

Minimax rate of convergence and the performance of empirical risk minimization in phase retrieval

Guillaume Lecué* Shahar Mendelson†

Abstract

We study the performance of Empirical Risk Minimization in both noisy and noiseless 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 subgaussian noise w and a fixed but unknown x_0 that belongs to a given $T \subset \mathbb{R}^n$.

We show that ERM performed in T produces \hat{x} whose Euclidean distance to either x_0 or $-x_0$ depends on the gaussian mean-width of T 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 sharp minimax lower bound for the phase retrieval problem. As examples, we study the class of d -sparse vectors in \mathbb{R}^n and the unit ball in ℓ_1^n .

Keywords: Empirical process; minimax theory; phase recovery.

AMS MSC 2010: 62H12.

Submitted to EJP on May 13, 2014, final version accepted on May 27, 2015.

Supersedes arXiv:1311.5024.

1 Introduction

There are many areas of engineering in which only the intensity of signals can be observed: the phase is either difficult to measure or it is simply lost in the measurement process. For example, phase-less data is the type of information one observes in X-ray diffraction images – like the ones that led to the discovery of the double-helix structure (cf. [33]). Other examples of natural problems leading to data that does not contain the phase can be found in [4, 5, 10, 27].

In phase retrieval, one attempts to identify a vector x_0 using noisy, quadratic measurements of x_0 . The given data is a random sample of cardinality N , $(a_i, y_i)_{i=1}^N$, for measurement vectors a_i and

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

*CNRS, CMAP, Ecole Polytechnique, 91120 Palaiseau, France.

E-mail: guillaume.lecue@cmmap.polytechnique.fr

†Department of Mathematics, Technion, I.I.T, Haifa 32000, Israel. Supported by the Mathematical Sciences Institute – The Australian National University and by the Israel Science Foundation.

E-mail: shahar@tx.technion.ac.il

where $(w_i)_{i=1}^N$ is the noise vector. The hope is that although the phase of the data $\langle a_i, x_0 \rangle$ is not measured, it is possible to estimate x_0 up to a phase. To simplify our exposition, we only consider the ‘real’ version of the phase retrieval problem: x_0, a_1, \dots, a_n are assumed to be vectors in \mathbb{R}^n , and in which case, the goal is to identify a point that is close either to x_0 or to $-x_0$.

The phase retrieval problem is usually considered under structural assumptions on the set T from which x_0 is taken – most notably, x_0 is assumed to be a sparse vector. In that context, greedy algorithms have been introduced in [9, 14], but with no theoretical guarantee of success, and with tuning parameter issues. Later, efficient algorithms that reconstruct x_0 (up to a phase) from the noiseless data $(|\langle a_i, x_0 \rangle|^2)_{i=1}^N$ were suggested: semidefinite programs obtained by convex relaxations such as *PhaseLift* [3, 6] and *PhaseCut* [32, 10] can be shown to perform well both from a theoretical point of view and from a practical one (see, also, [31]). Techniques from matrix completion have also been used ([2, 26]), and recently, the “small ball method” (see [15, 18, 21, 22, 23]) has been applied to the sparse phase retrieval setup in [29].

In the present work, we consider a general set $T \subset \mathbb{R}^n$ rather than studying sets that are associated with sparse vectors, and assume that $x_0 \in T$. Our aim is to investigate phase retrieval from a theoretical point of view, relative to a well behaved, random sampling method. We develop a common analysis for both noisy and noiseless measurements; for example, our general results imply exact recovery (up to the sign) in the noiseless case and error rates that are minimax optimal (up to logarithmic terms) in the noisy case under sparsity constraints. There are very few known results in the noisy setup, and in what follows we will compare ours to the ones from [8].

To formulate the problem, we need the following definitions.

Definition 1.1. Let μ be a measure on \mathbb{R}^n and set a to be a random vector distributed according to μ . The measure μ (or the random vector a) is isotropic if for every $x \in \mathbb{R}^n$, $\mathbb{E}\langle x, a \rangle^2 = \|x\|_2^2$. It is L -subgaussian for some $L \geq 1$ if for every $u \geq 1$, $\Pr(|\langle x, a \rangle| \geq Lu\|x\|_{L_2}) \leq 2 \exp(-u^2/2)$.

For a real-valued random variable w , the ψ_2 norm of w is defined by $\|w\|_{\psi_2} = \inf\{c > 0 : \mathbb{E} \exp(w^2/c^2) \leq 2\}$.

Given a set $T \subset \mathbb{R}^n$ and a fixed, but unknown $x_0 \in T$, y_i are 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 random 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 Euclidean 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 we shall use is *Empirical Risk Minimization* (ERM), which produces \hat{x} that minimizes the empirical risk in T : let ℓ_x be the squared loss associated with $f_x(a) = \langle x, a \rangle^2$; thus,

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

Set

$$\hat{x} \in \operatorname{argmin}_{x \in T} P_N \ell_x \text{ where } P_N \ell_x = \frac{1}{N} \sum_{i=1}^N (\langle a_i, x \rangle^2 - y_i)^2$$

and note that for every $x \in \mathbb{R}^n$,

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

Both components are difficult to handle directly, even when the underlying measure is subgaussian, because the two involve high powers of $\langle \cdot, a_i \rangle$: an effective power of 4 in the first component and of 3 in the second one. In contrast, in the standard linear regression problem, $\ell_x = (\langle a, x \rangle + w)^2$, and the corresponding components have powers of 2 and 1 respectively, resulting in a much simpler analysis.

Rather than trying to employ the concentration of empirical means around the actual ones, which is not sufficiently strong in this case, one uses a combination of a “small-ball estimate” for the empirical process $\left(N^{-1} \sum_{i=1}^N \langle x - x_0, a_i \rangle^2 \langle x + x_0, a_i \rangle^2\right)_{x \in T}$, and a more standard deviation argument for $\left(N^{-1} \sum_{i=1}^N w_i \langle x - x_0, a_i \rangle \langle x + x_0, a_i \rangle\right)_{x \in T}$ (see Section 3 and the formulation of Theorem A and Theorem B).

Taking the same path as in [8], we assume that linear forms satisfy the following.

Assumption 1.1. There is a constant $\kappa_0 > 0$ for which, 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 (see, for example, the discussion in [8]).

It is not surprising that the error rate of ERM in a phase retrieval problem 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 1.2. Let $G = (g_1, \dots, g_n)$ be the standard gaussian vector in \mathbb{R}^n . For $T \subset \mathbb{R}^n$, set

$$\ell(T) = \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n g_i t_i \right|.$$

We will consider two different types of normalized sets: firstly, following [8], a ‘global approach’ – and the reason for this name is that the resulting complexity parameter does not depend on the signal x_0 . This approach leads to the study of the sets

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

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

As will be explained later, there are natural examples of sets for which the global approach is not optimal. To handle such cases, our main result is based on a ‘local approach’, in which the normalized sets depend on the signal x_0 :

$$T_{-,R}(x_0) = \left\{ \frac{t - x_0}{\|t - x_0\|_2} : t \in T, R < \|t - x_0\|_2 \|t + x_0\|_2 \right\},$$

$$T_{+,R}(x_0) = \left\{ \frac{t + x_0}{\|t + x_0\|_2} : t \in T, R < \|t - x_0\|_2 \|t + x_0\|_2 \right\}.$$

These normalized sets play a significant role in the analysis of ERM. Indeed, setting

$\mathcal{L}_x = \ell_x - \ell_{x_0}$, the excess loss function associated with $x \in T$, 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, $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq R$.

This normalization allows one to study ‘relative fluctuations’ of $P_N \mathcal{L}_x$, in particular, the way the fluctuations scale with $\|x - x_0\|_2 \|x + x_0\|_2$.

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 [8] – but the outcome is far from satisfactory. Roughly put, there are two types of subsets of \mathbb{R}^n one is interested in, and that appear in applications. The first 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 . When the ‘local’ sets are not much smaller than $T_{-,R}$ and $T_{+,R}$, the ‘global’ approach 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 .

