_x \in F$. Also, since $\mu$ is $\psi_2$ with constant $\alpha$ then it follows by (2.1) that for all $\rho > 0$,

$$\gamma_2(F \cap \rho S_{L_2}, \|\ \|_{\psi_2}) \le \alpha\, \gamma_2(F \cap \rho S_{L_2}, \|\ \|_{L_2}) \le (\alpha/c_2)\, \ell_*(T_\rho).$$

Thus $\rho_k(\theta, T) \le r_k^*(c_2\theta/\alpha, T)$ and the result follows from Theorem 1.4. ∎

**Remark 2.5** It is clear from the proof of Theorem 1.4 that the upper estimates in (2.4) hold for $\theta \ge 1$ as well, with appropriate probability estimates and a modified expression for $r_k^*$ in (2.5). Note that of course in this case the lower estimate in (2.4) became vacuous. The same remark is valid for the estimate in (1.2) as well.

The last result immediately leads to an estimate for the diameter of random sections of a set $T$ in $\mathbb{R}^n$, given by kernels of random operators $\Gamma$, and which is a $\psi_2$-counterpart of the main result from [PT1] (see also [PT2]).

**Corollary 2.6** *There exist absolute constants $\tilde{c}, \tilde{c}' > 0$ for which the following holds. Let $T \subset \mathbb{R}^n$ be a star-shaped set and let $\mu$ be an isotropic $\psi_2$ probability measure with constant $\alpha \ge 1$. Set $k \ge 1$, put $X_1, \ldots, X_k$ to be independent, distributed according to $\mu$ and define $\Gamma$ by (2.2). Then, with probability at least $1 - \exp(-\tilde{c}k/\alpha^4)$,*

$$\operatorname{diam}(\ker \Gamma \cap T) \le \inf\Big\{\rho > 0 : \rho \ge \tilde{c}'\, \alpha^2 \ell_*(T_\rho)/\sqrt{k}\Big\}. \tag{2.7}$$

*Moreover, with the same probability,* $\operatorname{diam}(\ker \Gamma \cap T) \le \tilde{c}'\, \alpha^2 \ell_*(T)/\sqrt{k}$.

The Gaussian case (that is, when $\mu$ is the standard Gaussian measure on $\mathbb{R}^n$), although not explicitly stated in [PT1], follows immediately from the proof in that article.

A version of estimate (2.6) in the Gaussian case was discussed in [Ta2] where it is proved, in our notation, that

$$|y - v| \leq C \left( |\Gamma y - \Gamma v|^2/k + \ell_*(T)/\sqrt{k} \right),$$

where $C$ is an absolute constant.

A version of Corollary 2.6, though with a worse dependence on the parameter $\alpha$, follows from the low $M^*$ estimates in [MP]. To obtain similar bounds to the one in Corollary 2.6 one additionally requires the fact that the random operator $\Gamma$ shrinks the diameter of the set $T$ by a factor of $\sqrt{k/n}$. As already observed in [MP], for random $\pm 1$-vectors this follows from the result in [A]. For the reconstruction problem, however, the bounds obtained by this approach do not seem to give the best dependence of $k$ on $\theta$ in Theorem 2.3 and Corollary 2.7; and this dependence of the number of measurements required to achieve the accuracy given in advance is important from a computational point of view. For instance, in Theorem 2.3 it yields the conclusion (2.4) for $|x| \geq r_k^*(\theta^2/c\alpha^2)$ instead of $|x| \geq r_k^*(\theta/c\alpha^2)$. As a consequence, in Corollary 2.7 below, one would get the weaker estimate $k \geq (c'\alpha^4/\theta^4)\ell_*(T)^2$ instead of $k \geq (c'\alpha^4/\theta^2)\ell_*(T)^2$.

**Proof of Corollary 2.6.** Applying Theorem 2.3 with $\theta = 1/2$, say, we get that if $x \in T$ and $|x| \geq r_k^*(1/(2c\alpha^2))$ then $\Gamma x \neq 0$. Thus for $x \in \ker \Gamma \cap T$ we have $|x| \leq r_k^*(1/(2c\alpha^2))$ and the first conclusion follows by adjusting the constants.

Observe that since the function $\ell_*(T_\rho)/\rho = \ell_*\big((1/\rho)T \cap S^{n-1}\big)$ is decreasing in $\rho$ then $r_k^*$ actually satisfies the equality in the defining formula (2.7). Combining this and the obvious estimate $\ell_*(T_\rho) \leq \ell_*(T)$, concludes the "moreover" part. ∎

Finally, let us note a special case of Theorem 2.3 for subsets of the sphere.

