Minimax rate of convergence and the performance of ERM in phase recovery

Guillaume Lecué[1,3]   Shahar Mendelson[2,4,5]

November 20, 2013

Abstract

We study the performance of Empirical Risk Minimization in noisy phase retrieval problems, indexed by subsets of $\mathbb{R}^n$ and relative to subgaussian sampling; that is, when the given data is $y_i = \langle a_i, x_0 \rangle^2 + w_i$ for a subgaussian random vector a, independent noise w and a fixed but unknown x_0 that belongs to a given subset of $\mathbb{R}^n$.

We show that ERM produces $\hat{x}$ whose Euclidean distance to either x_0 or $-x_0$ depends on the gaussian mean-width of the indexing set and on the signal-to-noise ratio of the problem. The bound coincides with the one for linear regression when $\|x_0\|_2$ is of the order of a constant. In addition, we obtain a minimax lower bound for the problem and identify sets for which ERM is a minimax procedure. As examples, we study the class of d-sparse vectors in $\mathbb{R}^n$ and the unit ball in ℓ_1^n.

1   Introduction

Phase retrieval has attracted much attention recently, as it has natural applications in areas that include X-ray crystallography, transmission electron microscopy and coherent diffractive imaging (see, for example, the discussion in [2] and references therein).

In phase retrieval, one attempts to identify a vector x_0 that belongs to an arbitrary set $T \subset \mathbb{R}^n$ using noisy, quadratic measurements of x_0. The

---

[1]CNRS, CMAP, Ecole Polytechnique, 91120 Palaiseau, France.
[2]Department of Mathematics, Technion, I.I.T, Haifa 32000, Israel.
[3]Email: guillaume.lecue@cmap.polytechnique.fr
[4]Email: shahar@tx.technion.ac.il

given data is a sample of cardinality N, $(a_i, y_i)_{i=1}^N$, for vectors $a_i \in \mathbb{R}^n$ and

$$y_i = |\langle a_i, x_0 \rangle|^2 + w_i, \tag{1.1}$$

for a noise vector $(w_i)_{i=1}^N$.

Our aim is to investigate phase retrieval from a theoretical point of view, relative to a well behaved, random sampling method. To formulate the problem explicitly, let μ be an isotropic, L-subgaussian measure on $\mathbb{R}^n$ and set a to be a random vector distributed according to μ. Thus, for every $x \in \mathbb{R}^n$, $\mathbb{E}\langle x, a \rangle^2 = \|x\|_2^2$ (isotropicity) and for every $u \geq 1$, $Pr(|\langle x, a \rangle| \geq Lu\|\langle x, a \rangle\|_{L_2}) \leq 2\exp(-u^2/2)$ (L-subgaussian).

Given a set $T \subset \mathbb{R}^n$ and a fixed, but unknown $x_0 \in T$, y_i are the random noisy measurements of x_0: for a sample size N, $(a_i)_{i=1}^N$ are independent copies of a and $(w_i)_{i=1}^N$ are independent copies of a mean-zero variable w that are also independent of $(a_i)_{i=1}^N$.

Clearly, due to the nature of the given measurements, x_0 and $-x_0$ are indistinguishable, and the best that one can hope for is a procedure that produces $\hat{x} \in T$ that is close to one of the two points.

The goal here is to find such a procedure and identify the way in which the distance between $\hat{x}$ and either x_0 or $-x_0$ depends on the structure of T, the measure μ and the noise.

The procedure studied here is empirical risk minimization (ERM), which produces $\hat{x}$ that minimizes the empirical risk in T:

$$P_N \ell_x = \frac{1}{N} \sum_{i=1}^N \left(\langle a_i, x \rangle^2 - y_i \right)^2.$$

The loss is the standard squared loss functional, which, in this case, satisfies

$$\ell_x(a, y) = (f_x(a) - y)^2 = (\langle x, a \rangle^2 - \langle x_0, a \rangle^2 - w)^2 = (\langle x - x_0, a \rangle \langle x + x_0, a \rangle - w)^2.$$

Comparing the empirical and actual structures on T is a vital component in the analysis of ERM. In phase recovery, the centered empirical process that is at the heart of this approach is defined for any $x \in T$ by,

$$P_N(\ell_x - \ell_{x_0}) = \frac{1}{N} \sum_{i=1}^N \langle x - x_0, a_i \rangle^2 \langle x + x_0, a_i \rangle^2 - \frac{2}{N} \sum_{i=1}^N w_i \langle x - x_0, a_i \rangle \langle x + x_0, a_i \rangle.$$

Both the first and second components are difficult to handle directly, even when the underlying measure is subgaussian, because of the powers involved (an effective power of 4 in the first component and of 3 in the second one).

2

Therefore, rather than trying to employ the concentration of empirical means around the actual ones, which might not be sufficiently strong in this case, one uses a combination of a small-ball estimate for the 'high order' part of the process, and a more standard deviation argument for the low-order component (see Section 3 and the formulation of Theorem A and Theorem B).

We assume that linear forms satisfy a certain small-ball estimate, and in particular, do not assign too much weight to small neighbourhoods of 0.

Assumption 1.1 *There is a constant $\kappa_0 > 0$ satisfying that for every $s, t \in \mathbb{R}^n$,*
$$\mathbb{E}|\langle a, s\rangle\langle a, t\rangle| \geq \kappa_0 \|s\|_2 \|t\|_2.$$

Assumption 1.1 is not very restrictive and holds for many natural choices of random vectors in $\mathbb{R}^n$, like the gaussian measure or any isotropic log-concave measure on $\mathbb{R}^n$ (see, for example, the discussion in [2]).

It is not surprising that the error rate of ERM depends on the structure of T, and because of the subgaussian nature of the random measurement vector a, the natural parameter that captures the complexity of T is the gaussian mean-width associated with normalizations of T.

Definition 1.1 *Let $G = (g_1, ..., g_n)$ be the standard gaussian vector in $\mathbb{R}^n$. For $T \subset \mathbb{R}^n$, set*
$$\ell(T) = \mathbb{E} \sup_{t \in T} \Big| \sum_{i=1}^n g_i t_i \Big|.$$

The normalized sets in question are
$$T_{-,R} = \left\{ \frac{t - s}{\|t - s\|_2} \ : \ t, s \in T, \quad R < \|t - s\|_2 \|t + s\|_2 \right\},$$
$$T_{+,R} = \left\{ \frac{t + s}{\|t + s\|_2} \ : \ t, s \in T, \quad R < \|t - s\|_2 \|t + s\|_2 \right\},$$

which have been used in [2], or their 'local' versions,
$$T_{-,R}(x_0) = \left\{ \frac{t - x_0}{\|t - x_0\|_2} \ : \ t \in T, \quad R < \|t - x_0\|_2 \|t + x_0\|_2 \right\},$$
$$T_{+,R}(x_0) = \left\{ \frac{t + x_0}{\|t + x_0\|_2} \ : \ t \in T, \quad R < \|t - x_0\|_2 \|t + x_0\|_2 \right\}.$$

The sets in question play a central role in the exclusion argument that is used in the analysis of ERM. Setting $\mathcal{L}_x = \ell_x - \ell_{x_0}$, the excess loss function

associated with ℓ and $x \in T$, it is evident that $P_N \mathcal{L}_{\hat{x}} \leq 0$ (because $\mathcal{L}_{x_0} = 0$ is a possible competitor). If one can find an event of large probability and $R > 0$ for which $P_N \mathcal{L}_x > 0$ if $\|x - x_0\|_2 \|x + x_0\|_2 \geq R$, then on that event, $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq R$, which is the estimate one is looking for.

The normalization allows one to study 'relative fluctuations' of $P_N \mathcal{L}_x$ – in particular, the way these fluctuations scale with $\|x - x_0\|_2 \|x + x_0\|_2$. This is achieved by considering empirical means of products of functions $\langle u, \cdot \rangle \langle v, \cdot \rangle$, for $u \in T_{+,R}(x_0)$ and $v \in T_{-,R}(x_0)$.

The obvious problem with the 'local' sets $T_{+,R}(x_0)$ and $T_{-,R}(x_0)$ is that x_0 is not known. As a first attempt of bypassing this problem, one may use the 'global' sets $T_{+,R}$ and $T_{-,R}$ instead, as had been done in [2].

Unfortunately, this global approach is not completely satisfactory. Roughly put, there are two types of subsets of $\mathbb{R}^n$ one is interested in, and that appear in applications. The first type consists of sets for which the 'local complexity' is essentially the same everywhere, and the sets $T_{+,R}, T_{-,R}$ are not very different from the seemingly smaller $T_{+,R}(x_0)$, $T_{-,R}(x_0)$, regardless of x_0.

A typical example of such a set is d-sparse vectors – a set consisting of all the vectors in $\mathbb{R}^n$ that are supported on at most d-coordinates. For every $x_0 \in T$ and $R > 0$, the sets $T_{+,R}(x_0), T_{-,R}(x_0)$, and $T_{+,R}, T_{-,R}$ are contained in the subset of the sphere consisting of $2d$-sparse vectors, which is a relatively small set.

For this kind of set, the 'global' approach, using $T_{+,R}$ and $T_{-,R}$, suffices, and the choice of the target x_0 does not really influence the rate in which $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2$ decays to 0 with N.

In contrast, sets of the second type one would like to study, have vastly changing local complexity, with the typical example being a convex, centrally symmetric set (i.e. if $x \in T$ then $-x \in T$).

Consider, for example, the case $T = B_1^n$, the unit ball in ℓ_1^n. It is not surprising that for small R, the sets $T_{+,R}(0)$ and $T_{-,R}(0)$ are very different from $T_{-,R}(e_1)$ and $T_{+,R}(e_1)$: the ones associated with the centre 0 are the entire sphere, while for $e_1 = (1, 0, ..., 0)$, $T_{+,R}(e_1)$ and $T_{-,R}(e_1)$ consist of vectors that are well approximated by sparse vectors (whose support depend on R), and thus are rather small subsets of the sphere .

The situation that one encounters in B_1^n is generic for convex centrally-symmetric sets. The sets become locally 'richer' the closer the centre is to 0, and at 0, for small enough R, $T_{+,R}(0)$ and $T_{-,R}(0)$ are the entire sphere. Since the sets $T_{+,R}$ and $T_{-,R}$ are blind to the location of the centre, and are, in fact, the union over all possible centres of the local sets, they are simply too big to be used in the analysis of ERM in convex sets. A correct estimate on the performance of ERM for such sets requires a more delicate

4

local analysis and additional information on $\|x_0\|_2$. Moreover, the rate of convergence of ERM *truly depends* on $\|x_0\|_2$ in the phase recovery problem via the signal-to-noise ratio $\|x_0\|_2 / \sigma$.

We begin by formulating our results using the 'global' sets $T_{+,R}$ and $T_{-,R}$. Let $T_+ = T_{+,0}$ and $T_- = T_{-,0}$, set

$$E_R = \max\{\ell(T_{+,R}), \ell(T_{-,R})\}, \quad E = \max\{\ell(T_+), \ell(T_-)\}$$

and observe that as nonempty subsets of the sphere $\ell(T_{-,r}), \ell(T_{+,r}) \geq \mathbb{E}|g| = \sqrt{2/\pi}$.

The first result presented here is that the error rate of ERM for the phase retrieval problem in T depends on the fixed points

$$r_2(\gamma) = \inf\left\{r > 0 : E_r \leq \gamma\sqrt{N}r\right\}$$

and

$$r_0(Q) = \inf\left\{r > 0 : E_r \leq Q\sqrt{N}\right\},$$

for constants γ and Q that will be specified later.

Recall that the ψ_2-norm of a random variable w is defined by $\|w\|_{\psi_2} = \inf\{c > 0 : \mathbb{E}\exp(w^2/c^2) \leq 2\}$ and set $\sigma = \|w\|_{\psi_2}$.

Theorem A. For every $L > 1$, $\kappa_0 > 0$ and $\beta > 1$, there exist constants c_0, c_1 and c_2 that depend only on L, κ_0 and β for which the following holds. Let a and w be as above and assume that w has a finite ψ_2-norm. If ℓ is the squared loss and $\hat{x}$ is produced by ERM, then with probability at least

$$1 - 2\exp(-c_0\min\{\ell^2(T_{+,r_2^*}), \ell^2(T_{-,r_2^*})\}) - 2N^{-\beta+1},$$

$$\|\hat{x} - x_0\|_2\|\hat{x} + x_0\|_2 \leq r_2^* := \max\{r_0(c_1), r_2(c_2/\sigma\sqrt{\log N})\}.$$

When $\|w\|_\infty < \infty$ the term $\sigma\sqrt{\log N}$ may be replaced by $\|w\|_\infty$.