A typical example is the 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 relatively small in its own right, and the ‘global’ approach suffices.

In contrast, there are simple sets that have diverse local complexities, 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 $\ell_1^n = (\mathbb{R}^n, \|\cdot\|_1)$. 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 $e_1 = (1, 0, \dots, 0)$, $T_{+,R}(e_1)$ and $T_{-,R}(e_1)$ consist of vectors that are well approximated by sparse vectors (whose support depends on R), and thus are rather small subsets of the sphere.

This situation 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 local analysis and additional information on $\|x_0\|_2$.

In fact, we will show that this is true in general: the error rate of ERM *does depend* on $\|x_0\|_2$ 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 an upper estimate on the error rate of ERM using the global approach. Just as linear regression, the rates are determined by solutions of a fixed point equations

$$r_1(\gamma) = \inf \left\{ r > 0 : E_r \leq \gamma \sqrt{Nr} \right\}$$

and

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

for constants γ and Q that will be specified later.

Theorem A.[Global approach] *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 $T \subset \mathbb{R}^n$, set a to be an isotropic, L -subgaussian random vector on \mathbb{R}^n , put $x_0 \in T$ and let w and y be as in (1.1). Let $r_2^* = \max\{r_0(c_1), r_1(c_2/\sigma\sqrt{\log N})\}$. If \hat{x} is produced by ERM using the sample $(a_i, y_i)_{i=1}^N$, then with probability at least*

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

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq r_2^*.$$

When the subgaussian assumption on w is replaced by an L_∞ one, the term $\sigma\sqrt{\log N}$ may be replaced by $\|w\|_\infty$.

The upper estimate of $\max\{r_0, r_1\}$ in Theorem A represents two ranges of noise. It follows from the definition of the fixed points that r_0 is dominant if $\sigma \lesssim r_0/\sqrt{\log N}$, and if σ is larger, r_1 is dominant. As explained in [16] 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 well-separated 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 noise w , and the dominating term becomes r_1 . Of course, there are cases in which $r_0 = 0$ for N sufficiently large. This is precisely when exact recovery is possible in a noise-free problem. And, in such cases, the error of ERM tends to zero with σ .

Note that if T has a well behaved ‘global complexity’, and since $E_R \leq E$ for every $R > 0$, it follows that when $N \gtrsim E^2$, $r_0 = 0$ and that $r_1(\gamma) \leq E/(\gamma\sqrt{N})$. Therefore, on the event from Theorem A,

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq c(L)\sigma \frac{E}{\sqrt{N}} \sqrt{\log N}.$$

This estimate suffices for many applications. For example, when T is the set of d -sparse vectors, one may show (see, e.g. [8]) that

$$E \leq c\sqrt{d \log(en/d)}$$

for an absolute constant c .

Hence, by Theorem A, when $N \geq c_1(L)d \log(en/d)$, with high probability,

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq c_2(L)\sigma \sqrt{\frac{d \log(en/d)}{N}} \sqrt{\log N},$$

and, in particular, in the free-noise case (that is, when $\sigma = 0$), ERM results in exact reconstruction, meaning that either $\hat{x} = x_0$ or $-\hat{x} = x_0$.

The proof of this observation and that it 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 [8] in three ways.

- 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 [8], it scaled like $c/N^{1/4}$ for very large values of N .
- The estimate scales linearly in σ while the rate obtained in [8] does not decay with σ for $\sigma \leq 1$.

- The probability estimate has been improved, though it is still likely to be suboptimal.

The main motivation in [8] was dealing with phase retrieval for sparse classes, a goal for which Theorem A with its global approach is well suited. However, when considering the question of more general classes, the global approach is simply too coarse. We therefore turn to the ‘local’ approach, which requires slightly modified complexity parameters.

Definition 1.3. *Let*

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

and

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

The parameters r_N^* and s_N^* have been used in [16] to obtain a sharp estimate on the performance of ERM for linear regression in an arbitrary convex set, and relative to L -subgaussian measurements. The added structure in phase retrieval requires an additional parameter:

$$v_N^*(\zeta) = \inf \left\{ v > 0 : \ell(T \cap vB_2^n) \leq \zeta v^3 \sqrt{N} \right\}.$$

Our main result is the following:

Theorem B.[Local approach] 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 $1 - 2 \exp(-c_3 N \eta^2 (s_N^*(\eta))^2) - 2N^{-\beta+1}$,

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

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

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

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

Theorem B follows from a ‘local’ version of Theorem A, a more transparent description of the localized sets $T_{-,R}(x_0)$ and $T_{+,R}(x_0)$, together with a result connecting $\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2$ and $\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\}$ as a function of $\|x_0\|_2$ (see Lemma 4.2 below).

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 $\sigma \rightarrow 0$ or when $N \rightarrow \infty$.

Also, and using the same argument,

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

Thus, if $x_0 \neq 0$ and for every $\sigma > 0$, the condition $(\sigma/\|x_0\|_2) \geq c_0 r_N(Q)/\sqrt{\log N}$ is satisfied when N is large enough, and the first part of Theorem B describes the ‘long term’ behaviour of ERM.

In the typical situation, 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 improve the results from [8], it is natural to ask whether they 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. Then, we end the article with two examples of classes that are of interest in phase retrieval: the set of 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 varying ‘local complexity’.

2 Preliminaries

Throughout this article, absolute constants are denoted by C, c, c_1, \dots 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 norm endowed on \mathbb{R}^n , and set $\ell_p^n = (\mathbb{R}^n, \|\cdot\|_p)$. B_p^n denotes the unit ball in ℓ_p^n and S^{n-1} is the Euclidean sphere in \mathbb{R}^n .

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 [19, 30].

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 $t \geq 1$.

Note that a random variable X is L -subgaussian if it has a bounded ψ_2 norm and $\|X\|_{\psi_2} \leq L\|X\|_{L_2}$. Moreover, 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, \dots, X_N

are independent copies of X and $1 \leq \alpha \leq 2$, then

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

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., [30]).

Theorem 2.1. *There exists an absolute constant c_0 for which the following holds. If X_1, \dots, X_N are independent copies of a ψ_1 random variable X , then for every $t > 0$,*

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

An important example of a probability space is the discrete space $\Omega = \{1, \dots, N\}$, 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 proof 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 $(|v_i|)_{i=1}^N$. The next straightforward observation shows that the ψ_α^N norm captures information on the coordinates of $(v_i^*)_{i=1}^N$.

Lemma 2.2. *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 $\Omega = \{1, \dots, N\}$. 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

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

It is enough to show that \mathcal{B} is bounded from above 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 $1/\beta^2 < 1$,

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

Therefore,

$$\begin{aligned} \sum_{i=1}^N \exp(v_i^2/B^2) &= \sum_{i=1}^N \exp((v_i^*)^2/B^2) \leq \sum_{i=1}^N \exp(\beta^{-2} \log(eN/i)) \\ &\leq \sum_{i=1}^N \left(\frac{eN}{i}\right)^{1/\beta^2} \leq (eN)^{1/\beta^2} \cdot \frac{N^{1-1/\beta^2}}{1-1/\beta^2} \leq \frac{Ne^{1/\beta^2}}{1-1/\beta^2} < 2N, \end{aligned}$$

provided that $\beta \geq c_1$. Hence, by the definition of the ψ_2^N norm, if $\beta \geq c_1$, $\|v\|_{\psi_2^N} < B$ which is impossible. \square

2.1 Empirical and Subgaussian processes

The sampling method we use is with respect to an isotropic and L -subgaussian measure and the noise w has a bounded ψ_2 -norm. Thus, it is natural to study some properties of subgaussian processes.

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., [25]).