**Corollary 2.7** *Let $T \subset S^{n-1}$ and let $\mu$, $\alpha$, $k$, $X_i$, $\Gamma$ and $\theta$ be the same as in Theorem 2.3. As long as $k$ satisfies $k \geq (c'\alpha^4/\theta^2)\ell_*(T)^2$, then with probability at least $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$, for all $x \in T$,*

$$1 - \theta \leq \frac{|\Gamma x|^2}{k} \leq 1 + \theta, \tag{2.8}$$

*where $c, \bar{c} > 0$ are absolute constants.*

**Proof.** Let $c, \bar{c} > 0$ be the constants from Theorem 2.3. Observe that the condition on $k$, with $c' = c^2$, is equivalent to $r_k^*(\theta/c\alpha^2, \tilde{T}) \le 1$, where $\tilde{T} = \{\lambda x : x \in T, 0 \le \lambda \le 1\}$. Then (2.8) immediately follows from (2.4). ■

# 3 Approximate reconstruction

Next, we show how one can apply Theorem 2.3 to reconstruct any fixed $v \in T$ for any set $T \subset \mathbb{R}^n$, where the data at hand are linear subgaussian measurements of the form $\langle X_i, v \rangle$. We use below the notation of Theorem 2.3.

The reconstruction method we choose is as follows: for a fixed $\varepsilon > 0$, find some $t \in T$ such that

$$\left( \frac{1}{k} \sum_{i=1}^{k} \left( \langle X_i, v \rangle - \langle X_i, t \rangle \right)^2 \right)^{1/2} \le \varepsilon.$$

The fact that we only need to find $t$ for which $(\langle X_i, t \rangle)_{i=1}^k$ is close to $(\langle X_i, v \rangle)_{i=1}^k$ rather than equal to it, is potentially important from the practical point of view, because it may lead to is a far simpler computational problem.

Let us show why such an method can be used to solve the approximate reconstruction problem.

Consider $\bar{T} = \{\lambda(t - s) : t, s \in T,\ 0 \le \lambda \le 1\}$ and observe that by Theorem 2.3, for every $0 < \theta < 1$, with high probability, every such $t \in T$ satisfies that

$$|t - v| \le \frac{\varepsilon}{\sqrt{1-\theta}} + r_k^*(\theta/c\alpha^2, \bar{T}).$$

Hence, to bound the reconstruction error, one needs to estimate $r_k^*(\theta/c\alpha^2, \bar{T})$. Of course, if $T$ happens to be convex and symmetric then $\bar{T} \subset 2T$ which is star-shaped and thus

$$|t - v| \le \frac{\varepsilon}{\sqrt{1-\theta}} + r_k^*(\theta/c\alpha^2, 2T).$$

In a more general case, when $T$ is symmetric and quasi-convex with constant $a \ge 1$, (i.e., $T + T \subset 2aT$ and $T$ is star-shaped), then

$$|t - v| \le \frac{\varepsilon}{\sqrt{1-\theta}} + r_k^*(\theta/c\alpha^2, aT).$$

Therefore, in the quasi-convex case, the ability to approximate any point in $T$ using this kind of random sampling depends on the expectation of the supremum of a Gaussian process indexed by the intersection of $T$ and a sphere of a radius $\rho$ as a function of the radius. For a general set $T$, the reconstruction error is controlled by the behavior of the expectation of the supremum of the Gaussian process indexed by the intersection of $\bar{T}$ with spheres of radius $\rho$, and this function of $\rho$ is just the modulus of continuity of the Gaussian process indexed by the set $\{\lambda t : 0 \leq \lambda \leq 1,\ t \in T\}$ (i.e, the expectation of the supremum of the Gaussian process indexed by the set $\{\lambda(t-s) : 0 \leq \lambda \leq 1,\ t,s \in T,\ |t-s| = \rho\}$).

The parameters $r_k^*(\theta, T)$ can be estimated for the unit balls of classical normed or quasi-normed spaces. The two example we consider here are the unit ball in $\ell_1^n$, denoted by $B_1^n$, and the unit balls in the weak-$\ell_p^n$ spaces $\ell_{p,\infty}^n$ for $0 < p < 1$, denoted by $B_{p,\infty}^n$. Recall that $B_{p,\infty}^n$ is the set of all $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ such that the cardinality $|\{i : |x_i| \geq s\}| \leq s^{-p}$ for all $s > 0$, and observe that $B_{p,\infty}^n$ is a quasi convex body with constant $a = 2^{1/p}$. Let us mention that there is nothing "magical" about the examples we consider here. Those are simply the cases considered in [CT1, RV].

In order to bound $r_k^*$ for these sets we shall use the approach from [GLMP], and combine it with Theorem 2.3 to recover and extend the results from [CT1, RV].