The upper estimate of $\max\{r_0, r_2\}$ in Theorem A represents two ranges of noise. It follows from the definition of the fixed points that r_0 is dominant if $\sigma \leq r_0/\sqrt{\log N}$. As explained in [7] for linear regression, r_0 captures the difficulty of recovery in the noise free case, when the only reason for errors is that there are several far-away functions in the class that coincide with the target on the noiseless data. When the noise level σ surpasses that threshold, errors occur because of the interaction class members have with the noise, and the dominating term becomes r_2.

Of course, there are cases in which $r_0 = 0$ for N sufficiently large. This is precisely when exact recovery is possible in the noise-free environment. And, in such cases, the error of ERM tends to zero with σ.

The behavior of ERM in the noise-free case is one of the distinguishing features of sets with well behaved 'global complexity' – because E is not too large. Since $E_R \leq E$ for every $R > 0$, it follows that when $N \gtrsim E^2$, $r_0 = 0$ and that $r_2(\gamma) \leq E/(\gamma\sqrt{N})$. Therefore, on the event from Theorem A,

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \lesssim \sigma \frac{E}{\sqrt{N}} \sqrt{\log N}.$$

This estimate suffices for many applications. For example, when T is the set of d-sparse vectors, one may show (see, e.g. [2]) that

$$E \lesssim \sqrt{d \log(en/d)}.$$

Hence, by Theorem A, when $N \gtrsim d \log\left(en/d\right)$, with high probability,

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \lesssim \sigma \sqrt{\frac{d \log(en/d)}{N}} \sqrt{\log N}.$$

The proof of this observation regarding d-sparse vectors, and that this estimate is sharp in the minimax sense (up to the logarithmic term) may be found in Section 6.

One should note that Theorem A improves the main result from [2] in three ways. First of all, the error rate (the estimate on $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2$) established in Theorem A is $\sim E/\sqrt{N}$ (up to logarithmic factors), whereas in [2], it scaled like $c/N^{1/4}$ for very large values of N. Second, the error rate scales linearly in the noise level σ in Theorem A. On the other hand, the rate obtained in [2] does not decay with σ for $\sigma \leq 1$. Finally, the probability estimate has been improved, though it is still likely to be suboptimal.

Although the main motivation for [2] was dealing with phase retrieval for sparse classes, and for which Theorem A is well suited, we next turn to the question of more general classes, the most important example of which is a convex, centrally-symmetric class. For such a class, the global localization is simply too big to yield a good bound.

Definition 1.2 *Let*

$$r_N^*(Q) = \inf\left\{r > 0 : \ell(T \cap rB_2^n) \leq Qr\sqrt{N}\right\},$$

$$s_N^*(\eta) = \inf\left\{s > 0 : \ell(T \cap sB_2^n) \leq \eta s^2\sqrt{N}\right\},$$

6

and

$$v_N^*(\zeta) = \inf\left\{v > 0 : \ell(T \cap v B_2^n) \leq \zeta v^3 \sqrt{N}\right\},$$

The parameters r_N^* and s_N^* have been used in [7] to obtain a sharp estimate on the performance of ERM for linear regression in an arbitrary convex set, and relative to L-subgaussian measurements. This result is formulated below in a restricted context, analogous to the phase retrieval setup: a linear model $z = \langle a, x_0 \rangle + w$, for an isotropic, L-subgaussian vector a, independent noise w and $x_0 \in T$.

Let $\hat{x}$ be the output of ERM using the data $(a_i, z_i)_{i=1}^N$ and set $\|w\|_{\psi_2} = \sigma$.

Theorem 1.3 *For every $L \geq 1$ there exist constants c_1, c_2, c_3 and c_4 that depend only on L for which the following holds. Let $T \subset \mathbb{R}^d$ be a convex set, put $\eta = c_1/\sigma$ and set $Q = c_2$.*

1. *If $\sigma \geq c_3 r_N^*(Q)$ then with probability at least $1 - 4\exp(-c_4 N \eta^2 (s_N^*(\eta))^2)$,*

$$\|x - x_0\|_2 \leq s_N^*(\eta).$$

2. *If $\sigma \leq c_3 r_N^*(Q)$ then with probability at least $1 - 4\exp(-c_4 N Q^2)$,*

$$\|x - x_0\|_2 \leq r_N^*(Q).$$

Our main result is a phase retrieval version of Theorem 1.3.

Theorem B. For every $L \geq 1$, $\kappa_0 > 0$ and β there exist constants c_1, c_2, c_3, c_4, c_5 and Q that depend only on L and κ_0 and β for which the following holds. Let $T \subset \mathbb{R}^d$ be a convex, centrally-symmetric set, and let a and w be as in Theorem A. Assume that $(\sigma/\|x_0\|_2) \geq c_0 r_N^*(Q)/\sqrt{\log N}$, set $\eta = c_1 \|x_0\|_2/(\sigma\sqrt{\log N})$ and let $\zeta = c_1/(\sigma\sqrt{\log N})$.

1. If $\|x_0\|_2 \geq v_N^*(c_2)$, then with probability at least $1 - 2\exp(-c_3 N \eta^2 (s_N^*(\eta))^2) - 2N^{-\beta+1}$,

$$\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \leq c_4 s_N^*(\eta).$$

2. If $\|x_0\|_2 \leq v_N^*(c_2)$ then with probability at least $1 - 2\exp(-c_3 N \zeta^2 (v_N^*(\zeta))^2) - 2N^{-\beta+1}$,

$$\max\{\|x - x_0\|_2, \|x + x_0\|_2\} \leq c_4 v_N^*(\zeta).$$

If $(\sigma/\|x_0\|_2) \geq c_0 r_N^*(Q)/\sqrt{\log N}$ the same assertion as in 1. and 2. holds, with an upper bound of $r_N^*(Q)$ replacing $s_N^*(\eta)$ and $v_N^*(\zeta)$.

Theorem B follows from Theorem A and a more transparent description of the localized sets $T_{-,R}(x_0)$ and $T_{+,R}(x_0)$ (see Lemma 4.1).

To put Theorem B in some perspective, observe that v_N^* tends to zero. Indeed, since $\ell(T \cap rB_2^n) \leq \ell(T)$, it follows that $v_N^*(\zeta) \leq (\ell(T)/\sqrt{N}\zeta)^{1/3}$. Hence, for the choice of $\zeta \sim (\sigma\sqrt{\log N})^{-1}$ as in Theorem B,

$$v_N^* \leq \left(\sigma\ell(T)\sqrt{\frac{\log N}{N}}\right)^{1/3},$$

which tends to zero when $\sigma \to 0$ and when $N \to \infty$. Therefore, if $x_0 \neq 0$, the first part of Theorem B describes the 'long term' behaviour of ERM.

Also, and using the same argument,

$$r_N^*(Q) \leq \frac{\ell(T)}{Q\sqrt{N}}.$$

Thus, for every $\sigma > 0$ the problem becomes 'high noise' in the sense that the condition $(\sigma/\|x_0\|_2) \geq c_0 r_N(Q)/\sqrt{\log N}$ is satisfied when N is large enough.

In the typical situation, which is both 'high noise' and 'large $\|x_0\|_2$', the error rate depends on $\eta = c_1\|x_0\|_2/\sigma\sqrt{\log N}$. We believe that the $1/\sqrt{\log N}$ factor is an artifact of the proof, but the other term, $\|x_0\|_2/\sigma$ is the signal-to-noise ratio, and is rather natural.

Although Theorem A and Theorem B clearly improve the results from [2], it is natural to ask whether these are optimal in a more general sense. The final result presented here is that Theorem B is close to being optimal in the minimax sense. The formulation and proof of the minimax lower bound is presented in Section 5.

Finally, we end the article with two examples of classes that are of interest in phase retrieval: d-sparse vectors and the unit ball in ℓ_1^n. The first is a class with a fixed 'local complexity', and the second has a growing 'local complexity'.

2   Preliminaries

Throughout this article, absolute constants are denoted by $C, c, c_1, ...$ etc. Their value may change from line to line. The fact that there are absolute constants c, C for which $ca \leq b \leq Ca$ is denoted by $a \sim b$; $a \lesssim b$ means that $a \leq cb$, while $a \sim_L b$ means that the constants depend only on the parameter L.

For $1 \leq p \leq \infty$, let $\|\cdot\|_p$ be the ℓ_p^n norm endowed on $\mathbb{R}^n$, and for a function f (or a random variable X) on a probability space, set $\|f\|_{L_p}$ to be its L_p norm.

Other norms that play a significant role here are the Orlicz norms. For basic facts on these norms we refer the reader to [9, 15].

Recall that for $\alpha \geq 1$,

$$\|f\|_{\psi_\alpha} = \inf\{c > 0 : \mathbb{E}\exp(|f|^\alpha/c^\alpha) \leq 2\},$$

and it is straightforward to extend the definition for $0 < \alpha < 1$.

Orlicz norms measure the rate of decay of a function. One may verify that $\|f\|_{\psi_\alpha} \sim \sup_{p \geq 1} \|f\|_{L_p}/p^{1/\alpha}$. Moreover, for $t \geq 1$, $Pr(|f| \geq t) \leq 2\exp(-ct^\alpha/\|f\|_{\psi_\alpha}^\alpha)$, and $\|f\|_{\psi_\alpha}$ is equivalent to the smallest constant κ for which $Pr(|f| \geq t) \leq 2\exp(-t^\alpha/\kappa^\alpha)$ for every $t \geq 1$.

Definition 2.1 *A random variable is L-subgaussian if it has a bounded ψ_2 norm and $\|X\|_{\psi_2} \leq L\|X\|_{L_2}$.*

Observe that for L-subgaussian random variables, all the L_p norms are equivalent and their tails exhibits a faster decay than the corresponding gaussian. Indeed, if X is L-subgaussian,

$$\|X\|_{L_p} \lesssim \sqrt{p}\|X\|_{\psi_2} \lesssim L\sqrt{p}\|X\|_{L_2},$$

and for every $t \geq 1$,

$$Pr(|X| > t) \leq 2\exp(-ct^2/\|X\|_{\psi_2}^2) \leq 2\exp(-ct^2/(L^2\|X\|_{L_2}^2))$$

for a suitable absolute constant c.

It is standard to verify that for every f, g, $\|fg\|_{\psi_1} \lesssim \|f\|_{\psi_2}\|g\|_{\psi_2}$, and that if $X_1, ..., X_N$ are independent copies of X and $1 \leq \alpha \leq 2$, then

$$\|\max_{1 \leq i \leq N} X_i\|_{\psi_\alpha} \lesssim \|X\|_{\psi_\alpha} \log^{1/\alpha} N. \tag{2.1}$$

An additional feature of ψ_α random variables is concentration, namely that if $(X_i)_{i=1}^N$ are independent copies of a ψ_α random variable X, then $N^{-1}\sum_{i=1}^N X_i$ concentrates around $\mathbb{E}X$. One example of such a concentration result is the following Bernstein-type inequality (see, e.g., [15]).

Theorem 2.2 *There exists an absolute constant c_0 for which the following holds. If $X_1, ..., X_N$ are independent copies of a ψ_1 random variable X, then for every $t > 0$,*

$$Pr\left(\left|\frac{1}{N}\sum_{i=1}^N X_i - \mathbb{E}X\right| > t\|X\|_{\psi_1}\right) \leq 2\exp(-c_0 N \min\{t^2, t\}).$$

One important example of a probability space considered here is the discrete space $\Omega = \{1, ..., N\}$, endowed with the uniform probability measure. Functions on Ω can be viewed as vectors in $\mathbb{R}^N$ and the corresponding L_p and ψ_α norms are denoted by $\| \cdot \|_{L_p^N}$ and $\| \cdot \|_{\psi_\alpha^N}$.

A significant part of the proofs of Theorem A has to do with the behaviour of a monotone non-increasing rearrangement of vectors. Given $v \in \mathbb{R}^N$, let $(v_i^*)_{i=1}^N$ be a non-increasing rearrangement of $(|v_i|)_{i=1}^N$. It turns out that the ψ_α^N norm captures information on the coordinates of $(v_i^*)_{i=1}^N$.

Lemma 2.3 *For every $1 \leq \alpha \leq 2$ there exist constants c_1 and c_2 that depend only on α for which the following holds. For every $v \in \mathbb{R}^N$,*

$$c_1 \sup_{i \leq N} \frac{v_i^*}{\log^{1/\alpha}(eN/i)} \leq \|v\|_{\psi_\alpha^N} \leq c_2 \sup_{i \leq N} \frac{v_i^*}{\log^{1/\alpha}(eN/i)}.$$

Proof. We will prove the claim only for $\alpha = 2$ as the other cases follow an identical path.