Theorem 2.3. [20] *For every $L \geq 1$ there exist constants c_1 and c_2 that depend only on L 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 $I \subset \{1, \dots, N\}$,*

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

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

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

It follows from Theorem 2.3 that if $t \in T$ and $|I| \leq j_T$,

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

and if $|I| \geq j_T$,

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

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 2.4. [24] *Let $T_1, T_2 \subset \mathbb{R}^n$ and put $k^* = \min\{k^*(T_1), k^*(T_2)\}$. For $u \geq 8$, with*

probability at least $1 - 2 \exp(-c_0(L)u^2k^*)$,

$$\sup_{t \in T_1, s \in T_2} \left| \sum_{i=1}^N (\langle a_i, t \rangle \langle a_i, s \rangle - \mathbb{E} \langle a, t \rangle \langle a, s \rangle) \right| \leq c_1(L) \left(u^2 \ell(T_1) \ell(T_2) + u \sqrt{N} (d(T_1) \ell(T_2) + d(T_2) \ell(T_1)) \right).$$

Remark 2.5. Let $(\varepsilon_i)_{i=1}^N$ be independent, symmetric, $\{-1, 1\}$ -valued random variables. It follows from the results in [20] that with the same probability estimate as in Theorem 2.4 and relative to the product measure $(\varepsilon \otimes X)^N$,

$$\sup_{t \in T_1, s \in T_2} \left| \sum_{i=1}^N \varepsilon_i \langle a_i, t \rangle \langle a_i, s \rangle \right| \leq c_1(L) \left(u^2 \ell(T_1) \ell(T_2) + u \sqrt{N} (d(T_1) \ell(T_2) + d(T_2) \ell(T_1)) \right).$$

Assume that $(k^*(T_1))^{1/2} = \ell(T_1)/d(T_1) \geq \ell(T_2)/d(T_2) = (k^*(T_2))^{1/2}$. Theorem 2.4 and Remark 2.5 show that with probability at least $1 - 2 \exp(-c_1 u^2 k_*(T_2))$,

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

and

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

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

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

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 information on $\|(\langle a_i, t \rangle)_{i=1}^N\|_{L_p^N}$ and $\|(\langle a_i, t \rangle)_{i=1}^N\|_{\psi_\alpha^N}$.

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, $\|v\|_{L_p^m} \lesssim_p \|v\|_{\psi_\alpha^m}$. Moreover, 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. [12]), describes such regularity properties when $p = \alpha = 1$.

Lemma 2.6. *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 $I \subset \{1, \dots, m\}$ of cardinality at least $c_1 m$, and for every $i \in I$, $|v_i| \geq c_2 \|v\|_{L_1^m}$.*

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,

$$\begin{aligned} (m/2) \|v\|_{L_1^m} &\leq \sum_{\ell \geq j+1} v_\ell^* = \sum_{\ell \in I_\eta} |v_\ell| + \sum_{\ell \in I_\eta^c} |v_\ell| \leq v_{j+1}^* |I_\eta| + \eta \|v\|_{L_1^m} |I_\eta^c| \\ &\lesssim \|v\|_{L_1^m} |I| \left(\beta \log(e\beta) \frac{|I_\eta|}{|I|} + \eta \frac{|I_\eta^c|}{|I|} \right). \end{aligned}$$

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. \square

Next, let us turn to decomposition results for the vectors $(\langle a_i, t \rangle)_{i=1}^N$. 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 2.7. *For every $L > 1$ there exist constants c_1 and c_2 that depend only on L 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 $J_t \subset \{1, \dots, N\}$ 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 2.3 and the characterization of the ψ_2^N norm of a vector using the monotone rearrangement of its coordinates (Lemma 2.2),

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

To complete 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 2.8. *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, and

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

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

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

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.

$$\frac{\kappa_0}{2} \|s+t\|_2 \|s-t\|_2 \leq \frac{1}{N} \sum_{i=1}^N |\langle s+t, a_i \rangle \langle s-t, a_i \rangle| \lesssim_L \|s+t\|_2 \|s-t\|_2. \quad (2.4)$$

From here on we denote by $\Omega_{1,R}$ the event on which Corollary 2.8 holds for the sets $T_{+,R}$ and $T_{-,R}$ and for samples of cardinality $N \gtrsim_L E_R^2 / \kappa_0^2$.

Lemma 2.9. *There exist constants c_0 and c_1 that depend only on L and c_1, κ_0 and κ_1 for which the following holds. If $N \geq c_0 E_R^2 / \kappa_0^2$, then for $(a_i)_{i=1}^N \in \Omega_{1,R}$ and every $s, t \in T$ for which $\|s-t\|_2 \|s+t\|_2 \geq R$, there is $I_{s,t} \subset \{1, \dots, N\}$ 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 2.9 is an empirical small-ball estimate, as it shows that with high probability, and for every pair s, t as above, a proportional number 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

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

Let $y = v_1 + v_2$ and $x = u_1 + u_2$ as in Corollary 2.8, set $j_0 = \max\{j_{-,R}, j_{+,R}\}$ and put $J = \text{supp}(v_1) \cup \text{supp}(u_1)$. Observe that $|J| \leq 2j_0$ and that

$$\begin{aligned} \sum_{j \in J} |y(j)| \cdot |x(j)| &\leq \sum_{j \in \text{supp}(v_1)} |v_1(j)x(j)| + \sum_{j \in \text{supp}(u_1)} |y(j)u_1(j)| \\ &\leq \|v_1\|_2 \left(\sum_{i=1}^{2j_0} (x^2(j))^* \right)^{1/2} + \|u_1\|_2 \left(\sum_{i=1}^{2j_0} (y^2(j))^* \right)^{1/2} \\ &\lesssim_L \ell(T_{-,R}) \|s - t\|_2 \cdot \sqrt{j_0 \log(eN/j_0)} \|s + t\|_2 \\ &\quad + \ell(T_{+,R}) \|s + t\|_2 \cdot \sqrt{j_0 \log(eN/j_0)} \|s - t\|_2 \\ &\lesssim_L E_R^2 \|s - t\|_2 \|s + t\|_2 \leq \frac{\kappa_0 N}{4} \|s - t\|_2 \|s + t\|_2, \end{aligned}$$

because, by the definition of j_0 , $\sqrt{j_0 \log(eN/j_0)} \lesssim \max\{\ell(T_{-,R}), \ell(T_{+,R})\}$ and $N \geq c_0 E_R^2 / \kappa_0^2$ 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 2.6 for $\beta = c_1(L, \kappa_0)$. The claim follows immediately from that lemma. \square

3 Proof of Theorem A

It is well understood that when analyzing properties of ERM relative to a loss, studying the excess loss functional $x \in T \mapsto \mathcal{L}_x = \ell_x - \ell_{x_0}$ is rather natural. Here, the loss function is $\ell_x(a, y) = (\langle a, x \rangle^2 - y)^2$ and 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 $\mathcal{L}_{x_0} = 0$. Moreover, since $P_N \mathcal{L}_{x_0} = 0$ is a potential minimizer of $\{P_N \mathcal{L}_x : x \in T\}$ (because $x_0 \in T$), the minimizer \hat{x} satisfies that $P_N \mathcal{L}_{\hat{x}} \leq 0$. This gives a way of excluding parts of T as potential empirical minimizers: it suffices to show that with high probability, those parts belong to the set $\{x : P_N \mathcal{L}_x > 0\}$. The exclusion may be achieved by showing that $P_N \mathcal{L}_x$ is equivalent to (or at least larger than) $\mathbb{E}\mathcal{L}_x$, as the latter is positive for points that are not true minimizers.