**Theorem 3.1** *There is an absolute constant $\bar{c}$ for which the following holds. Let $1 \leq k \leq n$ and $0 < \theta < 1$, and set $\varepsilon > 0$. Let $\mu$ be an isotropic $\psi_2$ probability measure on $\mathbb{R}^n$ with constant $\alpha$, and let $X_1, \ldots, X_k$ be independent, distributed according to $\mu$. For any $0 < p < 1$, with probability at least $1 - \exp(-\bar{c}\theta^2 k/\alpha^4)$, if $v, y \in B_{p,\infty}^n$ satisfy that $(\sum_{i=1}^k \langle X_i, v - y\rangle^2 / k)^{1/2} \leq \varepsilon$, then*

$$|y - v| \leq \frac{\varepsilon}{\sqrt{1-\theta}} + 2^{1/p+1}\Big(\frac{1}{p} - 1\Big)^{-1} \left(C_{\alpha,\theta}\, \frac{\log(C_{\alpha,\theta} n/k)}{k}\right)^{1/p-1/2},$$

*where $C_{\alpha,\theta} = c\alpha^4/\theta^2$ and $c > 0$ is an absolute constant.*

*If $v, y \in B_1^n$ satisfy the same assumption then with the same probability estimate,*

$$|y - v| \leq \frac{\varepsilon}{\sqrt{1-\theta}} + \left(C_{\alpha,\theta}\, \frac{\log(C_{\alpha,\theta} n/k)}{k}\right)^{1/2}.$$

To prove Theorem 3.1, which follows by estimating the critical level $r_k^*$ for the specific sets under consideration, we require the following elementary fact.

**Lemma 3.2** *Let $0 < p < 1$ and $1 \le m \le n$. Then, for every $x \in \mathbb{R}^n$,*

$$\sup_{z \in B^n_{p,\infty} \cap \rho B^n_2} \langle x, z \rangle \le 2\rho \left( \sum_{i=1}^m x_i^{*2} \right)^{1/2},$$

*where $\rho = (1/p - 1)^{-1} m^{1/2-1/p}$ and $(x_i^*)_{i=1}^n$ is a non-increasing rearrangement of $(|x_i|)_{i=1}^n$.*

*Moreover,*

$$\sup_{z \in B^n_1 \cap \rho B^n_2} \langle x, z \rangle \le 2\rho \left( \sum_{i=1}^m x_i^{*2} \right)^{1/2},$$

*with $\rho = 1/\sqrt{m}$.*

**Proof.** We will present a proof for the case $0 < p < 1$. The case of $B^n_1$ is similar.

Recall a well known fact that for two sequences of positive numbers $a_i, b_i$ such that $a_1 \ge a_2 \ge \ldots$, the sum $\sum a_i b_{\pi(i)}$ is maximal over all permutations $\pi$ of the index set, if $b_{\pi(1)} \ge b_{\pi(2)} \ge \ldots$. It follows that, for any $\rho > 0, m \ge 1$ and $z \in B^n_{p,\infty} \cap \rho B^n_2$,

$$\begin{aligned}
\langle x, z \rangle &\le \rho \left( \sum_{i=1}^m x_i^{*2} \right)^{1/2} + \sum_{i>m} \frac{x_i^*}{i^{1/p}} \\
&\le \left( \sum_{i=1}^m x_i^{*2} \right)^{1/2} \left( \rho + \frac{1}{\sqrt{m}} \sum_{i>m} \frac{1}{i^{1/p}} \right) \\
&\le \left( \sum_{i=1}^m x_i^{*2} \right)^{1/2} \left( \rho + \left( \frac{1}{p} - 1 \right)^{-1} \frac{1}{m^{1/p-1/2}} \right).
\end{aligned}$$

By the definition of $\rho$, this completes the proof. ■

Consider the set of elements in the unit ball with "short support", defined by

$$U_m = \left\{ x \in S^{n-1} : |\{i : x_i \ne 0\}| \le m \right\}.$$

Note that Lemma 3.2 combined with a duality argument implies that for every $1 \le m \le n$ and every $I \subset \{1, ..., n\}$ with $|I| \le m$,

$$\sqrt{|I|}\, B^n_1 \cap S^{n-1} \subset 2 \operatorname{conv} U_m \cap S^{n-1}. \tag{3.1}$$

The next step is to bound the expectation of the supremum of the Gaussian process indexed by $U_m$.

**Lemma 3.3** *There exists an absolute constant $c$ such that for every $1 \leq m \leq n$,*

$$\ell_*(\operatorname{conv} U_m) \leq c\sqrt{m \log(cn/m)}.$$

**Proof.** Recall that for every $1 \leq m \leq n$, there is a set $\Lambda_m$ of cardinality at most $5^m$ such that $B_2^m \subset 2 \operatorname{conv} \Lambda_m$ (for example, a successive approximation shows that we may take as $\Lambda_m$ an $1/2$-net in $B_2^m$). Hence there is a subset of $B_2^n$ of cardinality at most $5^m \binom{n}{m}$ such that $U_m \subset 2 \operatorname{conv} \Lambda_m$. It is well known (see for example [LT]) that for every $T \subset B_2^n$,

$$\ell_*(\operatorname{conv} T) = \ell_*(T) \leq c\sqrt{\log(|T|)},$$

and thus,

$$\ell_*(\operatorname{conv} U_m) \leq c\sqrt{\log\left(5^m \binom{n}{m}\right)},$$

from which the claim follows. ■

Finally, we are ready to estimate $r_k^*(\theta, B_{p,\infty}^n)$ and $r_k^*(\theta, B_1^n)$.