Let $v \in \mathbb{R}^N$ and denote by Pr the uniform probability measure on $\Omega = \{1, \ldots, N\}$. By the tail characterization of the ψ_2 norm,

$$N^{-1} |\{j : |v_j| > t\}| = Pr(|v| > t) \leq 2 \exp(-ct^2/\|v\|_{\psi_2^N}^2).$$

Hence, for $t_i = c^{-1/2} \|v\|_{\psi_2^N} \sqrt{\log(eN/i)}$, $|\{j : |v_j| > t_i\}| \leq 2i/e \leq i$, and for every $1 \leq i \leq N$, $v_i^* \leq t_i$. Therefore,

$$\sup_{i \leq N} \frac{v_i^*}{\sqrt{\log(eN/i)}} \leq c^{-1/2} \|v\|_{\psi_2^N},$$

as claimed.

In the reverse direction, consider

$$\mathcal{B} = \left\{ \beta > 0 : \forall \, 1 \leq i \leq N, \ \|v\|_{\psi_2^N} \geq \beta v_i^* / \sqrt{\log(eN/i)} \right\}.$$

It is enough to show that $\mathcal{B}$ is bounded by a constant that is independent of v. To that end, fix $\beta \in \mathcal{B}$ and without loss of generality, assume that $\beta > 2$. Set $B = \sup_{i \leq N} \beta v_i^* / \sqrt{\log(eN/i)}$ and since $\beta \in \mathcal{B}$, $\|v\|_{\psi_2^N} \geq B$.

Also, since $1/\beta^2 < 1$,

$$\sum_{i=1}^N \left(\frac{1}{i}\right)^{1/\beta^2} \leq 1 + \int_1^N \left(\frac{1}{x}\right)^{1/\beta^2} dx \leq \frac{N^{1-1/\beta^2}}{1 - 1/\beta^2}.$$

10

Therefore,

$$\sum_{i=1}^{N} \exp(v_i^2/B^2) = \sum_{i=1}^{N} \exp((v_i^*)^2/B^2) \leq \sum_{i=1}^{N} \exp(\beta^{-2} \log(eN/i))$$

$$\leq \sum_{i=1}^{N} \left(\frac{eN}{i}\right)^{1/\beta^2} \leq (eN)^{1/\beta^2} \cdot \frac{N^{1-1/\beta^2}}{1-1/\beta^2} \leq \frac{Ne^{1/\beta^2}}{1-1/\beta^2} < 2N,$$

provided that $\beta \geq c_1$. Thus, if $\beta \geq c_1$, $\|v\|_{\psi_2^N} < B$ which is a contradiction, showing that $\mathcal{B}$ is bounded by c_1. ∎

2.1 Empirical and Subgaussian processes

The sampling method used here is isotropic and L-subgaussian, meaning that the vectors $(a_i)_{i=1}^{N}$ are independent and distributed according to a probability measure μ on $\mathbb{R}^n$ that is both isotropic and L-subgaussian [15]:

Definition 2.4 *Let μ be a probability measure on $\mathbb{R}^n$ and let a be distributed according to μ. The measure μ is isotropic if for every $t \in \mathbb{R}^n$, $\mathbb{E}\langle a, t\rangle^2 = \|t\|_2^2$. It is L-subgaussian if for every $t \in \mathbb{R}^n$ and every $u \geq 1$, $Pr(|\langle a, t\rangle| \geq Lu\|\langle t, a\rangle\|_2) \leq 2\exp(-u^2/2)$.*

Given $T \subset \mathbb{R}^n$, let $d_T = \sup_{t\in T} \|t\|_2$ and put $k_*(T) = (\ell(T)/d_T)^2$. The latter appears naturally in the context of Dvoretzky type theorems, and in particular, in Milman's proof of Dvoretzky's Theorem (see, e.g., [11]).

Theorem 2.5 *[10] For every $L \geq 1$ there exist constants c_1 and c_2 that depend only on L and for which the following holds. For every $u \geq c_1$, with probability at least $1 - 2\exp(-c_2 u^2 k_*(T))$, for every $t \in T$ and every $I \subset \{1, ..., N\}$,*

$$\left(\sum_{i \in I} \langle t, a_i\rangle^2\right)^{1/2} \leq Lu^3 \left(\ell(T) + d_T \sqrt{|I| \log(eN/|I|)}\right).$$

For every integer N, let j_T be the largest integer j in $\{1, ..., N\}$ for which

$$\ell(T) \geq d_T \sqrt{j \log(eN/j)}.$$

It follows from Theorem 2.5 that if $t \in T$ and $|I| \leq j_T$,

$$(\sum_{i \in I} \langle t, a_i\rangle^2)^{1/2} \lesssim_{L,u} \ell(T),$$

11

and if $|I| \geq j_T$,

$$(\sum_{i \in I} \langle t, a_i \rangle^2)^{1/2} \lesssim_{L,u} d_T \sqrt{|I| \log(eN/|I|)}.$$

Therefore, if $v = (\langle t, a_i \rangle)_{i=1}^N$ and $(v_i^*)_{i=1}^N$ is a monotone non-increasing rearrangement of $(|v_i|)_{i=1}^N$, then

$$v_i^* \leq \left(\frac{1}{i} \sum_{j=1}^i (v_j^*)^2 \right)^{1/2} \lesssim_{L,u} \begin{cases} \frac{\ell(T)}{\sqrt{i}} & \text{if } i \leq j_T \\ d_T \sqrt{\log(eN/i)} & \text{otherwise.} \end{cases} \tag{2.2}$$

This observation will be used extensively in what follows.

The next fact deals with product processes.

Theorem 2.6 *[10] There exist absolute constants c_0, c_1 and c_2 for which the following holds. If $T_1, T_2 \subset \mathbb{R}^n$, $1 \leq 2^j \leq N$ and $u \geq c_0$, then with probability at least $1 - 2\exp(-c_1 u^2 2^j)$,*

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^N \langle a_i, t \rangle \langle a_i, s \rangle - \mathbb{E}\langle a, t \rangle \langle a, s \rangle \right|$$

$$\leq c_2 L^2 u^2 \left(\ell(T_1)\ell(T_2) + u\sqrt{N} \left(\ell(T_1)d(T_2) + \ell(T_2)d(T_1) + 2^{j/2}d(T_1)d(T_2) \right) \right)$$

and

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^N |\langle a_i, t \rangle \langle a_i, s \rangle| - \mathbb{E}|\langle a, t \rangle \langle a, s \rangle| \right|$$

$$\leq c_2 L^2 u^2 \left(\ell(T_1)\ell(T_2) + u\sqrt{N} \left(\ell(T_1)d(T_2) + \ell(T_2)d(T_1) + 2^{j/2}d(T_1)d(T_2) \right) \right).$$

Remark 2.7 *Let $(\varepsilon_i)_{i=1}^N$ be independent, symmetric, $\{-1, 1\}$-valued random variables. It follows from the results in [10] that with the same probability estimate in Theorem 2.6 and relative to the product measure $(\varepsilon \otimes X)^N$,*

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^N \varepsilon_i \langle a_i, t \rangle \langle a_i, s \rangle \right|$$

$$\lesssim L^2 u^2 \left(\ell(T_1)\ell(T_2) + u\sqrt{N} \left(\ell(T_1)d(T_2) + \ell(T_2)d(T_1) + 2^{j/2}d(T_1)d(T_2) \right) \right).$$

Assume that $(k^*(T_1))^{1/2} = \ell(T_1)/d(T_1) \geq \ell(T_2)/d(T_2)$. Setting $2^{j/2} = \ell(T_1)/d(T_1)$, Theorem 2.6 and Remark 2.7 yield that with probability at least $1 - 2\exp(-c_1 u^2 k_*(T_1))$,

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^{N} \langle a_i, t \rangle \langle a_i, s \rangle - \mathbb{E}\langle a, t \rangle \langle a, s \rangle \right| \lesssim L^2 u^2 \ell(T_1) \left(\ell(T_2) + u\sqrt{N}d(T_2) \right),$$

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^{N} |\langle a_i, t \rangle \langle a_i, s \rangle| - \mathbb{E}|\langle a, t \rangle \langle a, s \rangle| \right| \lesssim L^2 u^2 \ell(T_1) \left(\ell(T_2) + u\sqrt{N}d(T_2) \right)$$

and

$$\sup_{t \in T_1, \ s \in T_2} \left| \sum_{i=1}^{N} \varepsilon_i \langle a_i, t \rangle \langle a_i, s \rangle \right| \lesssim L^2 u^2 \ell(T_1) \left(\ell(T_2) + u\sqrt{N}d(T_2) \right). \qquad (2.3)$$

One case which is of particular interest is when $T_1 = T_2 = T$, and then, with probability at least $1 - 2\exp(-c_1 u^2 k_*(T))$,

$$\sup_{t \in T} \left| \sum_{i=1}^{N} \langle a_i, t \rangle^2 - \mathbb{E}\langle a, t \rangle^2 \right| \lesssim L^2 u^2 \left(\ell^2(T) + u\sqrt{N}d(T)\ell(T) \right).$$

2.2   Monotone rearrangement of coordinates

The first goal of this section is to investigate the coordinate structure of $v \in \mathbb{R}^m$, given information on its norm in various L_p^m and ψ_α^m spaces. The vectors we will be interested in are of the form $(\langle a_i, t \rangle)_{i=1}^{N}$ for $t \in T$, and for which, thanks to the results presented in Section 2.1, one has the necessary information at hand.

It is standard to verify that if $\|v\|_{\psi_\alpha^m} \leq A$, then $\|v\|_p \lesssim_{p,\alpha} A \cdot m^{1/p}$. Thus, $\|v\|_{L_p^m} \lesssim_p \|v\|_{\psi_\alpha^m}$.

It turns out that if the two norms are equivalent, v is regular in some sense. The next lemma, which is a version of the Paley-Zygmund Inequality, (see, e.g. [4]), describes the regularity properties needed here in the case $p = \alpha = 1$.

Lemma 2.8 *For every $\beta > 1$ there exist constants c_1 and c_2 that depend only on β and for which the following holds. If $\|v\|_{\psi_1^m} \leq \beta\|v\|_{L_1^m}$, there exists $I \subset \{1, ..., m\}$ of cardinality at least $c_1 m$, and for every $i \in I$, $|v_i| \geq c_2\|v\|_{L_1^m}$.*

13

Proof. Recall that $\|v\|_{\psi_1^m} \sim \sup_{1 \leq i \leq m} v_i^* / \log(em/i)$. Hence, for every $1 \leq j \leq m$,

$$\sum_{\ell=1}^{j} v_\ell^* \lesssim \|v\|_{\psi_1^m} \sum_{\ell=1}^{j} \log(em/\ell) \lesssim \beta \|v\|_{L_1^m} j \log(em/j).$$

Therefore,

$$m\|v\|_{L_1^m} = \sum_{\ell=1}^{m} |v_\ell| = \sum_{\ell \leq j} v_\ell^* + \sum_{\ell=j+1}^{m} v_\ell^* \leq c_0 \beta \|v\|_{L_1^m} j \log(em/j) + \sum_{\ell=j+1}^{m} v_\ell^*.$$

Setting $c_1(\beta) \sim 1/(\beta \log(e\beta))$ and $j = c_1(\beta)m$,

$$c_0 \beta \|v\|_{L_1^m} j \log(em/j) \leq (m/2)\|v\|_{L_1^m}.$$

Thus, $\sum_{\ell=j+1}^{m} v_\ell^* \geq (m/2)\|v\|_{L_1^m}$, while

$$v_{j+1}^* \leq \frac{1}{j+1} \sum_{\ell \leq j+1} v_\ell^* \lesssim \beta \log(e\beta) \|v\|_{L_1^m}.$$

Let I be the set of the $m - j$ smallest coordinates of v. Fix $\eta > 0$ to be named later, put $I_\eta \subset I$ to be the set of coordinates in I for which $|v_i| \geq \eta \|v\|_{L_1^m}$ and denote by I_η^c its complement in I. Therefore,

$$(m/2)\|v\|_{L_1^m} \leq \sum_{\ell \geq j+1} v_\ell^* = \sum_{\ell \in I_\eta} |v_\ell| + \sum_{\ell \in I_\eta^c} |v_\ell| \leq v_{j+1}^* |I_\eta| + \eta \|v\|_{L_1^m} |I_\eta^c|$$

$$\lesssim \|v\|_{L_1^m} |I| \left(\beta \log(e\beta) \frac{|I_\eta|}{|I|} + \eta \frac{|I_\eta^c|}{|I|} \right).$$

Hence,

$$\frac{m}{2} \lesssim |I| \left(\beta \log(e\beta) \frac{|I_\eta|}{|I|} + \eta \left(1 - \frac{|I_\eta|}{|I|} \right) \right) \lesssim m \left((\beta \log(e\beta) - \eta) \frac{|I_\eta|}{|I|} + \eta \right).$$

If $\eta = \min\{1/4, (\beta/2) \log(e\beta)\}$, then $|I_\eta| \geq (\eta/2)|I| \geq c_2(\beta)m$, as claimed. ∎

Next, let us turn to decomposition results for vectors of the form $(\langle a_i, t \rangle)_{i=1}^{N}$. Recall that for a set $T \subset \mathbb{R}^N$, j_T is the largest integer for which $\ell(T) \geq d_T \sqrt{j \log(eN/j)}$.