The second ingredient is a decomposition of the excess loss to a sum of two processes: a quadratic process and a multiplier one [16]: if F is a class of functions,

$f^* = \operatorname{argmin}_{f \in F} \mathbb{E}(f(a) - y)^2$ and $f \in F$, then

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

When $y = \langle x_0, a \rangle^2 + w$ as we have here, each $f_x \in F$ given by $f_x = \langle x, \cdot \rangle^2$, and thus

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

If w is a mean-zero random variable that is independent of a , then by Assumption 1.1,

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

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

The final ingredient is a localization argument. To show that $P_N \mathcal{L}_x > 0$ on a large subset $T' \subset T$ (which implies that \hat{x} cannot be in T'), it suffices to obtain a high probability lower bound on

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

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 here is a localized set $T_R = \{x \in T : \|x - x_0\|_2 \|x + x_0\|_2 \geq R\}$ for a well-chosen R .

3.1 Control of the quadratic and multiplier processes

Let us establish a lower bound on the quadratic process and an upper bound on the multiplier process, both indexed by the set T_R .

Theorem 3.1. *There exists a constant c_0 that depends only on L , and constants c_1, κ_1 that depend 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 2.9 and its proof is omitted.

As for the multiplier process, one has the following:

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 processes, 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., [19]), it follows that for every fixed $(w_i)_{i=1}^N$ and $(a_i)_{i=1}^N$,

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

Applying Remark 2.5, if $N \gtrsim_L E_R$ then with $(\varepsilon \otimes a)^N$ -probability of at least $1 - 2 \exp(-c_1 L^2 \min\{\ell^2(T_{+,R}), \ell^2(T_{-,R})\})$,

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

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 $x \in T_R$,

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

□

Recall that $r_1(\gamma) = \inf\{r > 0 : E_r \leq \gamma \sqrt{N} r\}$ and let $R \geq r_1(c_1 \kappa_0 / (c_2 L^2 \sqrt{\beta}))$.

Observe that 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

$$\rho = \|x - x_0\|_2 \|x + x_0\|_2 \geq R,$$

then for every $x \in T_R$,

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

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

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

$P_N \mathcal{L}_x > 0$ and thus $\hat{x} \notin T_R$. Theorem A now follows from the definition of $r_1(\gamma)$ 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. Using an almost identical argument to the one used above, one may replace the sets $T_{+,R}$ and $T_{-,R}$ by $T_{+,R}(x_0)$ and $T_{-,R}(x_0)$. To that end, given $x_0 \in T$ let $E_r(x_0) = \max\{\ell(T_{+,R}(x_0)), \ell(T_{-,R}(x_0))\}$ and set

$$r_1(x_0, \gamma) = \inf \left\{ r > 0 : E_r(x_0) \leq \gamma \sqrt{Nr} \right\},$$

and

$$r_0(x_0, Q) = \inf \left\{ r > 0 : E_r(x_0) \leq Q \sqrt{Nr} \right\}.$$

Theorem 4.1. *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. Set*

$$r_2^* = \max\{r_0(x_0, c_1), r_1(x_0, c_2/\sigma \sqrt{\log N})\},$$

and under the same assumptions as in Theorem A, with probability at least

$$1 - 2 \exp(-c_0 \min\{\ell^2(T_{+,r_2^*}(x_0)), \ell^2(T_{-,r_2^*}(x_0))\}) - 2N^{-\beta+1},$$

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq r_2^*.$$

Next, let us analyze the structure of the local sets $T_{+,R}(x_0)$ and $T_{-,R}(x_0)$. A first step in that direction is the following:

Lemma 4.2. *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. Without loss of generality assume 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 $\|x + x_0\|_2 \leq 16\|x_0\|_2$. 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 pointwise 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 ‘moreover’ part, observe that

$$\begin{aligned} \|x\|_2^2 - 2\|x_0\|_2 \|x\|_2 &\leq \min\{\|x - x_0\|_2^2, \|x + x_0\|_2^2\} \leq \|x - x_0\|_2 \|x + x_0\|_2 \\ &\leq (\|x\|_2 + \|x_0\|_2)^2, \end{aligned}$$

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

It is clear from Lemma 4.2 that 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 .

Therefore, if T is a convex and centrally-symmetric, then $T_{+,R}(x_0) = T_{-,R}(x_0)$, and the corresponding gaussian averages satisfy

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

Recall that the fixed point conditions appearing in Theorem 4.1 are

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

and

$$r_1(\gamma) = r_1(x_0, \gamma) = \inf\{R : E_R(x_0) \leq \gamma \sqrt{NR}\}, \tag{4.2}$$

and consider the slightly suboptimal choice $\gamma = c_2 / \sigma \sqrt{\log N}$. The assertion of Theorem 4.1 is that with high probability, ERM produces \hat{x} for which

$$\|\hat{x} - x_0\|_2 \|\hat{x} + x_0\|_2 \leq \max\{r_1(\gamma), r_0\}.$$

If $\|x_0\|_2 \geq \sqrt{R}$, then the condition in (4.1) is

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

while (4.2) is

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

Set

$$r_N^*(Q) = \inf\{r > 0 : \ell(T \cap r B_2^n) \leq Q r \sqrt{N}\},$$

and

$$s_N^*(\eta) = \inf\{s > 0 : \ell(T \cap s B_2^n) \leq \eta s^2 \sqrt{N}\}.$$

Therefore,

$$r_0 = 2\|x_0\|_2 r_N^*(c_3)$$

and

$$r_1 (c_2 / (\sigma \sqrt{\log N})) \leq 2\|x_0\|_2 s_N^*(c_4 \|x_0\|_2 / \sigma \sqrt{\log N}).$$

For

$$R = 2\|x_0\|_2 \max\{r_N^*(c_3), s_N^*(c_4 \|x_0\|_2 / \sigma \sqrt{\log N})\},$$

it remains to verify that $\|x_0\|_2^2 \geq R$; that is,

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

To that end, observe that if

$$r_N^*(c_3) \leq \frac{c_3 \sigma}{2c_4 \|x_0\|_2} \sqrt{\log N}, \quad (4.6)$$

then $r_N^*(c_3) \leq s_N^*(c_4 \|x_0\|_2 / \sigma \sqrt{\log N})$. Indeed, by the convexity of T , $r_N^*(Q) \leq \rho$ if and only if $\ell(T \cap \rho B_2^n) \leq Q \rho \sqrt{N}$, and a similar statement holds for s_N^* (see the discussion in [16]). Therefore, if $\ell(T \cap r_N^*(Q)) > \eta (r_N^*(Q))^2 \sqrt{N}$ then $s_N^*(\eta) \geq r_N^*(Q)$, which is indeed the case because $\ell(T \cap r_N^*(Q)) \geq Q r_N^*(Q) \sqrt{N} / 2$, $Q = c_3$ and $\eta = c_4 \|x_0\|_2 / \sigma \sqrt{\log N}$.

Under (4.6), (4.5) becomes $2s_N^*(c_4 \|x_0\|_2 / \sigma \sqrt{\log N}) \leq \|x_0\|_2$ which, by the definition of s_N^* , holds if and only if

$$\ell\left(T \cap \frac{\|x_0\|_2}{2} B_2^n\right) \leq \frac{c_4 \|x_0\|_2}{\sigma \sqrt{\log N}} \cdot \frac{\|x_0\|_2^2}{4} \sqrt{N};$$

i.e., when $\|x_0\|_2 \geq v_N^*(\zeta)$ for $\zeta = c_4 / 4\sigma \sqrt{\log N}$.

Hence, by Theorem 4.1 combined with Lemma 4.2, it follows that with high probability,

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

The other cases, in which either $\|x_0\|_2$ is ‘small’, or when r_0 dominates r_1 are treated in a similar fashion, and are omitted.

5 A minimax lower bound

In this section we obtain a general lower bound on the performance of any procedure in the phase retrieval problem. The estimate we present 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 E . If $A \subset E$, let $M(A, rB)$ be the maximal cardinality of a subset of A that is r -separated with respect to the norm in E .

Clearly, if $M(A, rB) \geq L$ there are $x_1, \dots, x_L \in A$ for which the sets $x_i + (r/3)B$ are disjoint.