**Lemma 3.4** *There exists an absolute constant $c$ such that for any $0 < p < 1$ and $1 \leq k \leq n$,*

$$r_k^*(\theta, B_{p,\infty}^n) \leq c\left(\frac{1}{p} - 1\right)^{-1} \left(\frac{\log(cn/\theta^2 k)}{\theta^2 k}\right)^{1/p-1/2}$$

*and*

$$r_k^*(\theta, B_1^n) \leq c\left(\frac{\log(cn/\theta^2 k)}{\theta^2 k}\right)^{1/2}.$$

**Proof.** Again, we present a proof for $B_{p,\infty}^n$, while the treatment of $B_1^n$ is similar and thus omitted.

Let $0 < p < 1$ and $1 \leq k \leq n$, and set $1 \leq m \leq n$ to be determined later. Clearly,

$$\left(\sum_{i=1}^m x_i^{*2}\right)^{1/2} = \sup_{y \in U_m} \langle x, y \rangle,$$

and thus, by Lemma 3.2, $\ell_*(B^n_{p,\infty} \cap \rho B^n_2) \leq 2\rho \ell_*(U_m)$, where $\rho = (1/p - 1)^{-1} m^{1/2-1/p}$. From the definition of $r^*_k(\theta)$ in Theorem 2.3, it suffices to determine $m$ (and thus $\rho$) such that

$$c\ell_*(U_m) \leq \theta\sqrt{k},$$

which by Lemma 3.3, comes to $c\sqrt{m\log(cn/m)} \leq \theta\sqrt{k}$ for some other numerical constant $c$. It is standard to verify that the last inequality holds true provided that

$$m \leq c\frac{\theta^2 k}{\log(cn/\theta^2 k)},$$

and thus

$$r^*_k(\theta, B^n_{p,\infty}) \leq c\left(\frac{1}{p} - 1\right)^{-1}\left(\frac{\log(cn/\theta^2 k)}{\theta^2 k}\right)^{1/p-1/2}.$$

■

**Proof of Theorem 3.1.** The proof follows immediately from Theorem 2.3 and Lemma 3.4. ■

# 4 Exact reconstruction

Let us consider the following problem from the error correcting code theory. A linear code is given by an $n \times (n-k)$ real matrix $A$. Thus, a vector $x \in \mathbb{R}^{n-k}$ generates the vector $Ax \in \mathbb{R}^n$. Suppose that $Ax$ is corrupted by a noise vector $z \in \mathbb{R}^n$ and the assumption we make is that $z$ is *sparse*, that is, has a short support, which we denote by supp($z$)=$\{i : z_i \neq 0\}$. The problem is to *reconstruct* $x$ from the data, which is the noisy output $y = Ax + z$.

For this purpose, consider a $k \times n$ matrix $\Gamma$ such that $\Gamma A = 0$. Thus $\Gamma z = \Gamma y$ and correcting the noise is reduced to identifying the sparse vector $z$ (rather than approximating it) from the data $\Gamma z$ - which is the problem we focus on here.

In this context, a linear programming approach called the *basis pursuit* algorithm, was recently shown to be relevant for this goal [CDS]. This method is the following minimization problem

$$(P) \qquad \min \|t\|_{\ell_1}, \quad \Gamma t = \Gamma z$$

(and recall that the $\ell_1$-norm is defined by $\|t\|_{\ell_1} = \sum_{i=1}^n |t_i|$ for any $t = (t_i)_{i=1}^n \in \mathbb{R}^n$).

For a more detailed and complete analysis of the reconstruction of sparse vectors by this basis pursuit algorithm, we refer to [CDS] and the recent papers [CT2, CT3].

In this section, we show that if $\Gamma$ is an isotropic $\psi_2$ matrix then with high probability, for any vector $z$ whose support has size less than $ck/\log(c'n/k)$ (for some absolute constants $c, c'$), the problem $(P)$ above has a unique solution that is equal to $z$. It means that such random matrices can be used to reconstruct any sparse vector, as long as the size of the support is not too large. This extends the recent result proved in [CT2] and [RV] for Gaussian matrices. One should note that in that case the main technical point in the proof (which is contained in (4.4), below) is actually a well-known result from [GG]. This observation yields a simpler proof to the main result in [CT2, RV]. The more general, subgaussian case requires more effort.