Lemma 2.9 *For every $L > 1$ there exist constants c_1 and c_2 that depend only on L and for which the following holds. Let $T \subset \mathbb{R}^n$ and set $W = \{t/\|t\|_2 : t \in T\} \subset S^{n-1}$. With probability at least $1 - 2\exp(-c_1\ell^2(W))$, for every $t \in T$, $(\langle a_i, t \rangle)_{i=1}^N = v_1 + v_2$ and v_1, v_2 have the following properties:*

1. *The supports of v_1 and v_2 are disjoint.*

2. *$\|v_1\|_2 \leq c_2\ell(W)\|t\|_2$ and $|\text{supp}(v_1)| \leq j_W$.*

3. *$\|v_2\|_{\psi_2^N} \leq c_2\|t\|_2$.*

Proof. Fix $t \in T$ and let $J_t \subset \{1, ..., N\}$ be the set of the largest j_W coordinates of $(|\langle a_i, t \rangle|)_{i=1}^N$. Set

$$\bar{v}_1 = (\langle a_j, t/\|t\|_2 \rangle)_{j \in J_t} \quad \text{and} \quad \bar{v}_2 = (\langle a_j, t/\|t\|_2 \rangle)_{j \in J_t^c}.$$

By Theorem 2.5 and the characterization of the ψ_2^N norm of a vector using the monotone rearrangement of its coordinates (Lemma 2.3),

$$\|\bar{v}_1\|_2 \lesssim L\ell(W), \quad \text{and} \quad \|\bar{v}_2\|_{\psi_2^N} \lesssim L.$$

To conclude the proof, set $v_1 = \|t\|_2\bar{v}_1$ and $v_2 = \|t\|_2\bar{v}_2$. ∎

Recall that for every $R > 0$,

$$T_{+,R} = \left\{ \frac{t+s}{\|t+s\|_2} : t, s \in T, \; \|t+s\|_2\|t-s\|_2 \geq R \right\},$$

and a similar definition holds for $T_{-,R}$. Set $j_{+,R} = j_{T_{+,R}}$, $j_{-,R} = j_{T_{+,R}}$ and $E_R = \max\{\ell(T_{+,R}), \ell(T_{-,R})\}$. Combining the above estimates leads to the following corollary.

Corollary 2.10 *For every $L > 1$ there exist constants c_1, c_2, c_3 and c_4 that depend only on L for which the following holds. Let $T \subset \mathbb{R}^n$ and $R > 0$, and consider $T_{+,R}$ and $T_{-,R}$ as above. With probability at least $1 - 4\exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\})$, for every $s, t \in T$ for which $\|t - s\|_2\|t + s\|_2 \geq R$,*

1. *$(\langle s - t, a_i \rangle)_{i=1}^N = v_1 + v_2$, for vectors v_1 and v_2 of disjoint supports satisfying*

$$|\text{supp}(v_1)| \leq j_{-,R}, \quad \|v_1\|_2 \leq c_2\ell(T_{-,R})\|s-t\|_2 \quad \text{and} \quad \|v_2\|_{\psi_2^N} \leq c_2\|s-t\|_2.$$

2. $(\langle s + t, a_i \rangle)_{i=1}^N = u_1 + u_2$, for vectors u_1 and u_2 of disjoint supports satisfying

$$|\mathrm{supp}(u_1)| \leq j_{+,R}, \quad \|u_1\|_2 \leq c_2 \ell(T_{+,R}) \|s+t\|_2 \quad \text{and} \quad \|u_2\|_{\psi_2^N} \leq c_2 \|s+t\|_2.$$

3. If $h_{s,t}(a) = \langle \frac{s+t}{\|s+t\|_2}, a \rangle \langle \frac{s-t}{\|s-t\|_2}, a \rangle$, then

$$\left| \frac{1}{N} \sum_{i=1}^N |h_{s,t}(a_i)| - \mathbb{E}|h_{s,t}| \right| \leq c_3 \left(\frac{E_R}{\sqrt{N}} + \frac{E_R^2}{N} \right).$$

In particular, recalling that for every $s, t \in T$,

$$\mathbb{E}|\langle s+t, a \rangle \langle s-t, a \rangle| \geq \kappa_0 \|s+t\|_2 \|s-t\|_2,$$

it follows that if $\sqrt{N} \geq c_4(L) E_R/\kappa_0$ then

4.

$$\frac{\kappa_0}{2} \|s+t\|_2 \|s-t\|_2 \leq \frac{1}{N} \sum_{i=1}^N |\langle s+t, a_i \rangle \langle s-t, a_i \rangle| \lesssim_L \|s+t\|_2 \|s-t\|_2.$$

$$(2.4)$$

From here on, denote by $\Omega_{1,R}$ the event on which Corollary 2.10 holds for the sets $T_{+,R}$ and $T_{-,R}$ and samples of cardinality $N \gtrsim_L E_R^2/\kappa_0^2$.

Lemma 2.11 *There exist constants c_0 depending only on L and c_1, κ_1 that depend only on κ_0 and L for which the following holds. If $N \geq c_0 E_R^2/\kappa_0^2$, then for $(a_i)_{i=1}^N \in \Omega_{1,R}$, for every $s, t \in T$ for which $\|s-t\|_2 \|s+t\|_2 \geq R$, there is $I_{s,t} \subset \{1, ..., N\}$ of cardinality at least $\kappa_1 N$, and for every $i \in I_{s,t}$,*

$$|\langle s-t, a_i \rangle \langle s+t, a_i \rangle| \geq c_1 \|s-t\|_2 \|s+t\|_2.$$

Lemma 2.11 is an empirical "small-ball" estimate, as it shows that with high probability, and for every pair s, t as above, many of the coordinates of $(|\langle a_i, s-t \rangle| \cdot |\langle a_i, s+t \rangle|)_{i=1}^N$ are large.
Proof. Fix $s, t \in T$ as above and set

$$y = (\langle s-t, a_i \rangle)_{i=1}^N, \quad \text{and} \quad x = (\langle s+t, a_i \rangle)_{i=1}^N.$$

Let $y = v_1 + v_2$ and $x = u_1 + u_2$ as in Corollary 2.10. Let $j_0 = \max\{j_{-,R}, j_{+,R}\}$ and put $J = \mathrm{supp}(v_1) \cup \mathrm{supp}(u_1)$. Observe that $|J| \leq 2j_0$ and that

$$\sum_{j \in J} |y(j)| \cdot |x(j)| \leq \sum_{j \in \mathrm{supp}(v_1)} |v_1(j)x(j)| + \sum_{j \in \mathrm{supp}(u_1)} |y(j)u_1(j)|$$

$$\leq \|v_1\|_2 \left(\sum_{i=1}^{2j_0} (x^2(j))^* \right)^{1/2} + \|u_1\|_2 \left(\sum_{i=1}^{2j_0} (y^2(j))^* \right)^{1/2}$$

$$\lesssim_L \ell(T_{-,R}) \|s - t\|_2 \cdot \sqrt{j_0 \log(eN/j_0)} \|s + t\|_2$$

$$+ \ell(T_{+,R}) \|s + t\|_2 \cdot \sqrt{j_0 \log(eN/j_0)} \|s - t\|_2$$

$$\lesssim_L E_R^2 \|s - t\|_2 \|s + t\|_2 \leq \frac{\kappa_0 N}{4} \|s - t\|_2 \|s + t\|_2,$$

because, by the definition of j_0, $\sqrt{j_0 \log(eN/j_0)} \lesssim \max\{\ell(T_{-,R}), \ell(T_{+,R})\}$ and since $N \geq c_0 E_R^2/\kappa_0^2$ for $c_0 = c_0(L)$ large enough.

Thus, by (2.4),

$$\sum_{j \in J^c} |y(j)x(j)| \geq N\kappa_0 \|s - t\|_2 \|s + t\|_2/4.$$

Set $m = |J^c|$ and let $z = (y(j)x(j))_{j \in J^c} = (v_2(j)u_2(j))_{j \in J^c}$. Since $N \gtrsim_L E_R^2/\kappa_0^2$, it is evident that $j_0 \leq N/2$; thus $N/2 \leq m \leq N$ and

$$\|z\|_{L_1^m} = \frac{1}{m} \sum_{j \in J^c} |y(j)x(j)| \geq \frac{N}{4m} \kappa_0 \|s - t\|_2 \|s + t\|_2 \gtrsim \kappa_0 \|s - t\|_2 \|s + t\|_2.$$

On the other hand,

$$\|z\|_{\psi_1^m} \leq \|(v_2 u_2(j))_{j \in J_c}\|_{\psi_1^m} \lesssim \|v_2\|_{\psi_2^m} \|u_2\|_{\psi_2^m} \lesssim_L \|s - t\|_2 \|s + t\|_2,$$

and z satisfies the assumption of Lemma 2.8 for $\beta = c_1(L, \kappa_0)$. The claim follows immediately from that lemma. ∎

3   Proof of Theorem A

It is well understood that when analyzing properties of ERM relative to a loss ℓ, studying the excess loss functional is rather natural. The excess loss shares the same empirical minimizer as the loss, but it has additional qualities: for every $x \in T$, $\mathbb{E}\mathcal{L}_x \geq 0$ and $\mathcal{L}_{x_0} = 0$.

Since 0 is a potential minimizer of $\{P_N \mathcal{L}_x : x \in T\}$, the minimizer $\hat{x}$ satisfies that $P_N \mathcal{L}_{\hat{x}} \leq 0$, giving one a way of excluding parts of T as potential empirical minimizers. One simply has to show that with high probability, those parts belong to the set $\{x : P_N \mathcal{L}_x > 0\}$, for example, by showing that $P_N \mathcal{L}_x$ is equivalent to $\mathbb{E}\mathcal{L}_x$, as the latter is positive for points that are not true minimizers.

The squared excess loss has a simple decomposition to two processes: a quadratic process and a multiplier one. Indeed, given a class of functions F and $f \in F$,

$$\big(f(a) - y\big)^2 - \big(f^*(a) - y\big)^2 = \big(f(a) - f^*(a)\big)^2 - 2\big(f(a) - f^*(a)\big)\big(f^*(a) - y\big).$$

where, as always, f^* is a minimizer of the functional $\mathbb{E}\big(f(a) - y\big)^2$ in F.

In the phase retrieval problem, $y = \big\langle x_0, a\big\rangle^2 + w$ for a noise w that is independent of a, and each $f_x \in F$ is given by $f_x = \big\langle x, \cdot \big\rangle^2$. Thus,

$$\begin{aligned}
\mathcal{L}_x(a, y) = \ell_x(a, y) - \ell_{x_0}(a, y) &= \big(f_x(a) - y\big)^2 - \big(f_{x_0}(a) - y\big)^2 \\
&= \big(\langle x - x_0, a\rangle\langle x + x_0, a\rangle\big)^2 - 2w\langle x - x_0, a\rangle\langle x + x_0, a\rangle.
\end{aligned}$$

Since w is a mean-zero random variable that is independent of a, and by Assumption 1.1,

$$\mathbb{E}\mathcal{L}_x(a, y) = \mathbb{E}\big|\langle x - x_0, a\rangle\langle x + x_0, a\rangle\big| \geq \kappa_0^2 \|x - x_0\|_2^2 \|x + x_0\|_2^2.$$

Therefore, $\mathbb{E}\big(f_x(a) - y\big)^2$ has a unique minimizer in F: $f^* = f_{x_0} = f_{-x_0}$.