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 with variance σ , which is independent of a , consider the target

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

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

Following [16], there are two possible sources of ‘statistical complexity’ that influence the error rate of the problem:

1. There are functions $f \in F$ 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 property does not depend on the choice of the noise: for every $f_0 \in F$ and $\mathbb{A} = (a_i)_{i=1}^N$, the key factor is the $L_2(\mu)$ diameter of the set

$$K(f_0, \mathbb{A}) = \{f \in F : (f(a_i))_{i=1}^N = (f_0(a_i))_{i=1}^N\},$$

and we shall denote that diameter by $d_N^*(\mathbb{A})$.

2. Let D be the unit ball in $L_2(\mu)$. The set

$$(F - f_0) \cap rD = \{f - f_0 : f \in F, \|f - f_0\|_{L_2} \leq r\} \tag{5.2}$$

is ‘rich enough’ at a scale that is proportional to its $L_2(\mu)$ diameter r . If the set is ‘rich’, the procedure is likely to make mistakes because of the interaction between class members and the noise (see the discussion in [16]).

The ‘size’ of the set (5.2) is measured using the cardinality of a maximal $L_2(\mu)$ -separated set it contains. To that end, for $H \subset F$ set

$$C(H, r) = \sup_{h \in H} r \log^{1/2} M(H \cap (h + \theta_0 rD), rD),$$

for some $\theta_0 \geq 2$ and $r > 0$. The constant θ_0 we shall use will be specified later.

For every $H \subset F$, let

$$q_N(H, \eta) = \inf \left\{ r > 0 : C_0(H, r) \leq \eta r^2 \sqrt{N} \right\} \tag{5.3}$$

if the infimum is smaller than $\text{diam}(H, L_2(\mu))$; otherwise, set $q_N(H, \eta) = \text{diam}(H, L_2(\mu))$.

Remark 5.2. It follows from Sudakov’s inequality (see, e.g. [19]) that for every $H \subset L_2(\mu)$ and every $r > 0$,

$$C_0(H, r) \lesssim \sup_{h \in H} \mathbb{E} \|G\|_{H \cap (h + \theta_0 rD)},$$

which hints to the connection between q_N and s_N^* .

The lower bound is an outcome of the following fact:

Theorem 5.3. [16] 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 $\theta_1 > 0$

that depends only on θ_0 for which, for every procedure \tilde{f} ,

$$\sup_{f_0 \in F} \mathbb{P}_{f_0}^{\otimes N} \left(\left\| f_0 - \tilde{f} \right\|_{L_2} \geq \max\{q_N(F, \theta_1/\sigma), (d_N^*(\mathbb{A})/4)\} \right) \geq 1/5. \quad (5.4)$$

Minimax bounds of a similar flavour may be found in [28], [34] in the context of density estimation, and also in [1].

To apply this general principle to the phase retrieval problem generated by $f_0(a) + w$, note that $f_0(a) = \langle x_0, a \rangle^2 \equiv f_{x_0}(a)$ for some unknown vector $x_0 \in T \subset \mathbb{R}^n$, while the estimators generated by the procedure are $\tilde{f} = \langle \tilde{x}, \cdot \rangle^2$. Also, observe that for every $x_1, x_2 \in T$,

$$\|f_{x_0} - f_{x_1}\|_{L_2}^2 = \mathbb{E}(\langle x_0, a \rangle^2 - \langle x_1, a \rangle^2)^2 = \mathbb{E}\langle x_0 - x_1, a \rangle^2 \langle x_0 + x_1, a \rangle^2$$

and therefore, to apply Theorem 5.3, one has to identify the L_2 structure of the set

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

We will do so by assuming the following:

Assumption 5.1. There exist constants C_1 and $C_2 > 2$ for which, for every $s, t \in \mathbb{R}^n$,

$$C_1 \|s - t\|_2 \|s + t\|_2 \leq (\mathbb{E}\langle s - t, a \rangle^2 \langle s + t, a \rangle^2)^{1/2} \leq C_2 \|s - t\|_2 \|s + t\|_2.$$

It is straightforward to verify that if a is an L -subgaussian vector in \mathbb{R}^n that satisfies Assumption 1.1, then it automatically satisfies Assumption 5.1 for constants C_1 and C_2 that depend only on L .

As will be made clear later, the lower bound on $\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2$ changes with $\|x_0\|_2$. Therefore, one has to consider each shell $V_0 = T \cap R_0 S^{n-1}$ for $R_0 > 0$ and the corresponding class of functions $F' = \{f_u : u \in V_0\}$ separately. The lower bound is obtained by selecting the ‘worst’ V_0 and $x_0 \in V_0$, which will be used to generate the target $y = \langle x_0, a \rangle^2 + w$.

Note that by Assumption 5.1, for every $u, v \in T$,

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

Hence, for every $r > 0$,

$$\left\{ v \in V_0 : \|v - x_0\|_2 \|v + x_0\|_2 \leq \frac{\theta_0 r}{C_2} \right\} \subset \{v \in V_0 : f_v \in f_{x_0} + \theta_0 r D\}. \quad (5.5)$$

Set $\mathcal{D}(u, v) = \|u - v\|_2 \|u + v\|_2$ and put $\mathcal{B}_x(\rho) = \{u \in \mathbb{R}^n : \mathcal{D}(u, x) \leq \rho\}$. Invoking Theorem 5.3 and (5.5), it suffices to construct a well-separated set in $V_0 \cap \mathcal{B}_{x_0}(\theta_0 r/C_2)$, in the sense that for every $i \neq j$, $\mathcal{D}(x_i, x_j)$ is large enough.

Formally, Let $\theta_0 > 2$ to be named later and set θ_1 as in Theorem 5.3. If there are $x_0 \in V_0$ and $\{x_1, \dots, x_k\} \subset V_0$ that satisfy

1. $\|x_i - x_0\|_2 \|x_i + x_0\|_2 \leq \theta_0 r/C_2$, and
2. for every $1 \leq i < j \leq k$, $\|x_i - x_j\|_2 \|x_i + x_j\|_2 \geq r/C_1$,

then

$$\sup_{f_0 \in F'} \log M(F' \cap (f_0 + \theta_0 r D), r D) \geq \log k. \quad (5.6)$$

As a consequence, if r also satisfies that

$$\log k > N \left(\frac{\theta_1 r}{\sigma} \right)^2,$$

then $r \leq q_N(F', \theta_1/\sigma)$ and with probability at least $1/5$, the error of any procedure \tilde{x} is at least r ; that is, for any procedure there will be some $v_0 \in V_0$ for which, given data generated by $\langle v_0, \cdot \rangle + w$, with probability $1/5$

$$\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2 \geq r.$$

What now remains is to identify when such a separated set exists and to relate $\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2$ to $\min\{\|\tilde{x} - x_0\|_2, \|\tilde{x} + x_0\|_2\}$.

We begin with the following simple observation:

Lemma 5.4. *If $\|u\|_2 = R_0$ and $\|v\|_2 \leq R_0$ then*

$$R_0 \min\{\|u - v\|_2, \|u + v\|_2\} \leq \mathcal{D}(u, v) \leq 2R_0 \min\{\|u - v\|_2, \|u + v\|_2\}$$

Proof. Assume without loss of generality that $\|u - v\|_2 \leq \|u + v\|_2$. Thus, $\langle u, v \rangle \geq 0$ and

$$\|u + v\|_2^2 = \|u\|_2^2 + \|v\|_2^2 + 2\langle u, v \rangle \geq R_0^2.$$

Therefore, $\|u + v\|_2 \geq R_0$ and by the triangle inequality, $\|u + v\|_2 \leq 2R_0$, completing the proof. \square

To formulate the next observation, let C_1 and C_2 be as in Assumption 5.1. Set $R_0 > 0$ and consider $V_0 = T \cap R_0 S^{n-1}$ and $F' = \{f_v : v \in V_0\}$.