**Theorem 4.1** *There exist absolute constants $c, C$ and $\bar{c}$ for which the following holds. Let $\mu$ be an isotropic $\psi_2$ probability measure with constant $\alpha \geq 1$. For $1 \leq k \leq n$, set $X_1, \ldots, X_k$ to be independent, distributed according to $\mu$ and let $\Gamma = \sum_{i=1}^k \langle X_i, \cdot \rangle e_i$. Then with probability at least $1 - \exp(-\bar{c}\, k/\alpha^4)$, any vector $z$ satisfying*

$$|\operatorname{supp}(z)| \leq \frac{Ck}{\alpha^4 \log(cn\alpha^4/k)}$$

*is the unique minimizer of the problem*

$$(P) \qquad \min \|t\|_{\ell_1}, \quad \Gamma t = \Gamma z.$$

Let us remark that problem $(P)$ is equivalent to the following one

$$(P') \qquad \min_{t \in \mathbb{R}^{n-k}} \|y - At\|_{\ell_1}$$

where $\Gamma A = 0$. Thus we obtain the reconstruction result:

**Corollary 4.2** *Let $A$ be a $n \times (n-k)$ matrix. Set $\Gamma$ to be a $k \times n$ matrix that satisfies the conclusion of the previous Theorem with the constants $c$ and $C$, and for which $\Gamma A = 0$. For any $x \in \mathbb{R}^{n-k}$ and any $y = Ax + z$, if $|\operatorname{supp}(z)| \leq \frac{Ck}{\log(cn/k)}$, then $x$ is the unique minimizer of the problem*

$$\min_{t \in \mathbb{R}^{n-k}} \|y - At\|_{\ell_1}.$$

## 4.1 Proof of Theorem 4.1

As in [CT2], the proof consists of finding a simple condition for a fixed matrix $\Gamma$ to satisfy the conclusion of our Theorem. We then apply a result from the previous section to show that random matrices satisfy this condition.

The first step is to provide some criteria which ensure that the problem $(P)$ has a unique solution as specified in Theorem 4.1. This convex optimization problem can be represented as a linear programming problem. Indeed, let $z \in \mathbb{R}^n$ and set

$$I^+ = \{i : z_i > 0\}, \quad I^- = \{i : z_i < 0\}, \quad I = I^+ \cup I^-. \tag{4.1}$$

Denote by $\mathcal{C}$ the cone of constraint

$$\mathcal{C} = \{t \in \mathbb{R}^n : \sum_{i \in I^+} t_i - \sum_{i \in I^-} t_i + \sum_{i \in I^c} |t_i| \leq 0\}$$

corresponding to the $\ell_1$-norm.

Note that if $|t|$ is small enough then $\|z+t\|_{\ell_1} = \sum_{i \in I^+}(z_i+t_i) - \sum_{i \in I^-}(z_i + t_i) + \sum_{i \in I^c} |t_i|$. Thus, the solution of $(P)$ is unique and equals to $z$ if and only if

$$\ker \Gamma \cap \mathcal{C} = \{0\} \tag{4.2}$$

By the Hahn-Banach separation Theorem, the latter is equivalent to the existence of a linear form $\tilde{w} \in \mathbb{R}^n$ vanishing on $\ker \Gamma$ and positive on $\mathcal{C} \setminus \{0\}$.

After appropriate normalization, it is easy to check that such an $\tilde{w}$ satisfies that $\tilde{w} = \sum_{i=1}^k \alpha_i X_i$, $\langle \tilde{w}, e_i \rangle = 1$ for all $i \in I^+$, $\langle \tilde{w}, e_i \rangle = -1$ for all $i \in I^-$, and $|\langle \tilde{w}, e_i \rangle| < 1$ for all $i \in I^c$. Setting $w = \sum_{i=1}^k \alpha_i e_i$ and noticing that $\langle \tilde{w}, e_i \rangle = \langle w, \Gamma e_i \rangle$ we arrive at the following criterion.

**Lemma 4.3** *Let $\Gamma$ be a $k \times n$ matrix and $z \in \mathbb{R}^n$. With the notation (4.1), the problem*

$$(P) \qquad \min \|t\|_{\ell_1}, \quad \Gamma t = \Gamma z$$

*has a unique solution which equals to $z$, if and only if there exists $w \in \mathbb{R}^k$ such that*

$$\forall i \in I^+ \ \langle w, \Gamma e_i \rangle = 1, \quad \forall i \in I^- \ \langle w, \Gamma e_i \rangle = -1, \quad \forall i \in I^c \ |\langle w, \Gamma e_i \rangle| < 1.$$

The second preliminary result we require follows from Corollary 2.7 and the estimates of the previous section.