To show that $P_N \mathcal{L}_x > 0$ on a large subset $T' \subset T$, it suffices to obtain a high probability lower bound on

$$\inf_{x \in T'} \frac{1}{N} \sum_{i=1}^{N} \big(\langle x - x_0, a_i\rangle\langle x + x_0, a_i\rangle\big)^2$$

that dominates a high probability upper bound on

$$\sup_{x \in T'} \left| \frac{2}{N} \sum_{i=1}^{N} w_i \langle x - x_0, a_i\rangle\langle x + x_0, a_i\rangle \right|.$$

The set T' that will be used is $T_R = \{x \in T : \|x - x_0\|_2 \|x + x_0\|_2 \geq R\}$ for a well-chosen R.

Theorem 3.1 *There exist a constant c_0 depending only on L, and constants c_1, κ_1 depending only on κ_0 and L for which the following holds. For every $R > 0$ and $N \geq c_0 E_R^2/\kappa_0^2$, with probability at least*

$$1 - 4\exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\}),$$

for every $x \in T_R$,

$$\frac{1}{N}\sum_{i=1}^{N}\langle x_0 - x, a_i\rangle^2 \langle x_0 + x, a_i\rangle^2 \geq c_1\|x_0 - x\|_2^2\|x_0 + x\|_2^2.$$

Theorem 3.1 is an immediate outcome of Lemma 2.11

Theorem 3.2 *There exist absolute constants c_1 and c_2 for which the following holds. For every $\beta > 1$, with probability at least*

$$1 - 2\exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\}) - 2N^{-(\beta-1)},$$

for every $x \in T_R$,

$$\left|\frac{1}{N}\sum_{i=1}^{N} w_i\langle x - x_0, a_i\rangle\langle x + x_0, a_i\rangle\right| \leq c_2\sqrt{\beta}\|w\|_{\psi_2}\sqrt{\log N}\cdot\frac{E_R}{\sqrt{N}}\|x - x_0\|_2\|x + x_0\|_2.$$

Proof. By standard properties of empirical process, and since w is mean-zero and independent of a, it suffices to estimate

$$\sup_{x \in T_R}\left|\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i|w_i|\langle x - x_0, a_i\rangle\langle x + x_0, a_i\rangle\right|,$$

for independent signs $(\varepsilon_i)_{i=1}^{N}$. By the contraction principle for Bernoulli processes (see, e.g., [9]), it follows that for every fixed $(w_i)_{i=1}^{N}$ and $(a_i)_{i=1}^{N}$,

$$Pr_\varepsilon\left(\sup_{x \in T_R}\left|\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i|w_i|\langle\frac{x - x_0}{\|x - x_0\|_2}, a_i\rangle\langle\frac{x + x_0}{\|x + x_0\|_2}, a_i\rangle\right| > u\right)$$

$$\leq 2Pr_\varepsilon\left(\max_{i \leq N}|w_i| \cdot \sup_{x \in T_R}\left|\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i\langle\frac{x - x_0}{\|x - x_0\|_2}, a_i\rangle\langle\frac{x + x_0}{\|x + x_0\|_2}, a_i\rangle\right| > \frac{u}{2}\right).$$

Applying Remark 2.7, if $N \gtrsim_L E_R$ then with $(\varepsilon \otimes a)^N$-probability of at least $1 - 2\exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\})$,

$$\sup_{x \in T_R}\left|\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i\langle\frac{x - x_0}{\|x - x_0\|_2}, a_i\rangle\langle\frac{x + x_0}{\|x + x_0\|_2}, a_i\rangle\right| \leq c_2 L^2\frac{E_R}{\sqrt{N}}.$$

Also, because w is a ψ_2 random variable,

$$Pr(w_1^* \geq t\|w\|_{\psi_2}) \leq 2N \exp(-t^2/2),$$

and thus, $w_1^* \leq \sqrt{2\beta \log N}\|w\|_{\psi_2}$ with probability at least $1 - 2N^{-\beta+1}$.

Combining the two estimates and a Fubini argument, it follows that with probability at least $1 - 2\exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\}) - 2N^{-\beta+1}$, for every $x \in T_R$,

$$\left|\frac{1}{N}\sum_{i=1}^{N} w_i\langle x - x_0, a_i\rangle\langle x + x_0, a\rangle\right| \leq c_3 L^2 \sqrt{\beta}\|w\|_{\psi_2}\sqrt{\log N}\frac{E_R}{\sqrt{N}}\cdot\|x-x_0\|_2\|x+x_0\|_2.$$

■

On the intersection of the two events appearing in Theorem 3.1 and Theorem 3.2, if $N \gtrsim_{\kappa_0,L} E_R^2$ and setting

$$\rho = \|x - x_0\|_2\|x + x_0\|_2 \geq R \geq r_2(c_1\kappa_0/(c_2 L^2\sqrt{\beta})),$$

then for every $x \in T_R$,

$$P_N \mathcal{L}_x \geq \left(c_1\kappa_0^2\rho - c_2 L^2\sqrt{\beta}\|w\|_{\psi_2}\sqrt{\log N}\frac{E_R}{\sqrt{N}}\right)\rho$$

$$\geq \left(c_1\kappa_0^2 R - c_2 L^2\sqrt{\beta}\|w\|_{\psi_2}\sqrt{\log N}\frac{E_R}{\sqrt{N}}\right)R.$$

Therefore, if $N \gtrsim_{L,\kappa_0} E_R^2$ and

$$E_R \leq c_3(L, \kappa_0)\frac{R}{\|w\|_{\psi_2}}\sqrt{\frac{N}{\beta \log N}}, \tag{3.1}$$

then $P_N \mathcal{L}_x > 0$ and $\hat{x} \notin T_R$. Theorem A follows from the definition of $r_2(\gamma)$ for a well chosen γ.

4   Proof of Theorem B

Most of the work required for the proof of Theorem B has been carried out in Section 3. A literally identical argument, in which one replaces the sets $T_{+,R}$ and $T_{-,R}$ with $T_{+,R}(x_0)$ and $T_{-,R}(x_0)$ may be used, leading to an analogous version of Theorem A, with the obvious modifications: the complexity parameter is $\max\{\ell(T_{+,R}(x_0)), \ell(T_{-,R}(x_0))\}$ for the right choice of R, and

the probability estimate is $1 - 2\exp(-c_0\min\{\ell^2(T_{+,R}(x_0)), \ell^2(T_{-,R}(x_0))\}) - N^{-\beta+1}$.

All that remains to complete the proof of Theorem B is to analyze the structure of the local sets and identify the fixed points r_0 and r_2. A first step in that direction is the following:

Lemma 4.1 *There exist absolute constants c_1 and c_2 for which the following holds. For every $R > 0$ and $\|x_0\|_2 \geq \sqrt{R}/4$,*

1. *If $\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \geq R$ then $\|x - x_0\|_2\|x + x_0\|_2 \geq c_1 R$.*

2. *If $\|x - x_0\|_2\|x + x_0\|_2 \geq R$ then $\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \geq c_2 R$.*

Moreover, if $\|x_0\|_2 \leq \sqrt{R}/4$ then $\|x - x_0\|_2\|x + x_0\|_2 \geq R$ if and only if $\|x\|_2 \gtrsim \sqrt{R}$.

Proof. Assume without loss of generality that $\|x - x_0\|_2 \leq \|x + x_0\|_2$.
If $\|x - x_0\|_2 \leq \|x_0\|_2$ then

$$\|x_0\|_2 \leq 2\|x_0\|_2 - \|x - x_0\|_2 \leq \|x + x_0\|_2 \leq \|x - x_0\|_2 + 2\|x_0\|_2 \leq 3\|x_0\|_2.$$

Hence, $\|x_0\|_2 \sim \|x + x_0\|_2$, and

$$\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \sim \|x - x_0\|_2\|x + x_0\|_2.$$

Otherwise, $\|x - x_0\|_2 > \|x_0\|_2$.
If, in addition,

$$\|x_0\|_2 \geq (\|x - x_0\|_2\|x + x_0\|_2)^{1/2}/4,$$

then
$$4\|x_0\|_2 \geq (\|x - x_0\|_2\|x + x_0\|_2)^{1/2} \geq \|x_0\|_2^{1/2}\|x + x_0\|_2^{1/2},$$

and thus $\|x + x_0\|_2 \leq 16\|x_0\|_2$. Since $\|x_0\|_2 < \|x - x_0\|_2 \leq \|x + x_0\|_2$, it follows that $\|x + x_0\|_2 \sim \|x - x_0\|_2 \sim \|x_0\|_2$, and again,

$$\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \sim \|x - x_0\|_2\|x + x_0\|_2.$$

Therefore, the final case, and the only one in which there is no point-wise equivalence between $\|x - x_0\|_2\|x + x_0\|_2$ and $\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\}$, is when $\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \geq \|x_0\|_2$ and $\|x_0\|_2 \leq (\|x - x_0\|_2\|x + x_0\|_2)^{1/2}/4$. In that case, if $\|x_0\|_2 \geq \sqrt{R}/4$ then

$$\|x_0\|_2\min\{\|x - x_0\|_2, \|x + x_0\|_2\} \geq \|x_0\|_2^2 \geq R/16,$$

and

$$\|x - x_0\|_2 \|x + x_0\|_2 \geq 4\|x_0\|_2^2 \geq R/4,$$

from which the first part of the claim follows immediately.

For the second one, observe that

$$\|x\|_2^2 - 2\|x_0\|_2\|x\|_2 \leq \|x - x_0\|_2\|x + x_0\|_2 \leq \|x\|_2^2 + 2\|x\|_2\|x_0\|_2 + \|x_0\|_2^2,$$

and if $\|x_0\|_2 \leq \sqrt{R}/4$, the equivalence is evident. ∎

In view of Lemma 4.1, the way the product $\|x - x_0\|_2\|x + x_0\|_2$ relates to $\min\{\|x - x_0\|_2, \|x + x_0\|_2\}$ depends on $\|x_0\|_2$. If $\|x_0\|_2 \geq \sqrt{R}/4$, then

$$\{x \in T : \|x - x_0\|_2\|x + x_0\|_2 \leq R\} \subset \{x \in T : \min\{\|x - x_0\|_2, \|x + x_0\|_2\} \leq c_1 R/\|x_0\|_2\},$$

and if $\|x_0\|_2 \leq \sqrt{R}/4$,

$$\{x \in T : \|x - x_0\|_2\|x + x_0\|_2 \leq R\} \subset \{x \in T : \|x\|_2 \leq c_1\sqrt{R}\},$$

for a suitable absolute constant c_1.

When T is convex and centrally-symmetric, the corresponding complexity parameter - the gaussian average of $T_{+,R}(x_0) = T_{-,R}(x_0)$ is

$$E_R(x_0) \lesssim \begin{cases} \frac{\|x_0\|_2}{R} \cdot \ell(2T \cap (c_1 R/\|x_0\|_2)B_2^n) & \text{if } \|x_0\|_2 \geq \sqrt{R}, \\[2ex] \frac{1}{\sqrt{R}}\ell(2T \cap c_1\sqrt{R}B_2^n) & \text{if } \|x_0\| < \sqrt{R}. \end{cases}$$

The fixed point conditions appearing in Theorem A now become

$$r_0 = \inf\{R : E_R(x_0) \leq c_2\sqrt{N}\} \tag{4.1}$$

and

$$r_2(\gamma) = \inf\{R : E_R(x_0) \leq \gamma\sqrt{N}R\}, \tag{4.2}$$

where one selects the slightly suboptimal $\gamma = c_2/\sigma\sqrt{\log N}$. The assertion of Theorem A is that with high probability, ERM produces $\hat{x}$ for which

$$\|\hat{x} - x_0\|_2\|\hat{x} + x_0\|_2 \leq \max\{r_2(\gamma), r_0\}.$$

If $\|x_0\|_2 \geq \sqrt{R}$, the fixed-point condition (4.1) is

$$\ell(2T \cap (c_1 R/\|x_0\|_2)B_2^n) \leq c_3\left(\frac{R}{\|x_0\|_2}\right)\sqrt{N} \tag{4.3}$$

22

while (4.2) is,

$$\frac{\|x_0\|_2}{R}\ell(2T \cap (c_1 R/\|x_0\|_2)B_2^n) \le (c_4/\sigma\sqrt{\log N}) \cdot \sqrt{N}R. \qquad (4.4)$$

Recall that

$$r_N^*(Q) = \inf\{r > 0 : \ell(T \cap rB_2^n) \le Qr\sqrt{N}\},$$

and

$$s_N^*(\eta) = \inf\{s > 0 : \ell(T \cap sB_2^n) \le \eta s^2\sqrt{N}\}.$$

Therefore, it is straightforward to verify that

$$r_0 = 2\|x_0\|_2 r_N^*(c_3) \text{ and } r_2\big(c_2/(\sigma\sqrt{\log N})\big) = 2\|x_0\|_2 s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N}).$$

Setting $R = 2\|x_0\|_2 \max\{r_N^*(c_3), s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N})\}$, it remains to ensure that $\|x_0\|_2^2 \ge R$; that is,

$$2\max\{s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N}), r_N^*(c_3)\} \le \|x_0\|_2. \qquad (4.5)$$

Observe that if

$$r_N^*(c_3) \le \frac{c_3\sigma}{c_4\|x_0\|_2}\sqrt{\log N}, \qquad (4.6)$$

then $r_N^*(c_3) \le s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N})$. Indeed, applying the convexity of T, it is standard to verify that $r_N^*(Q)$ is attained and $r_N^*(Q) \le \rho$ if and only if $\ell(T \cap \rho B_2^n) \le Q\rho\sqrt{N}$ – with a similar statement for s_N^* (see, e.g., the discussion in [7]). Therefore, $s_N^*(\eta) \ge r_N^*(Q)$ if and only if $\ell(T \cap r_N^*(Q)) \ge \eta(r_N^*(Q))^2\sqrt{N}$. The latter is evident because $\ell(T \cap r_N^*(Q)) = Qr_N^*(Q)\sqrt{N}$ and recalling that $Q = c_3$ and $\eta = c_4\|x_0\|_2/\sigma\sqrt{\log N}$.