Lemma 5.5. *If $x_0 \in V_0$, $r < C_2 R_0^2 / \theta_0$ and $\{x_1, \dots, x_k\} \subset V_0 \cap (x_0 + \theta_0 r / 2 C_2 R_0) B_2^n$ is $r / C_1 R_0$ separated in ℓ_2^n , then*

$$M(F' \cap (f_{x_0} + \theta_0 r D), r D) \geq k.$$

Proof. Set $R = r / C_2$ and observe that $R_0 \geq \sqrt{\theta_0 R}$. By Lemma 5.4, if $x \in V_0$ satisfies that

$$\|x_0\|_2 \min\{\|x - x_0\|_2, \|x + x_0\|_2\} \leq \theta_0 R / 2,$$

then $\|x - x_0\|_2 \|x + x_0\|_2 \leq \theta_0 R$, and in particular, $\|f_x - f_{x_0}\|_{L_2} \leq \theta_0 r$. Hence,

$$V_0 \cap (x_0 + (\theta_0 R / 2 \|x_0\|_2) B_2^n) \subset \{x \in V_0 : \|x_0 - x\|_2 \|x_0 + x\|_2 \leq \theta_0 R\}.$$

Note that if $x_i, x_j \in V_0 \cap (x_0 + (\theta_0 R / 2 \|x_0\|_2) B_2^n)$ then by the triangle inequality

$$\|x_i + x_j\|_2 \geq 2\|x_0\|_2 - \theta_0 R / \|x_0\|_2 \geq R_0 \tag{5.7}$$

because $R_0 = \|x_0\|_2 \geq \sqrt{\theta_0 R}$.

Hence, if $\{x_1, \dots, x_k\} \subset V_0 \cap (x_0 + (\theta_0 R / 2 R_0) B_2^n)$ is $r / C_1 R_0$ -separated in ℓ_2^n , it is evident that

$$\|x_i + x_j\|_2 \|x_i - x_j\|_2 \geq R_0 \cdot \frac{r}{C_1 R_0} = \frac{r}{C_1}.$$

The corresponding functions satisfy that $f_{x_i} \in f_{x_0} + \theta_0 r D$, and $\|f_{x_i} - f_{x_j}\|_{L_2} \geq r$ as required. \square

Let $c_0 = C_1/2$ and set

$$C(R_0, \rho) = \sup_{x_0 \in V_0} \rho \log^{1/2} (M(V_0 \cap (x_0 + c_0 \rho B_2^n), \rho B_2^n))$$

where for every R_0 , $V_0 = T \cap R_0 S^{n-1}$. Finally, let

$$q_N^*(R_0, \eta) = \inf \left\{ \rho > 0 : C(R_0, \rho) \leq \eta \rho^2 \sqrt{N} \right\}$$

if the infimum is smaller than R_0 ; otherwise, set $q_N^*(R_0, \eta) = R_0$.

Theorem C. There exist constants c_1, c_2 and c_3 that depend only on C_1 and C_2 for which the following holds. Let $R_0 > 0$ and assume that $V_0 = T \cap R_0 S^{n-1}$ is nonempty. For any procedure \tilde{x} , there is some $x_0 \in V_0$, for which, with probability at least $1/5$, given the data generated by $y = \langle x_0, a \rangle^2 + w$,

$$\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2 \geq c_1 \|x_0\|_2 q_N^* \left(\|x_0\|_2, \frac{c_2 \|x_0\|_2}{\sigma} \right)$$

and

$$\min\{\|\tilde{x} - x_0\|_2, \|\tilde{x} + x_0\|_2\} \geq c_1 q_N^* \left(\|x_0\|_2, \frac{c_3 \|x_0\|_2}{\sigma} \right).$$

In particular, the error rates in T cannot be better than $\sup_{R_0} R_0 q_N^*(R_0, c_2 R_0/\sigma)$ and of $\sup_{R_0} q_N^*(R_0, c_3 R_0/\sigma)$, respectively (up to some numerical constant).

Proof. First, let us specify the choice of constants that have been used above: given $C_1 \leq C_2 < 2$ as in Assumption 5.1, set $\theta_0 = C_2$ and let θ_1 be as in Theorem 5.3.

Fix R_0 for which $T \cap R_0 S^{n-1}$ is nonempty and let $\eta = \theta_1 C_1 R_0/\sigma$. It is evident from the definition of q_N^* that there is some $x_0 \in V_0 = T \cap R_0 S^{n-1}$ for which

$$\log^{1/2} (M(V_0 \cap (x_0 + c_0 (q_N^*/2) B_2^n), (q_N^*/2) B_2^n)) \geq \eta q_N^* \sqrt{N}/2,$$

(and recall that $c_0 = C_1/2$).

Set $r = c_0 R_0 q_N^*$ and clearly $r < C_2 R_0^2/\theta_0$ as required in Lemma 5.5. Also, $c_0 (q_N^*/2) = r/2R_0 = (\theta_0/C_2) \cdot r/2R_0$, $q_N^*/2 = r/2c_0 R_0 = r/C_1 R_0$ and $\eta q_N^*/2 = (\theta_1/\sigma)r$. Applying Lemma 5.5

$$\log M(F' \cap (f_{x_0} + \theta_0 r D), r D) \geq N \left(\frac{\theta_1}{\sigma} \right)^2 r^2,$$

where $F' = \{f_u : u \in T \cap R_0 S^{n-1}\}$.

Therefore, by Theorem 5.3, given any procedure \tilde{x} there is some $v_0 \in V_0$ for which with probability at least $1/5$,

$$\|f_{\tilde{x}} - f_{v_0}\|_{L_2} = \|\tilde{x} - v_0\|_2 \|\tilde{x} + v_0\|_2 \geq r = \frac{C_1}{2} \|v_0\|_2 q_N^* \left(\|v_0\|_2, C_1 \|v_0\|_2 \frac{\theta_1}{\sigma} \right). \quad (5.8)$$

Finally, one has to show that

$$\min\{\|\tilde{x} - v_0\|_2, \|\tilde{x} + v_0\|_2\} \gtrsim q_N^* \left(\|v_0\|_2, C_1 \|v_0\|_2 \frac{\theta_1}{\sigma} \right). \quad (5.9)$$

And indeed, since $\|\tilde{x} - v_0\|_2 \|\tilde{x} + v_0\|_2 \geq r$ and $r \lesssim R_0^2 = \|v_0\|_2^2$, the proof of (5.9) follows that path of Lemma 4.2 and is omitted.

The claim now follows by taking the worst possible choices of R_0 – that is, the choices that lead to the largest values in (5.8) and (5.9). \square

6 Examples

Here, we will present two simple applications of the upper and lower bounds on the performance of ERM in phase retrieval. Naturally, there are many other examples that follow in a similar way and that can be derived using very similar arguments. The choice of examples has been made to illustrate the question of optimality of ERM, as well as an indication of the similarities and differences between phase retrieval 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 is of a class with a ‘constant’ local complexity.

Let $T = W_d$ be the set of d -sparse vectors in \mathbb{R}^n ; that is, a set consisting of vectors supported on at most d nonzero coordinates.