**Theorem 4.4** *There exist absolute constants $c, C$ and $\bar{c}$ for which the following holds. Let $\mu$, $\alpha$, $k$ and $\Gamma$ be as in Theorem 4.1. Then, for every $0 < \theta < 1$, with probability at least $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$, every $x \in 2\,\mathrm{conv}\,U_{4m} \cap S^{n-1}$ satisfies that*

$$(1-\theta)|x|^2 \le \frac{|\Gamma x|^2}{k} \le (1+\theta)|x|^2, \tag{4.3}$$

*provided that*

$$m \le C \frac{\theta^2 k/\alpha^4}{\log\left(cn\alpha^4/\theta^2 k\right)}.$$

**Proof.** Applying Corollary 2.7 to $T = 2\,\mathrm{conv}\,U_{4m} \cap S^{n-1}$, we only have to check that $k \ge (c'\,\alpha^4/\theta^2)\ell_*(T)^2$, which from Lemma 3.3 reduces to verifying that $k \ge (c'\alpha^4/\theta^2)cm\log(cn/m)$. The conclusion now follows from the same computation as in the proof of Lemma 3.4.

**Proof of Theorem 4.1.** Observe that if $t \in \mathcal{C} \cap S^{n-1}$ then $\|t\|_{\ell_1} \le 2\sum_{i\in I}|t_i| \le 2\sqrt{|I|}$, where $I$ is the support of $z$. Hence,

$$\mathcal{C} \cap S^{n-1} \subset \sqrt{4|I|}\, B_1^n \cap S^{n-1}.$$

This inclusion and condition (4.2) clearly imply that if

$$\ker\Gamma \cap \left(\sqrt{4|I|}\, B_1^n \cap S^{n-1}\right) = \emptyset. \tag{4.4}$$

then the solution of $(P)$ is unique and equals to $z$. Condition (4.4) exactly means that the Euclidean diameter of $\ker\Gamma \cap \sqrt{4|I|}\, B_1^n$ is less than 1. When $\Gamma$ is a Gaussian matrix, the main result from [GG] tells us that this is satisfied as far as $|I| \le Ck/\log(cn/k)$ for some numerical constants $c, C > 0$. This concludes the proof of Theorem 4.1 in that case. To treat the more general case we are studying here, note that by (3.1) we have

$$\sqrt{4|I|}\, B_1^n \cap S^{n-1} \subset 2\,\mathrm{conv}\,U_{4m} \cap S^{n-1}.$$

Therefore, if $\Gamma$ does not vanish on any point of $2\,\mathrm{conv}\,U_{4m} \cap S^{n-1}$ then $z$ is the unique solution of $(P)$. Applying Theorem 4.4, the lower bound in (4.3) shows that indeed, $\Gamma$ does not vanish on any point of the required set, provided that

$$m \le \frac{Ck}{\alpha^4 \log(cn\alpha^4/k)}$$

for some suitable constants $c$ and $C$. ∎

## 4.2 The geometry of faces of random polytopes

Next, we investigate the geometry of random polytopes. Let $\Gamma$ be a $k \times n$ isotropic $\psi_2$ matrix. For $1 \le i \le n$ let $v_i = \Gamma(e_i)$ be the vector columns of the matrix $\Gamma$ and set $K^+(\Gamma)$ (resp. $K(\Gamma)$) to be the convex hull (resp., the symmetric convex hull) of these vectors.

In this situation, the random model that makes sense is when $X = (x_i)_{i=1}^n$, where $(x_i)_{i=1}^n$ are independent, identically distributed random variables for which $\mathbb{E}|x_i|^2 = 1$ and $\|x_i\|_{\psi_2} \le \alpha$. It is standard to verify that in this case $X = (x_i)_{i=1}^n$ is an isotropic $\psi_2$ vector with constant $\alpha$, and moreover, each vertex of the polytope is given by $v_i = (x_{i,j})_{j=1}^k$.

A polytope is called *$m$-neighborly* if any set of less than $m$ vertices is the vertex set of a face. In the symmetric setting, we will say that a symmetric polytope is $m$-symmetric-neighborly if any set of less than $m$ vertices containing no-opposite pairs, is the vertex set of a face.

The condition of Lemma 4.3 may be reformulated by saying that the set $\{v_i : i \in I^+\} \cup \{-v_i : i \in I^-\}$ is the vertex set of a face of the polytope $K(\Gamma)$. Thus, the condition for the exact reconstruction using the basis pursuit method for any vector $z$ with $|\operatorname{supp}(z)| \le m$ may be reformulated as a geometric property of the polytope $K(\Gamma)$ (see [CT2, RV]); namely, that *for all disjoint subsets $I^+$ and $I^-$ of $\{1, \dots, n\}$ such that $|I^+| + |I^-| \le m$, the set $\{v_i : i \in I^+\} \cup \{-v_i : i \in I^-\}$ is the vertex set of a face of the polytope $K(\Gamma)$.* That is, $K(\Gamma)$ is $m$-symmetric-neighborly. A similar analysis may be done in the non-symmetric case, for $K^+(\Gamma)$, where now $I^-$ is empty.