Under (4.6), an assumption which has been made in the formulation of Theorem B, (4.5) becomes $2s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N}) \le \|x_0\|_2$ and, by the definition of s_N^*, this is the case if and only if

$$\ell(T \cap \|x_0\|_2 B_2^n) \le \frac{c_4\|x_0\|_2}{\sigma\sqrt{\log N}} \cdot \|x_0\|_2^2\sqrt{N};$$

that is simply when $\|x_0\|_2 \ge v_N^*(\zeta)$ for $\zeta = c_4/\sigma\sqrt{\log N}$.

Hence, by Theorem A, combined with Lemma 4.1, it follows that with high probability,

$$\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \le 2s_N^*(c_4\|x_0\|_2/\sigma\sqrt{\log N}).$$

The other cases, when either $\|x_0\|_2$ is 'small', or when r_0 dominates r_2 are treated is a similar fashion, and are omitted. ∎

5 Minimax lower bounds

In this section we study the optimality of ERM as a phase retrieval procedure, in the minimax sense. The estimate obtained here is based on the maximal cardinality of separated subsets of the class with respect to the $L_2(\mu)$ norm.

Definition 5.1 *Let B be the unit ball in a normed space. For any subset A of the space, let $M(A, rB)$ be the maximal cardinality of a subset of A that is r-separated with respect to the norm associated with B.*

Observe that if $M(A, rB) \geq L$ there are $x_1, ..., x_L \in A$ for which the sets $x_i + (r/3)B$ are disjoint. A similar statement is true in the reverse direction.

Let F be a class of functions on (Ω, μ) and let a be distributed according to μ. For $f_0 \in F$ and a centred gaussian variable w, which has variance σ and is independent of a, consider the gaussian regression model

$$y = f_0(a) + w. \tag{5.1}$$

Any procedure that performs well in the minimax sense, must do so for any choice of $f_0 \in F$ in (5.1).

Following [7], there are two possible sources of 'statistical complexity' that influence the error rate of gaussian regression in F.

1. Firstly, that there are functions in F that, despite being far away from f_0, still satisfy $f_0(a_i) = f(a_i)$ for every $1 \leq i \leq N$, and thus are indistinguishable from f_0 on the data.

   This statistical complexity is independent of the noise, and for every $f_0 \in F$ and $\mathbb{A} = (a_i)_{i=1}^N$, it is captured by the $L_2(\mu)$ diameter of the set
   $$K(f_0, \mathbb{A}) = \{f \in F : (f(a_i))_{i=1}^N = (f_0(a_i))_{i=1}^N\},$$
   which is denoted by $d_N^*(\mathbb{A})$.

2. Secondly, that the set $(F - f_0) \cap rD = \{f - f_0 : f \in F, \|f - f_0\|_{L_2} \leq r\}$ is 'rich enough' at a scale that is proportional to its $L_2(\mu)$ diameter r.

   The richness of the set is measured using the cardinality of a maximal $L_2(\mu)$-separated set. To that end, let D be the unit ball in $L_2(\mu)$, set
   $$C(r, \theta_0) = \sup_{f_0 \in F} r \log^{1/2} M(F \cap (f_0 + \theta_0 rD), rD)$$
   and put
   $$q_N^*(\eta) = \inf \left\{ r > 0 : C(r, \theta_0) \leq \eta r^2 \sqrt{N} \right\}. \tag{5.2}$$

Theorem 5.2 *[7] For every $f_0 \in F$ let $\mathbb{P}_{f_0}^{\otimes N}$ be the probability measure that generates samples $(a_i, y_i)_{i=1}^N$ according to (5.1). For every $\theta_0 \geq 2$ there exists a constant $\theta_1 > 0$ for which*

$$\inf_{\hat{f}} \sup_{f_0 \in F} \mathbb{P}_{f_0}^{\otimes N} \left(\left\| f_0 - \hat{f} \right\|_2 \geq \max\{q_N^*(\theta_1/\sigma), (d_N^*(\mathbb{A})/4)\} \right) \geq 1/5 \qquad (5.3)$$

where $\inf_{\hat{f}}$ is the infimum over all possible estimators constructed using the given data.

Earlier versions of this minimax bound may be found in [14], [16] and [1].

To apply this general principle to the phase recovery problem, note that the regression function is $f_0(x) = \langle x_0, x \rangle^2 := f_{x_0}(x)$ for some unknown vector $x_0 \in T \subset \mathbb{R}^n$, while the estimators are $\hat{f} = \langle \hat{x}, \cdot \rangle^2$. Also, observe that for every $x_1, x_2 \in T$,

$$\|f_{x_0} - f_{x_1}\|_{L^2(\mu)}^2 = \mathbb{E}\big(\langle x_0, a\rangle^2 - \langle x_1, a\rangle^2\big)^2 = \mathbb{E}\langle x_0 - x_1, a\rangle^2 \langle x_0 + x_1, a\rangle^2$$

and therefore, one has to identify the L_2 structure of the set

$$F - f_{x_0} = \{\langle x, \cdot \rangle^2 - \langle x_0, \cdot \rangle^2 : x \in T\}.$$

To obtain the desired bound, it suffices to assume the following:

Assumption 5.1 *There exist constants C_1 and C_2 for which, for every $s, t \in \mathbb{R}^n$,*

$$C_1^2 \|s - t\|_2^2 \|s + t\|_2^2 \leq \mathbb{E}\langle s - t, a\rangle^2 \langle s + t, a\rangle^2 \leq C_2^2 \|s - t\|_2^2 \|s + t\|_2^2.$$

It is straightforward to verify that if a is an L-subgaussian vector on $\mathbb{R}^n$ that satisfies Assumption 1.1, then it automatically satisfies Assumption 5.1.

The norm $\|x_0\|_2$ plays a central role in the analysis of the rates of convergence of the ERM in phase recovery. Therefore, the minimax lower bounds presented here are not only for the entire model T but for every shell $V_0 = T \cap R_0 S^{n-1}$. A minimax lower bound over T follows by taking the supremum over all possible choices of R_0.

To apply Theorem 5.2, observe that by Assumption 5.1, for every $u, v \in T$,

$$C_1 \|u - v\|_2 \|u + v\|_2 \leq \|f_v - f_u\|_{L_2} \leq C_2 \|u - v\|_2 \|u + v\|_2.$$

Fix $R_0 > 0$ and consider $V_0 = T \cap R_0 S^{n-1}$. Clearly, for every $r > 0$ and every $x_0 \in V_0$,

$$\left\{ u \in V_0 : \|u - x_0\|_2 \|u + x_0\|_2 \leq \frac{\theta_0 r}{C_2} \right\} \subset \{ u \in V_0 : f_u \in F' \cap (f_{x_0} + \theta_0 r D) \} \tag{5.4}$$

where $F' = \{ f_u : u \in V_0 \}$.

Fix $\theta_0 > 2$ to be named later, and let θ_1 be as in Theorem 5.2. If there are $x_0 \in V_0$ and $\{x_1, ..., x_M\} \subset V_0$ that satisfy

1. $\|x_i - x_0\|_2 \|x_i + x_0\|_2 \leq \theta_0 r / C_2$,

2. for every $1 \leq i < j \leq M$, $\|x_i - x_j\|_2 \|x_i + x_j\|_2 \geq r/C_1$, and

3. $\log M > N(\theta_1 r / \sigma)^2$,

then $\sup_{f_0 \in F'} r \log^{1/2} M(F' \cap (f_0 + \theta_0 r D), rD) > \theta_1 r \sqrt{N}/\sigma$, and the best possible rate in phase recovery in V_0 is larger than r.

Fix $x_0 \in V_0$ and $r > 0$, and let $R = r/C_2$. We will present two different estimates, based on $R_0 = \|x_0\|_2$, the 'location' of x_0.

Centre of 'small norm'. Recall that $\theta_0 > 2$ and assume first that $R_0 = \|x_0\|_2 \leq \sqrt{\theta_0 R}/4$. Note that

$$V_0 \cap (\sqrt{R}/8)B_2^n \subset \left\{ u \in V_0 : \|u - x_0\|_2 \|u + x_0\|_2 \leq \frac{\theta_0 r}{C_2} \right\},$$

and thus it suffices to constructed a separated set in $V_0 \cap (\sqrt{R}/8)B_2^n$.

Set $x_1, ..., x_L$ to be a maximal $c_3\sqrt{R}$-separated subset of $V_0 \cap (\sqrt{R}/8)B_2^n$ for a constant c_3 that depends only on C_1 and C_2 and which will be specified later; thus, $L = M(V_0 \cap (\sqrt{R}/8)B_2^n, c_3\sqrt{R}B_2^n)$.

Lemma 5.3 *There is a subset $I \subset \{1, ..., L\}$ of cardinality $M \geq L/2 - 1$ for which $(x_i)_{i \in I}$ satisfies 1. and 2..*

Proof. Since $x_i \in (\sqrt{R}/8)B_2^n$ and $\|x_0\|_2 \leq \sqrt{\theta_0 R}/4$,

$$\|x_i - x_0\|_2 \|x_i + x_0\|_2 \leq ((\sqrt{\theta_0} + 1)\sqrt{R}/4)^2 \leq \theta_0 R = \theta_0 r / C_2,$$

and thus 1. is satisfied for every $1 \leq i \leq L$.

To show that 2. holds for a large subset, it suffices to find $I \subset \{1, ..., L\}$ of cardinality at least $L/2 - 1$ such that for every $i, j \in I$,

$$\|x_i - x_j\|_2 \geq c_3\sqrt{R}/2 \quad \text{and} \quad \|x_i + x_j\|_2 \geq c_3\sqrt{R}/2$$

for a well-chosen c_3.

To construct the subset, observe that if there are distinct integers $1 \leq i, j, k \leq L$ for which $\|x_i + x_j\|_2 < c_3\sqrt{R}/2$ and $\|x_i + x_k\|_2 < c_3\sqrt{R}/2$, then $\|x_j - x_k\|_2 < c_3\sqrt{R}$, which is impossible. Therefore, for every x_i there is at most a single index $j \in \{1, ..., L\}\backslash\{i\}$ satisfying that $\|x_i + x_j\|_2 < c_3\sqrt{R}/2$. With this observation the set I is constructed inductively.

Without loss of generality, assume that $I = \{1, ..., M\}$ for $M \geq L/2 - 1$. If $i \neq j$ and $1 \leq i, j \leq M$,

$$\|x_i - x_j\|_2\|x_i + x_j\|_2 \geq c_3^2 R/4 \geq 2r/C_1,$$

for the right choice of c_3, and thus $(x_i)_{i=1}^M$ satisfies 2.. ∎

Centre of 'large norm'. Next, assume that $R_0 = \|x_0\|_2 \geq \sqrt{\theta_0 R}/4$. By Lemma 4.1, there is an absolute constant $c_4 < 1/32$, for which, if $\|x_0\|_2 \geq \sqrt{\rho}/4$ and $\|x_0\|_2 \min\{\|x-x_0\|_2, \|x+x_0\|_2\} \leq c_4\rho$, then $\|x-x_0\|_2\|x+x_0\|_2 \leq \rho$. Therefore, applied to the choice $\rho = \theta_0 R$,

$$(V_0 \cap (x_0 + (c_4\theta_0 R/\|x_0\|_2)B_2^n)) \cup (V_0 \cap (-x_0 + (c_4\theta_0 R/\|x_0\|_2)B_2^n))$$
$$\subset \{u \in V_0 : \|x_0 - u\|_2\|x_0 + u\|_2 \leq \theta_0 R\},$$

and it suffices to a find a separated set in the former.