Corollary 6.1. *Under the assumptions of Theorem A and for $x_0 \in W_d$ the following holds. If $N \geq c_0(L)d \log(en/d)$, then with probability at least $1 - 2 \exp(-c(L)d \log(en/d)) - N^{-\beta+1}$, ERM produces \hat{x} that satisfies*

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

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

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

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

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

Proof. It is straightforward to verify (see, for instance, Lemma 3.3.1 in [7]) that

$$\ell(W_d \cap S^{n-1}) \sim \sqrt{d \log\left(\frac{en}{d}\right)}. \quad (6.4)$$

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 $I \subset \{1, \dots, n\}$ of cardinality d that is disjoint of $\text{supp}(x_0)$, the sets

$$\left\{ \frac{(x - x_0)_{i \in I}}{\|x - x_0\|_2} : x \in W_d \right\}, \quad \text{and} \quad \left\{ \frac{(x + x_0)_{i \in I}}{\|x + x_0\|_2} : x \in W_d \right\}$$

contain $(1/\sqrt{2})S^I$ – the Euclidean sphere of radius $1/\sqrt{2}$. Thus, for every $R > 0$,

$$\ell(T_{+,R}), \ell(T_{-,R}) \sim \sqrt{d \log(en/d)}. \quad (6.5)$$

To apply Theorem A one has to identify the fixed points $r_1(\gamma)$ and $r_0(Q)$ for the right

choice of γ and Q . Since $E_R \sim \sqrt{d \log(en/d)}$, it follows that for $N \gtrsim_L (d/Q^2) \log(en/d)$,

$$r_0(Q) = 0 \text{ and } r_1(\gamma) \sim \frac{1}{\gamma} \sqrt{\frac{d}{N} \log\left(\frac{en}{d}\right)}.$$

Therefore, with probability at least $1 - 2 \exp(-c(L)d \log(en/d)) - N^{-\beta+1}$, \hat{x} satisfies

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

Finally, the estimate on $\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\}$ is an immediate outcome of Lemma 4.2 and is omitted. \square

When $\|x_0\|_2$ is of the order of a constant, the error rate in (6.1) and (6.2) is identical to the one obtained in [16] for 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 error rate 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 (again, up to logarithmic terms).

When the noise level σ tends to zero, the error rates in linear regression and in phase retrieval both tend to zero. In particular, exact reconstruction occurs - that is $\hat{x} = x_0$ in linear regression and $\hat{x} = x_0$ or $\hat{x} = -x_0$ in phase retrieval - in the noise-free case when $N \gtrsim_L d \log(en/d)$.

The following result shows that the upper bounds obtained in Corollary 6.1 are the minimax rate, up to the $\sqrt{\log N}$ term.

Corollary 6.2. *Consider the phase retrieval model (1.1), for w that is gaussian, has variance σ and is independent of a . If the number of observations is $N \gtrsim d \log(en/d)$ and $R_0 > 0$, then for any procedure \tilde{x} there exists a d -sparse vector x_0 for which $\|x_0\|_2 = R_0$ and with probability at least $1/5$,*

$$\|\tilde{x} - x_0\|_2 \|\tilde{x} + x_0\|_2 \gtrsim \min\left(\sigma \sqrt{\frac{d \log(en/d)}{N}}, \|x_0\|_2^2\right)$$

and

$$\min\{\|\tilde{x} - x_0\|_2, \|\tilde{x} + x_0\|_2\} \gtrsim \min\left(\frac{\sigma}{\|x_0\|_2} \sqrt{\frac{d \log(en/d)}{N}}, \|x_0\|_2\right).$$

The proof of Corollary 6.2 follows from Theorem C and a standard entropy estimate, namely, that for every $r > 0$ and $c_0 \geq 2$,

$$\log^{1/2} M(W_d \cap c_0 r B_2^n, r B_2^n) \sim \sqrt{d \log(en/d)} \tag{6.6}$$

(see, e.g. Lemma 1.4.2 and Lemma 2.2.17 in [7]).

As a consequence, ERM is a minimax procedure for phase retrieval of d -sparse vectors, up to the $\log^{1/2} N$ factor in the upper estimate.

One specific choice of d which is of natural interest is when $d = n$ and $T = \mathbb{R}^n$. Thus, there is no a-priori information on the signal x_0 that one would like to recover.

This problem has been studied in [6], and it turns out that in the noiseless case, exact

reconstruction is possible using the PhaseLift procedure when $N \gtrsim n$.

Our results lead to exact recovery as well: by Corollary 6.1, if $N \geq_L n$, ERM produces \hat{x} for which, with high probability, either $\hat{x} = x_0$ or $-\hat{x} = x_0$.

In the noisy case, the problem has been studied in [6] when a is the standard gaussian vector (see Theorem 1.3 there). It follows that if $N \gtrsim n$, PhaseLift (together with a computation of the leading eigenvector) yields an estimator \bar{x} for which, with high probability,

$$\min\{\|\bar{x} - x_0\|, \|\bar{x} + x_0\|\} \lesssim \begin{cases} \|x_0\|_2 & \text{when } \|x_0\|_2^2 \leq \sigma \\ \frac{\sigma}{\|x_0\|_2} & \text{otherwise.} \end{cases} \quad (6.7)$$

Corollary 6.2 shows that this estimator is minimax when $N \sim n$, though the estimate is clearly suboptimal when N is much larger than n , as it does not tend to 0 with N .

Comparing (6.7) with Corollary 6.1, it follows from the latter that when a is L -subgaussian and $N \gtrsim_L n$, with probability at least $1 - 2 \exp(-c(L)n) - N^{-\beta+1}$, ERM produces \hat{x} that satisfies

$$\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \lesssim_L \begin{cases} \left[\sigma \sqrt{\frac{n \log N}{N}} \right]^{1/2} & \text{when } \|x_0\|_2^2 \lesssim \sigma \sqrt{\frac{n \log N}{N}} \\ \frac{\sigma}{\|x_0\|_2} \sqrt{\frac{n \log N}{N}} & \text{otherwise.} \end{cases} \quad (6.8)$$

which is minimax up to the $\sqrt{\log N}$ term for any $N \gtrsim_L n$.

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 whose local complexity changes with x_0 – it increases the closer x_0 is to 0. As an added value, one may obtain sharp estimates on $\ell(B_1^n \cap rB_2^n)$ at every scale $r > 0$. Indeed, one may show (see, for example, [13]) that

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

It follows that for B_1^n , one has

$$r_N^*(Q) \begin{cases} \sim \left(\frac{1}{Q^2 N} \log \left(\frac{n}{Q^2 N} \right) \right)^{1/2} & \text{if } n \geq C_0 Q^2 N \\ \lesssim \frac{1}{N} & \text{if } C_1 Q^2 N \leq n \leq C_0 Q^2 N \\ = 0 & \text{if } n \leq C_1 Q^2 N. \end{cases}$$

where C_0 and C_1 are absolute constants. The only range in which this estimate is not sharp is when $n \sim Q^2 N$, 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 $n \sim Q^2 N$ (see [17]), but since it is not our main interest, we will not pursue it further, and only consider the cases $n \leq C_1 Q^2 N$ and $n \geq C_0 Q^2 N$.

A straightforward computation shows that the two other fixed points from Definition 1.3 satisfy:

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

and

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

Theorem B, leading to an upper estimate on

$$(*) = \min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\},$$

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, $(*)$ is upper bounded by $r_N^*(Q)$, for Q that is an absolute constant. In particular, when $n \geq C_0 Q^2 N$, then with high probability

$$\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\} \lesssim_L \left(\frac{1}{N} \log(n/N)\right)^{1/2}.$$

Let us show that this estimate is optimal, and that ERM is a minimax procedure. To that end, the minimax lower bound $d_N^*(\mathbb{A})$ in Theorem 5.3 may be used, as no procedure can do better than $d_N^*(\mathbb{A})/4$, with probability greater than $1/5$.

Using the notation of section 5, in the phase retrieval problem one has

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

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

By Lemma 4.2, for $x_0 = (1/2, 0, \dots, 0) \in B_1^n$ (in fact, any vector x_0 in B_1^n for which $0 < \|x_0\|_2 \leq 1/2$ would do),

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

which is the Gelfand N -width of B_1^n . By the well-known result of Garnaev and Gluskin [11],

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

which is of the same order as r_N^* , (except when $n \sim N$, which is not treated here), implying that ERM is a minimax procedure.