**Lemma 4.5** *Let $\Gamma$, $K(\Gamma)$ and $K^+(\Gamma)$ be as above. Then the problem*

$$(P) \qquad \min \|t\|_{\ell_1}, \quad \Gamma t = \Gamma z$$

*has a unique solution which equals to $z$ for any vector $z$ (resp., $z \ge 0$) such that $|\operatorname{supp}(z)| \le m$, if and only if $K(\Gamma)$ (resp., $K^+(\Gamma)$) is $m$-symmetric-neighborly (resp., $m$-neighborly).*

Applying Theorem 4.1, we obtain

**Theorem 4.6** *There exist absolute constants $c, C$ and $\bar{c}$ for which the following holds. Let $\mu$ be an isotropic $\psi_2$ probability measure with constant $\alpha \ge 1$ and let $k$ and $\Gamma$ be as above. Then, with probability at least $1 - \exp(-\bar{c}\, k/\alpha^4)$,* *the polytopes $K^+(\Gamma)$ and $K(\Gamma)$ are $m$-neighborly and $m$-symmetric-neighborly, respectively, for every $m$ satisfying*

$$m \leq \frac{Ck}{\alpha^4 \log(cn\alpha^4/k)}.$$

The statement of Theorem 4.6 for $K(\Gamma)$ and for a Gaussian matrix $\Gamma$ is the main result of [RV]. However, our analysis shows that the same results holds for a random $\{-1, 1\}$-matrix. In such a case, $K^+(\Gamma)$ is the convex hull of $n$ random vertices of the discrete cube $\{-1, +1\}^k$, also known as a random $\{-1, 1\}$-polytope. With high probability, every $(m-1)$-dimensional face of $K^+(\Gamma)$ is a simplex and there are $\binom{n}{m}$ such faces, for $m \leq Ck/\log(cn/k)$.

**Remark 4.7** Let us mention some related results about random $\{-1, 1\}$-polytopes. A result of [BP] states that for such polytopes, the number of facets, which are the $k-1$-dimensional faces, may be super-exponential in the dimension $k$, for an appropriate choice of the number $n$ of vertices. Denote by $f_q(K^+(\Gamma))$ the number of $q$-dimensional faces of the polytope $K^+(\Gamma)$. The quantitative estimate in [BP] was recently improved in [GGM] where it is shown that there are positive constants $a, b$ such that for $k^a \leq n \leq \exp(bk)$, one has $\mathbb{E} f_{k-1}(K^+(\Gamma)) \geq (\ln n/\ln k^a)^{k/2}$. For lower dimensional faces a threshold of $f_q(K^+(\Gamma))/\binom{n}{q+1}$ was established in [K].

# References


[A] Artstein, S. The change in the diameter of a convex body under a random sign-projection, in GAFA Seminar Notes 2002-2003 (V.D. Milman and G. Schechtman eds.), Lecture notes in Mathematics 1850, (2004), 31-39.

[BBL] Boucheron, S., Bousquet, O. & Lugosi, G. Theory of Classification: a Survey of Recent Advances, to appear,

[BP] Bárány I. & Pór A. On 0-1 polytopes with many facets, Adv. Math., 161 (2001), 209-228.

[CDS] Chen S.S., Donoho D. L. & Saunders M. A. Atomic decomposition by basis pursuit. SIAM J. Scientific Computing 20 (1999), 33-61.

[CT1] Candes, E. & Tao, T. Near optimal recovery from random projections: universal encoding strategies. Preprint.

[CT2] Candes, E. & Tao, T. Decoding by Linear Programming. Preprint.

[CT3] Candes, E. & Tao, T. Error Correction via Linear Programming. Preprint.

[D1] Donoho D. L. For most large undetermind systems of linear equations the minimial $\ell_1$-norm solution is the sparsets solution, preprint.

[DE] Donoho D. L. & Elad, M. Optimally sparse representation in general (non-orthogonal) dictionaries via $\ell_1$ minimization. Proc. Natl. Acad. Sci. USA 100(2003) 2197-2202.

[DET] Donoho D. L., Elad, M & Temlyakov, V. Stable recovery of sparse overcomplete representations in the presence of noise, preprint.

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

[Gl] Gluskin, E. D. Norms of random matrices and width of finite dimensional sets. Math. USSR. Sbornik, Vol. 48 (1984), 173-182.

[GG] Garnaev, A. U. & Gluskin, E. D. The widths of a Euclidean ball (Russian) doklady acad. Nauk SSSR 277 (1984) 1048–1052, english translation: soviet Math. Doklady 30 (1984), 200–204.

[GGM] Gatzouras D., Giannopoulos A. & Markoulakis N. On the maximal number of facets of 0/1 polytopes. Preprint.