Note that if $x \in V_0 \cap (x_0 + (c_4\theta_0 R/\|x_0\|_2)B_2^n)$ then

$$\|x\|_2 \geq \|x_0\|_2 - c_4\theta_0 R/\|x_0\|_2 \geq \|x_0\|_2/2 \geq \sqrt{\theta_0 R/4}/4$$

because one can choose $c_4 \leq 1/32$, and $\|x\|_2 \leq 3\|x_0\|_2/2$. Moreover, if $x_1, x_2 \in V_0 \cap (x_0 + (c_4\theta_0 R/\|x_0\|_2)B_2^n)$, then

$$\|x_1 + x_2\|_2 \geq 2\|x_0\|_2 - 2c_4\theta_0 R/\|x_0\|_2 \geq \|x_0\|_2.$$

Applying Lemma 4.1, there is an absolute constant c_5, for which, if

$$\|x_0\|_2 \min\{\|x_i - x_j\|_2, \|x_i + x_j\|_2\} \geq \theta_0 R/4, \tag{5.5}$$

then

$$\|x_i - x_j\|_2\|x_i + x_j\|_2 \geq c_5\theta_0 R/4.$$

Hence, if $x_1, ..., x_M \in V_0 \cap (x_0 + (c_4\theta_0 R/\|x_0\|_2)B_2^n)$ is $\theta_0 R/4\|x_0\|_2$-separated, then (5.5) holds, and

$$\|x_i - x_j\|_2\|x_i + x_j\|_2 \geq c_5\theta_0 R/4 \geq r/C_1,$$

provided that θ_0 is a sufficiently large constant that depends only on C_1 and C_2.

27

Corollary 5.4 *There exist absolute constants θ_1, c_1, c_2 and c_3 that depend only on C_1 and C_2 and for which the following holds. Let $R_0 > 0$ and set $V_0 = T \cap R_0 S^{n-1}$. Define*

$$
q'(r) = \begin{cases} \log M(V_0 \cap c_1\sqrt{r}B_2^n, c_2\sqrt{r}B_2^n) & \text{if } R_0 \le c_3\sqrt{r}, \\[2mm] \sup_{x_0 \in V_0} \log M\left(V_0 \cap \left(x_0 + c_1\left(\frac{r}{R_0}\right)B_2^n\right), c_2\frac{r}{R_0}B_2^n\right) & \text{if } R_0 > c_3\sqrt{r} \end{cases}
$$

If $q'(r) \ge \theta_1 (r/\sigma)^2 N$, then the minimax rate in V_0 is larger than r.

Now, a minimax lower bound for the risk $\min\{\|\tilde{x} - x_0\|_2, \|\tilde{x} + x_0\|_2\}$ for all shells V_0 (and therefore, for T) may be derived using Lemma 4.1. To that end, let c_0 be a large enough absolute constant and

$$
C(R_0, r) = \sup_{x_0 \in T: \|x_0\|_2 = R_0} r \log^{1/2} M\big((T \cap R_0 S^{n-1}) \cap (x_0 + c_0 r B_2^n), r B_2^n\big).
$$

$$(5.6)$$

Definition 5.5 *Fix $R_0 > 0$. For every $\alpha, \beta > 0$ set*

$$
q_N^*(\alpha) = \inf\left\{r > 0 : C(R_0, r) \le \alpha r^2 \sqrt{N}\right\}
$$

and put

$$
t_N^*(\beta) = \inf\left\{r > 0 : C(R_0, r) \le \beta r^3 \sqrt{N}\right\}.
$$

Note that for any $c > 0$, if $t_N^*(c) \le 1$ then $t_N^*(c) \ge q_N^*(c)$ and $q_N^*\big(cR_0/\sigma\big) \ge t_N^*(c/\sigma)$ if and only if $q_N^*\big(cR_0/\sigma\big) \ge R_0$.

Theorem C. There exists an absolute constant c_1 for which the following holds. Let $R_0 > 0$.

1. If $R_0 \ge t_N^*\big(c_1/\sigma\big)$, then for any procedure $\tilde{x}$, there exists $x_0 \in T$ with $\|x_0\|_2 = R_0$ and for which, with probability at least $1/5$,

$$
\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2 \ge \|x_0\|_2 \, q_N^*\left(\frac{c_1 \|x_0\|_2}{\sigma}\right)
$$

and

$$
\min\{\|\tilde{x} - x_0\|_2, \|\tilde{x} + x_0\|_2\} \ge q_N^*\left(\frac{c_1 \|x_0\|_2}{\sigma}\right).
$$

2. If $R_0 \leq t_N^*(c_1/\sigma)$ then for any procedure $\tilde{x}$ there exists $x_0 \in T$ with $\|x_0\|_2 = R_0$ for which, with probability at least $1/5$,

$$\|\tilde{x} - x_0\|_2 \, \|\tilde{x} + x_0\|_2 \geq \left(t_N^* \left(\frac{c_1}{\sigma} \right) \right)^2$$

and

$$\|\tilde{x}\|_2 \, , \|\tilde{x} - x_0\|_2 \, , \|\tilde{x} + x_0\|_2 \geq t_N^* \left(\frac{c_1}{\sigma} \right).$$

Theorem C is a general minimax bound, and although it seems strange at first glance, the parameters appearing in it are very close to those used in Theorem C. Following the same path as in [7], let us show that Theorem C and Theorem C are almost sharp, under some mild structural assumptions on T.

First, recall Sudakov's inequality (see, e.g., [9]):

Theorem 5.6 *If $W \subset \mathbb{R}^n$ and $\varepsilon > 0$ then*

$$c\varepsilon \log^{1/2} M(W, \varepsilon B_2^n) \leq \ell(W),$$

where c is an absolute constant.

Fix $R_0 > 0$ set $V_0 = T \cap R_0 B_2^n$, and put

$$s_N^* = s_N^* \left(c_1 R_0 / (\sigma \sqrt{\log N}) \right) \quad \text{and} \quad v_N^* = v_N^* \left(c_1 R_0 / (\sigma \sqrt{\log N}) \right).$$

Assume that there is some $x_0 \in T$, with the following properties:

1. $\|x_0\|_2 = R$.

2. The localized sets $V_0 \cap (x_0 + s_N^* B_2^n)$ and $V_0 \cap (x_0 + v_N^* B_2^n)$ satisfy that

$$\ell(V_0 \cap (x_0 + s_N^* S^{n-1})) \sim \ell(T \cap s_N^* B_2^n)$$

and that
$$\ell(V_0 \cap (x_0 + v_N^* S^{n-1})) \sim \ell(T \cap v_N^* B_2^n),$$

which is a mild assumption on the complexity structure of T.

3. Sudakov's inequality is sharp at the scales s_N^* and v_N^*, namely,

$$s_N^* \log^{1/2} M(V_0 \cap (x_0 + c_0 s_N^* B_2^n), s_N^* B_2^n) \sim \ell(V_0 \cap (x_0 + s_N^* B_2^n)) \quad (5.7)$$

and a similar assertion holds for v_N^*.

29

In such a case, the rates of convergence obtained in Theorem B are minimax (up to the an extra $\sqrt{\log N}$ factor) thanks to Theorem C.

When Sudakov's inequality is sharp as in (5.7), we believe that ERM should be a minimax procedure in phase recovery, despite the logarithmic gap between Theorem B an Theorem C. A similar conclusion for linear regression was obtained in [7].

Sudakov's inequality is sharp in many cases - most notably, when $T = B_1^n$, but not always. It is not sharp even for standard sets like the unit ball in ℓ_p^n for $1 + (\log n)^{-1} < p < 2$.

6 Examples

Here, we will present two simple applications of the upper and lower bounds on the performance of ERM in phase recovery. Naturally, there are many other examples that follow in a similar way and that can be derived using very similar arguments. The choice of the examples has been made to illustrate the question of the optimality of Theorem A, B and C, as well as an indication of the similarities and differences between phase recovery and linear regression. Since the estimate used in these examples are rather well known, some of the details will not be presented in full.

6.1 Sparse vectors

The first example we consider represents classes with a local complexity that remains unchanged, regardless of the choice of x_0.

Let $T = W_d$ be the set of d-sparse vectors in $\mathbb{R}^n$ (for some $d \leq N/4$) – that is, vectors with at most d non-zero coordinates. Clearly, for every $R > 0$ $T_{+,R}, T_{-,R} \subset W_{2d} \cap S^{n-1}$. Also, for any $x_0 \in T$ and any $I \subset \{1, ..., n\}$ of cardinality d that is disjoint of $\mathrm{supp}(x_0)$,

$$(1/\sqrt{2})S^I \subset \left\{ \frac{(x - x_0)_{i \in I}}{\|x - x_0\|_2} : x \in W_d \right\}, \left\{ \frac{(x + x_0)_{i \in I}}{\|x + x_0\|_2} : x \in W_d \right\},$$

where S^I is the unit sphere supported on the coordinates I.

With this observation, a straightforward argument shows that,

$$\ell(T_{+,R}), \ell(T_{-,R}) \sim \sqrt{d \log(en/d)}.$$

Applying Theorem A, if follows that for $N \gtrsim d \log(en/d)$, with probability

at least $1 - 2\exp(-c(d\log(en/d))) - N^{-\beta+1}$, ERM produces $\hat{x}$ that satisfies

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \lesssim_{\kappa_0, L, \beta} \sigma \sqrt{\frac{d\log(en/d)}{N}} \sqrt{\log N} = (*). \qquad (6.1)$$

Moreover, with the same probability estimate, if $\|x_0\|_2^2 \gtrsim (*)$ then by Lemma 4.1,

$$\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \lesssim_{\kappa_0, L, \beta} \frac{\sigma}{\|x_0\|_2} \sqrt{\frac{d\log(en/d)}{N}} \sqrt{\log N} \qquad (6.2)$$

and if $\|x_0\|_2^2 \lesssim (*)$ then

$$\|\hat{x}\|_2^2, \|\hat{x} - x_0\|_2^2, \|\hat{x} + x_0\|_2^2 \lesssim_{\kappa_0, L} \sigma \sqrt{\frac{d\log(en/d)}{N}} \sqrt{\log N}. \qquad (6.3)$$

In particular, when $\|x_0\|_2$ is of the order of a constant, the rate of convergence in (6.1) and (6.2) is identical to the one obtained in [7] in the linear regression (up to a $\sqrt{\log N}$ term). In the latter, ERM achieves the minimax rate (with the same probability estimate) of

$$\|\hat{x} - x_0\|_2 \lesssim_L \sigma \sqrt{\frac{d\log(en/d)}{N}}.$$

Otherwise, when $\|x_0\|_2$ is large, the rate of convergence of the ERM in (6.2) is actually better than in linear regression, but, when $\|x_0\|_2$ is small, it is worse - deteriorating to the square root of the rate in linear regression (up to logarithmic terms).

When the noise level σ tends to zero, the rates of convergence in linear regression and phase recovery tend to zero as well. In particular, exact reconstruction happens – that is $\hat{x} = x_0$ in linear regression and $\hat{x} = x_0$ or $\hat{x} = -x_0$ in phase recovery – when $N \gtrsim d\log(en/d)$.

For the lower bound, it is well known that $\log^{1/2} M(W_d \cap c_0 r B_2^n, r B_2^n) \sim \sqrt{d\log(en/d)}$ for every $r > 0$ (and $c_0 \geq 2$). Combined with the results of the previous section, this suffices to show that the rate obtained in Theorem A is the minimax one (up to a $\sqrt{\log N}$ term in the "large noise" regime) and that the ERM is a minimax procedure for the phase retrieval problem when the signal x_0 is known to be d-sparse and $N \gtrsim d\log(en/d)$.

6.2   The unit ball of ℓ_1^n

Consider the set $T = B_1^n$, the unit ball of ℓ_1^n. Being convex and centrally symmetric, it is a natural example of a set with changing 'local complexity' –

which becomes very large when x_0 is close to 0. Moreover, it is an example in which one may obtain sharp estimates on $\ell(B_1^n \cap rB_2^n)$ at every scale. Indeed, one may show (see, for example, [5]) that

$$\ell\big(B_1^n \cap rB_2^n\big) \sim \begin{cases} \sqrt{\log(enr^2)} & \text{if } r^2n \geq 1 \\[2mm] r\sqrt{n} & \text{otherwise.} \end{cases}$$

It follows that for B_1^n, one has

$$r_N^*(Q) \quad \begin{cases} \sim \left(\frac{1}{Q^2N} \log\left(\frac{n}{Q^2N}\right)\right)^{1/2} & \text{if } n \geq C_0Q^2N \\[3mm] \lesssim \frac{1}{N} & \text{if } C_1Q^2N \leq n \leq C_0Q^2N \\[3mm] = 0 & \text{if } n \leq C_1Q^2N. \end{cases}$$

where C_0 and C_1 are absolute constants. The only range in which this estimate is not sharp is when $n \sim Q^2N$, because in that range, $r_N^*(Q)$ decays to zero very quickly. A more accurate estimate on $\ell(B_1^n \cap rB_2^n)$ can be performed when $n \sim Q^2N$ (see [8]), but since it is not our main interest, we will not pursue it further, and only consider the cases $n \leq C_1Q^2N$ and $n \geq C_0Q^2N$.