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

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

All the resulting estimates, summarized in Figure 1, follow from a straightforward

yet tedious computation. We will only sketch the case $\zeta^{2/3}N^{1/3} \leq \eta\sqrt{N} \leq C_1Q^2N$, which is equivalent to

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

$(*) \lesssim$	$\sigma/\ x_0\ _2 \leq c_0 r_N^*/\sqrt{\log N}$	$\sigma/\ x_0\ _2 \geq c_0 r_N^*/\sqrt{\log N}$
$\ x_0\ _2 \leq v_N^*$	r_N^*	s_N^*
$\ x_0\ _2 \geq v_N^*$	r_N^*	v_N^*

Figure 1: High probability bounds on $\min\{\|\hat{x} - x_0\|_2, \|\hat{x} + x_0\|_2\}$.

The upper bound on $(*)$ changes according to the way N scales with n :

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

$$(*) \leq \begin{cases} \left(\frac{\sigma^2 \log N}{\|x_0\|_2^2 N} \log\left(\frac{\sigma^2 n^2}{\|x_0\|_2^2 N}\right)\right)^{1/4} & \text{if } \|x_0\|_2 \geq \left(\frac{\sigma^2 \log N}{N} \log\left(\frac{\sigma^2 n^3}{N}\right)\right)^{1/6} \\ \left(\frac{\sigma^2 \log N}{N} \log\left(\frac{\sigma^2 n^3}{N}\right)\right)^{1/6} & \text{otherwise.} \end{cases} \quad (6.9)$$

2. $c_1 \|x_0\|_2 / (\sigma \sqrt{\log N}) \sqrt{N} \leq n \leq C_1 Q^2 N$. In that case $r_N^* = 0$. In particular $\sigma/\|x_0\|_2 > c_0 r_N^*/\sqrt{\log N}$ and therefore, $(*)$ is upper bounded as in (6.9).
3. $(c_1/(\sigma \sqrt{\log N}))^{2/3} N^{1/3} \leq n \leq c_1 \|x_0\|_2 / (\sigma \sqrt{\log N}) \sqrt{N}$. Again, in this case, $r_N^* = 0$ and

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

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

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

Observe that up to the extra $\sqrt{\log N}$ factor, these estimates are optimal in the minimax sense. Indeed, it is enough to apply Theorem C and verify, as in Example 2 in [20], that when $\|x_0\|_1 \leq 1/2$, for every $\varepsilon < 1/4$

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

References

- [1] L. Birgé, *Approximation dans les espaces métriques et théorie de l'estimation*, Z. Wahrsch. Verw. Gebiete, 65(2), 1983 MR-0722129
- [2] E.J. Candès and Y. C. Eldar, T. Strohmer, and V. Voroninski. *Phase retrieval via matrix completion*. SIAM J. Imaging Sci., 6(1), 2013. MR-3032952

- [3] E.J. Candès, T. Strohmer, and V. Voroninski. *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*. Communications on Pure and Applied Mathematics, 67(6), pp. 906–956, 2012. MR-3069958
- [4] E.J. Candès and X. Li and M. Soltanolkotabi. *Phase retrieval from Coded Diffraction Patterns*. To appear in *Applied and Computational Harmonic Analysis*. arXiv:1310.3240. MR-3352016
- [5] E.J. Candès and X. Li and M. Soltanolkotabi. *Phase retrieval via Wirtinger flow: theory and algorithms*. IEEE Transactions on Information Theory, 61(4), pp. 1985–2007, 2014. MR-3332993
- [6] E.J. Candès and X. Li. *Solving quadratic equations via phaselift when there are about as many equations as unknowns*. Foundations of Computational Mathematics, 14(5), pp. 1017–1026, 2012. MR-3260258
- [7] D. Chafaï, O. Guédon, G. Lecué, and A. Pajor. *Interactions between compressed sensing random matrices and high dimensional geometry*, volume 37 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2012. MR-3113826
- [8] Y.C. Eldar and S. Mendelson. *Phase Retrieval: Stability and Recovery Guarantees*. Applied and Computational Harmonic Analysis, Elsevier, Sep. 2013. MR-3175089
- [9] J.R. Fienup. *Phase retrieval algorithms: a comparison*. Applied Optics, 21(15):2758-2769, 1982.
- [10] F. Fogel and I. Waldspurger and A. D’Aspremont. *Phase Retrieval for Imaging Problems*. arXiv:1304.7735
- [11] A. Yu. Garnaev and E. D. Gluskin. *The widths of a Euclidean ball*. Dokl. Akad. Nauk SSSR, 277(5):1048–1052, 1984. MR-0759962
- [12] E. Giné, V. de la Peña. *Decoupling: From Dependence to Independence*. Springer-Verlag, 1999. MR-1666908
- [13] Y. Gordon, A.E. Litvak, S. Mendelson and A. Pajor. *Gaussian averages of interpolated bodies and applications to approximate reconstruction*. Journal of Approximation Theory, 149(1), 2007 MR-2371614
- [14] R. Gerchberg and W. Saxton. *A practical algorithm for the determination of phase from image and diffraction plane pictures*. Optik, 35:237-246, 1972.
- [15] V. Koltchinskii and S. Mendelson. *Bounding the smallest singular value of a random matrix without concentration*. To appear in *International Mathematics Research Notices*. arXiv:1312.3580.
- [16] G. Lecué, S. Mendelson, *Learning subgaussian classes : Upper and minimax bounds* arXiv:1305.4825.
- [17] 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. MR-3060158
- [18] G. Lecué, S. Mendelson. *Sparse recovery under weak moment assumption*. To appear in *Journal of the European Mathematical Society*. arXiv:1401.2188.
- [19] M. Ledoux, M. Talagrand. *Probability in Banach spaces. Isoperimetry and processes*. Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23. Springer-Verlag, Berlin, 1991. MR-1102015
- [20] S. Mendelson. *On the geometry of subgaussian coordinate projections*. Preprint.
- [21] S. Mendelson. *Learning without concentration*. To appear in *Journal of the ACM*. arXiv:1410.3192.
- [22] S. Mendelson. *Learning without concentration for general losses*. arXiv:1410.3192
- [23] S. Mendelson. *A remark on the diameter of random sections of convex bodies*. Geometric Aspects of Functional Analysis (GAFA Seminar Notes). Lecture notes in Mathematics 2116, pages 395–404, 2014.
- [24] S. Mendelson. *Upper bounds on product and multiplier empirical processes*. arXiv:1410.8003
- [25] V.D. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*. Lecture Notes in Mathematics 1200, Springer, 1986. MR-0856576

- [26] Y. Shechtman, Y.C. Eldar, A. Szameit, and M. Segev. *Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing*. Optics Express, 19(16):14807-14822, 2011.
- [27] Y. Shechtman and Y.C. Eldar and O. Cohen and H.N. Chapman and J. Miao and M. Segev. *Phase Retrieval with Application to Optical Imaging*, IEEE signal processing society, 32(3), pp. 87 – 109, May 2015.
- [28] A.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. MR-2724359
- [29] J. Tropp. *Convex recovery of a structured signal from independent random linear measurements*. To appear in *Sampling Theory, a Renaissance*. arXiv:1405.1102.
- [30] A. W. Van der Vaart and John A. Wellner. *Weak convergence and empirical processes*. Springer Verlag, 1996. MR-1385671
- [31] V. Voroninski and Z. Xu. *A strong restricted isometry property, with an application to phaseless compressed sensing*. arXiv:1404.3811.
- [32] I. Waldspurger, A. d’Aspremont, and S. Mallat. *Phase recovery, maxcut and complex semidefinite programming*. math. program. 149(1-2):47–81, Ser. A, 2015. MR-3300456
- [33] J.D. Watson and F.H. Crick. *A structure for Deoxyribose Nucleic Acid*. Nature, 171, 1953.
- [34] Y. Yang and A. Barron. *Information-theoretic determination of minimax rates of convergence*. Ann. Statist., 27(5):1564–1599, 1999. MR-1742500

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS¹)
- Non profit, sponsored by IMS², BS³, PKP⁴
- Purely electronic and secure (LOCKSS⁵)

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

²IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

³BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁴PK: Public Knowledge Project <http://pkp.sfu.ca/>

⁵LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>