[GM] Giannopoulos A.& Milman, V. Euclidean structures in finite dimensional normed spaces, Handbook of the Geometry of Banach spaces, Vol 1, (2001), 707-779, J. Lindenstrauss, W.B. Johnson (Eds.).

[GMT] Giannopoulos A., Milman, V. & Tsolomitis, A. Asymptotic formulas for the diameter of sections of symmetric convex bodies. J. of Funct. Analysis, 223 (2005), 86–108.

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

[Go1] Gordon, Y. On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$. Geometric aspects of functional analysis (1986/87), Lecture Notes in Math., 1317 (1988), 84–106.

[GLMP] Gordon, Y., Litvak, A., Mendelson, S. & Pajor, A. private communication.

[K] Kaibel V. Low-Dimensional Faces of Random 0/1-Polytopes. Integer Programming and Combinatorial Optimization. Proceedings of IPCO X, New York , D. Bienstock and G. Nemhauser (eds.), Lecture Notes in Computer Science 3064, Springer, 2004, 401-415

[Ka] Kashin, B. The widths of certain finite-dimensional sets and classes of smooth functions. Izv. Akad. Nauk SSSR Ser. Mat., 41 (1977), 334–351 (Russian).

[KM] Klartag, B. & Mendelson, S. Empirical processes and random projections, Journal of Functional Analysis, to appear.

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

[MS] MacWilliams, F.J. & Sloane, N.J.A. The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1977.

[M] Mendelson, S. Geometric parameters in Learning Theory, in GAFA seminar notes 2002-2003 (V.D. Milman and G. Schechtman eds.), Lecture notes in Mathematics 1850, (2004), 193-235.

[MPT] Mendelson, S., Pajor A. & Tomczak-Jaegermann N. *Reconstruction and subgaussian processes*, C. R. Acad. Sci. Paris, Sér. I Math., 340 (2005), 885-888.

[MT] Mendelson, S. & Tishby, N. Statistical Sufficiency for classes in empirical $L_2$ spaces, in Proceedings of the 13th annual conference on Computational Learning Theory COLT00 (Nicolò Cesa-Bianchi, Sally A. Goldman eds.), (2000) 81-89.

[Mi] Milman, V. Random subspaces of proportional dimension of finite dimensional normed spaces: approach through the isoperimetric inequality. Lecture Notes in Math., 1166 (1985), 106–115.

[Mi1] Milman, V. Spectrum of a position of a convex body and linear duality relations, Israel Mathematical Conference Proceedings Vol 3, (1990), 151-161.

[MP] Milman, V. D. & Pajor, A. Regularization of star bodies by random hyperplane cut off. *Studia Math.*, 159 (2003), 247–261.

[MiPi] Milman, V. D. & Pisier, G. Banach spaces with a weak cotype 2 property, Israel J. Math 54(2), (1986), 139–158.

[MiSc] Milman, V. D. & Schechtman, G. Asymptotic theory of finite dimensional normed spaces, Lecture Notes in Mathematics 1200 (1986), Springer.

[PT1] Pajor, A. & Tomczak-Jaegermann, N. Subspaces of small codimension of finite-dimensional Banach spaces, Proceedings of the American Mathematical Society 97(4), (1986), 637-642.

[PT2] Pajor, A. & Tomczak-Jaegermann, N. Nombres de Gelfand et sections euclidiennes de grande dimension. (French) [Gelfand numbers and high-dimensional Euclidean sections] Séminaire d'Analyse Fonctionelle 1984/1985, Publ. Math. Univ. Paris VII, 26, Univ. Paris VII, Paris, 1986, 37–47.

[P] Pisier, G. The volume of convex bodies and Banach space geometry, (1989), Cambridge University Press.

[RV] Rudelson, M. & Vershynin, R. Geometric approach to error correcting codes and reconstruction of signals. Preprint.

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

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

[Ta2] Talagrand, M. Sudakov type minoration for Gaussian chaos processes, Israel J. Math. 79 (1992), 207–224.

[Ta3] Talagrand, M. The generic chaining, Springer, 2005.

[VW] Van der Vaart, A. W. & Wellner, J. A. Weak convergence and empirical processes, Springer Verlag, 1996.

[Z] Ziegler, G. M. Lectures on Polytopes. Volume 152 of Graduate Texts in Mathematics. Springer-Verlag, New-York (1995) 2nd edition: 1998.

**S. Mendelson** Centre for Mathematics and its Applications, The Australian National University, Canberra, ACT 0200, Australia
shahar.mendelson@anu.edu.au

**A. Pajor** Laboratoire d'Analyse et Mathématiques Appliquées, Université de Marne-la-Vallée, 5 boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallee, Cedex 2, France
alain.pajor@univ-mlv.fr

**N. Tomczak-Jaegermann** Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1
nicole@ellpspace.math.ualberta.ca