A straightforward computation shows that the two other fixed points satisfy:

$$s_N^*(\eta) \sim \begin{cases} \left(\frac{1}{\eta^2N} \log\left(\frac{n^2}{\eta^2N}\right)\right)^{1/4} & \text{if } n \geq \eta\sqrt{N} \\[3mm] \sqrt{\frac{n}{\eta^2N}} & \text{if } n \leq \eta\sqrt{N} \end{cases}$$

and

$$v_N^*(\zeta) \sim \begin{cases} \left(\frac{1}{\zeta^2N} \log\left(\frac{n^3}{\zeta^2N}\right)\right)^{1/6} & \text{if } n \geq \zeta^{2/3}N^{1/3} \\[3mm] \left(\frac{n}{\zeta^2N}\right)^{1/4} & \text{if } n \leq \zeta^{2/3}N^{1/3}. \end{cases}$$

The estimates above will be used to derive rates of convergence for the ERM $\hat{x}$ (for the squared loss) of the form

$$\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \leq rate.$$

Upper bounds on the rate of convergence *rate* follow from Theorem B, and hold with high probability as stated in there. For the sake of brevity, we

will not present the probability estimates, but those can be easily derived from Theorem B.

Thanks to Theorem B, obtaining upper bounds on *rate* involves the study of several different regimes, depending on $\|x_0\|_2$, the noise level σ and the way the number of observations N compares with the dimension n.

The noise-free case: $\sigma = 0$. In this case, *rate* is upper bounded by $r_N^*(Q)$, for Q that is an absolute constant. In particular, when $n \geq C_0 Q^2 N$, the rate is less than
$$\left(N^{-1} \log\left(n/N\right)\right)^{1/2}.$$

At this point, it is natural to wonder whether there is a procedure that outperforms ERM in the noise-free case. The minimax lower bound $d_N^*(\mathbb{A})$ in Theorem 5.2 may be used to address this question, as no algorithm can do better than $d_N^*(\mathbb{A})/4$, with probability greater than $1/5$.

In the phase recovery problem and using the notation of section 5, one has

$$
\begin{aligned}
d_N^*(\mathbb{A}) &= \sup\left\{\|f_x - f_{x_0}\|_2 : x \in B_1^n, \ f_x(a_i) = f_{x_0}(a_i) , i = 1, \ldots, N\right\} \\
&\sim \sup\left\{\|x - x_0\|_2 \|x + x_0\|_2 : \ x \in B_1^n, \ |\langle a_i, x\rangle| = |\langle a_i, x_0\rangle|, \ i = 1, \ldots, N\right\} \\
&\gtrsim \inf_{L:\mathbb{R}^n \to \mathbb{R}^N} \sup\left\{\|x - x_0\|_2 \|x + x_0\|_2 : \ x \in B_1^n, \ L(x) = L(x_0)\right\}
\end{aligned}
$$

with an infimum taken over all linear operators $L : \mathbb{R}^n \to \mathbb{R}^N$.

By Lemma 4.1, for $x_0 = (1/2, 0, \ldots, 0) \in B_1^n$ (in fact, any vector x_0 in B_1^n for which $\|x_0\|_2$ is a positive constant smaller than $1/2$ would do)

$$
\begin{aligned}
d_N^*(\mathbb{A}) &\gtrsim \inf_{L:\mathbb{R}^n \to \mathbb{R}^N} \sup_{x \in B_1^n \cap (\ker L - x_0)} \min\{\|x - x_0\|_2 , \|x + x_0\|_2\} \\
&\gtrsim \inf_{L:\mathbb{R}^n \to \mathbb{R}^N} \sup_{x,y \in B_1^n \cap \ker L} \|x - y\|_2 = c_N(B_1^n)
\end{aligned}
$$

which is the Gelfand N-width of B_1^n. By a result due to Garnaev and Gluskin (see [3]),

$$
c_N(B_1^n) \sim
\begin{cases}
\min\left\{1, \sqrt{\frac{1}{N} \log\left(\frac{en}{N}\right)}\right\} & \text{if } N \leq n \\
\\
0 & \text{otherwise.}
\end{cases}
$$

which is of the same order as r_N^* (except when $n \sim N$, which is not treated here). Therefore, no algorithm can outperform ERM and ERM is a minimax procedure in this case.

Note that when $n \le c_1 Q^2 N$, exact reconstruction of x_0 or $-x_0$ is possible and it can happens only in that case (i.e. $\sigma = 0$ and $n \le c_1 Q^2 N$) because of the minimax lower bound provided by $d_N^*(\mathbb{A})$.

The noisy case: $\sigma > 0$. According to Theorem B, the rate of convergence *rate* depends on $r_N^* = r_N^*(Q)$ for some absolute constant Q, $s_N^* = s_N^*(\eta)$ for $\eta = c_1 \|x_0\|_2 / (\sigma\sqrt{\log N})$ and on $v_N^* = v_N^*(\zeta)$ for $\zeta = c_1 / (\sigma\sqrt{\log N})$. The outcome of Theorem B is presented in Figure 1.

rate $\lesssim$	$\sigma / \|x_0\|_2 \le c_0 r_N^* / \sqrt{\log N}$	$\sigma / \|x_0\|_2 \ge c_0 r_N^* / \sqrt{\log N}$
$\|x_0\|_2 \le v_N^*$	r_N^*	s_N^*
$\|x_0\|_2 \ge v_N^*$	r_N^*	v_N^*

Figure 1: High probability bounds on the rate of convergence of the ERM $\hat{x}$ for the square loss in phase recovery: $\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \le rate$.

As the proof of all the estimates is similar, a detailed analysis is only presented when $\zeta^{2/3} N^{1/3} \le \eta\sqrt{N} \le C_1 Q^2 N$, which is equivalent to

$$\left(\frac{\sigma^2 \log N}{c_1 N}\right)^{1/6} \le \|x_0\|_2 \le \frac{c_1 Q^2 \sigma\sqrt{N \log N}}{c_1}.$$

The upper bounds on *rate* change according to the way the number of observations N scales relative to n:

1. $n \ge C_0 Q^2 N$. In this situation, $r_N^* \sim \left(\log(n/N)/N\right)^{1/2}$. Therefore, if $\sigma / \|x_0\|_2 \lesssim \sqrt{\log(n/N)/(N \log N)}$ then $rate \le \left(\log(n/N)/N\right)^{1/2}$, and if $\sigma / \|x_0\|_2 \gtrsim \sqrt{\log(n/N)/(N \log N)}$,

$$rate \le \begin{cases} \left(\frac{\sigma^2 \log N}{\|x_0\|_2^2 N} \log\left(\frac{\sigma^2 n^2}{\|x_0\|_2^2 N}\right)\right)^{1/4} & \text{if } \|x_0\|_2 \ge \left(\frac{\sigma^2 \log N}{N} \log\left(\frac{\sigma^2 n^3}{N}\right)\right)^{1/6} \\ \left(\frac{\sigma^2 \log N}{N} \log\left(\frac{\sigma^2 n^3}{N}\right)\right)^{1/6} & \text{otherwise.} \end{cases}$$
(6.4)

2. $c_1 \|x_0\|_2 / (\sigma\sqrt{\log N})\sqrt{N} \le n \le C_1 Q^2 N$. In that case $r_N^* = 0$. In particular $\sigma / \|x_0\|_2 > c_0 r_N^* / \sqrt{\log N}$ and therefore, the rate is upper bounded as in (6.4).

3. $\left(c_1 / (\sigma\sqrt{\log N})\right)^{2/3} N^{1/3} \le n \le c_1 \|x_0\|_2 / (\sigma\sqrt{\log N})\sqrt{N}$. Again, in this case, $r_N^* = 0$. Therefore, one is in the situation of the small

34

signal-to-noise ratio and

$$rate \leq \begin{cases} \frac{\sigma}{\|x_0\|_2} \sqrt{\frac{n \log N}{N}} & \text{if } \|x_0\|_2 \geq \left(\frac{\sigma^2 \log N}{N} \log \left(\frac{\sigma^2 n^3}{N} \right) \right)^{1/6} \\[3mm] \left(\frac{\sigma^2 \log N}{N} \log \left(\frac{\sigma^2 n^3}{N} \right) \right)^{1/6} & \text{otherwise.} \end{cases}$$

4. $n \leq \left(c_1/(\sigma \sqrt{\log N}) \right)^{2/3} N^{1/3}$. Once again, $r_N^* = 0$, and

$$rate \leq \begin{cases} \frac{\sigma}{\|x_0\|_2} \sqrt{\frac{n \log N}{N}} & \text{if } \|x_0\|_2 \geq \left(\sigma \sqrt{\frac{n \log N}{N}} \right)^{1/2} \\[3mm] \left(\sigma \sqrt{\frac{n \log N}{N}} \right)^{1/2} & \text{otherwise.} \end{cases}$$

One may ask whether these estimates are optimal in the minimax sense, or perhaps there is another procedure that can outperform ERM. It appears that (up to an extra $\sqrt{\log N}$ factor), ERM is indeed optimal.

To see that, it is enough to apply Theorem C and verify that Sudakov's inequality is sharp in the following sense: (see the discussion following Theorem C): that if $\|x_0\|_1 \leq 1/2$, then for every $\varepsilon < 1/4$

$$\varepsilon \log^{1/2} M(B_1^n \cap (x_0 + c_0 \varepsilon B_2^n), \varepsilon B_2^n) \sim \ell \left(B_1^n \cap \varepsilon B_2^n \right).$$

This fact is relatively straightforward to verify (see, e.g., Example 2 in [10]).

Therefore, up to the extra $\sqrt{\log N}$ factor, which we believe is parasitic, ERM is a minimax phase-recovery procedure in B_1^n.

References

[1] L. Birgé, *Approximation dans les espaces métriques et théorie de l'estimation*, Z. Wahrsch. Verw. Gebiete (2) vol. 65, 1983

[2] Y. C. Eldar, S. Mendelson, *Phase Retrieval: Stability and Recovery Guarantees*, preprint.

[3] A. Yu. Garnaev and E. D. Gluskin. The widths of a Euclidean ball. *Dokl. Akad. Nauk SSSR*, 277(5):1048–1052, 1984.

[4] E. Giné, V. de la Peña, *Decoupling: From Dependence to Independence*, Springer-Verlag, 1999.

[5] Y.Gordon, A.E. Litvak, S. Mendelson and A. Pajor, *Gaussian averages of interpolated bodies and applications to approximate reconstruction*, J. Approx. Theory, 149 (1), 2007

[6] Vladimir Koltchinskii. *Oracle inequalities in empirical risk minimization and sparse recovery problems*, volume 2033 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].

[7] G. Lecué, S. Mendelson, *Learning subgaussian classes : Upper and minimax bounds*, preprint, available at http://arxiv.org/pdf/1305.4825.pdf.

[8] G. Lecué, S. Mendelson, *On the optimality of the empirical risk minimization procedure for the convex aggregation problem*, Annales de l'Institut Henri Poincaré Probabilités et Statistiques, 49 (1), 2013.

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

[10] S. Mendelson, *On the geometry of subgaussian coordinate projections*, preprint.

[11] V.D. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Mathematics 1200, Springer, 1986.

[12] M. Schmuckenschlaeger, *Bernstein inequalities for a class of random variables*, Proc. Amer. Math. Soc. Proceedings of the American Mathematical Society (4), vol. 117, 1993.

[13] Alain Pajor and Nicole Tomczak-Jaegermann. Subspaces of small codimension of finite-dimensional Banach spaces. *Proc. Amer. Math. Soc.*, 97(4):637–642, 1986.

[14] Alexandre B. Tsybakov. *Introduction to nonparametric estimation*. Springer Series in Statistics. Springer, New York, 2009. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

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

[16] Yuhong Yang and Andrew Barron. Information-theoretic determination of minimax rates of convergence. *Ann. Statist.*, 27(5):1564–1599, 1999.