

# Sharp MSE Bounds for Proximal Denoising

Samet Oymak

Babak Hassibi\*

Department of Electrical Engineering  
Caltech, Pasadena - 91125  
soymak@caltech.edu, hassibi@caltech.edu

## Abstract

Denoising has to do with estimating a signal  $\mathbf{x}_0$  from its noisy observations  $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$ . In this paper, we focus on the “structured denoising problem”, where the signal  $\mathbf{x}_0$  possesses a certain structure and  $\mathbf{z}$  has independent normally distributed entries with mean zero and variance  $\sigma^2$ . We employ a structure-inducing convex function  $f(\cdot)$  and solve  $\min_{\mathbf{x}} \{\frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \sigma\lambda f(\mathbf{x})\}$  to estimate  $\mathbf{x}_0$ , for some  $\lambda > 0$ . Common choices for  $f(\cdot)$  include the  $\ell_1$  norm for sparse vectors, the  $\ell_1 - \ell_2$  norm for block-sparse signals and the nuclear norm for low rank matrices. The metric we use to evaluate the performance of an estimate  $\mathbf{x}^*$  is the normalized mean-squared-error  $\text{NMSE}(\sigma) = \frac{\mathbb{E}\|\mathbf{x}^* - \mathbf{x}_0\|_2^2}{\sigma^2}$ . We show that NMSE is maximized as  $\sigma \rightarrow 0$  and we find the *exact* worst case NMSE, which has a simple geometric interpretation: the mean-squared-distance of a standard normal vector to the  $\lambda$ -scaled subdifferential  $\lambda\partial f(\mathbf{x}_0)$ . When  $\lambda$  is optimally tuned to minimize the worst-case NMSE, our results can be related to the constrained denoising problem  $\min_{f(\mathbf{x}) \leq f(\mathbf{x}_0)} \{\|\mathbf{y} - \mathbf{x}\|_2\}$ . The paper also connects these results to the generalized LASSO problem, in which, one solves  $\min_{f(\mathbf{x}) \leq f(\mathbf{x}_0)} \{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2\}$  to estimate  $\mathbf{x}_0$  from noisy linear observations  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{z}$ . We show that certain properties of the LASSO problem are closely related to the denoising problem. In particular, we characterize the normalized LASSO cost and show that it exhibits a “phase transition” as a function of number of observations. Our results are significant in two ways. First, we find a simple formula for the performance of a general convex estimator. Secondly, we establish a connection between the denoising and linear inverse problems.

**Keywords:** convex optimization, proximity operator, structured sparsity, statistical estimation, model fitting, stochastic noise, linear inverse, generalized LASSO

## 1 Introduction

Signals exhibiting structured behavior play a critical role in various applications. In particular, sparse signals, block-sparse signals, low-rank matrices and their many variations are often the underlying solutions of problems, with applications ranging from MRI to recommendation systems to DNA microarrays, [1–9]. Hence, a significant amount of work has been dedicated to developing and analyzing algorithms, that can take advantage of the signal structure. In this work, we will be considering the estimation of structured signals corrupted by additive noise via convex optimization. Under Gaussian noise assumption, we will provide an exact characterization of the estimation performances of useful convex algorithms based on *proximity operator*. Proximity operator corresponding to a function  $f(\cdot)$  at a point  $\mathbf{y}$  is given by,

$$\text{prox}_f(\mathbf{y}, \lambda) = \arg \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda f(\mathbf{x}) \quad (1.1)$$

Proximity operator was first introduced by Moreau in [10]. It has several *nice* properties, it can often be evaluated quickly and it constitutes the backbone of the proximal algorithms, [10–17]. The topic of this work is to understand and characterize the estimation capabilities of  $\text{prox}_f(\cdot)$  when  $\mathbf{y}$  is the noisy observations

---

\*This work was supported in part by the National Science Foundation under grants CCF-0729203, CNS-0932428 and CIF-1018927, by the Office of Naval Research under the MURI grant N00014-08-1-0747, and by a grant from Qualcomm Inc.

of an underlying structured signal  $\mathbf{x}_0$ . Our results will be particularly meaningful when  $f(\cdot)$  is *structure inducing*. The prime example is the  $\ell_1$  norm, in which case (1.1) is known as “soft-thresholding the entries of  $\mathbf{y}$ ” and has the following closed form solution, [4, 15, 31, 40, 73],

$$[\text{prox}_f(\mathbf{y}, \lambda)]_i = \begin{cases} y_i - \lambda & \text{if } y_i \geq \lambda \\ 0 & \text{if } |y_i| < \lambda \\ y_i + \lambda & \text{else} \end{cases} . \quad (1.2)$$

While sparse signal estimation is the fundamental question in compressed sensing and properties of  $\ell_1$  minimization has been studied extensively, recent literature shows that, various other signal forms show up in diverse set of applications and these applications are growing in a daily basis, [7, 43–46, 48, 49, 51, 52, 54–60, 70]. Consequently, a uniform treatment of the structured signal recovery problems is highly desirable. With this intention, we will loosely say,  $\mathbf{x}_0$  is a structured signal and  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function that exploits this particular structure. Such structured signal-function pairs include the sparse vectors and the  $\ell_1$  norm, and the low-rank matrices and the nuclear norm. Chandrasekaran et al. [8] and Bach [9] consider systematic ways of finding the structure-inducing function given the characteristics of the signal.

With this motivation, we will provide sharp estimation guarantees for the general problem (1.1) where  $\mathbf{y} = \mathbf{x}_0 + \sigma\mathbf{v}$ ,  $\mathbf{x}_0$  is a structured signal, and  $\mathbf{v}$  is the noise vector; whose entries are independent standard normal. As it has been observed in the relevant literature, [19, 29, 30, 34], when noise has variance  $\sigma^2$ , it is useful to consider the slight modification of (1.1),

$$\text{prox}_f(\mathbf{y}, \sigma\lambda) = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \sigma\lambda f(\mathbf{x}), \quad (1.3)$$

which takes the normalization for  $\sigma$  into account. We focus on finding a tight upper bound on the normalized-mean-squared-error (NMSE) defined as,

$$\frac{\mathbb{E}[\|\text{prox}_f(\mathbf{x}_0 + \sigma\mathbf{v}, \sigma\lambda) - \mathbf{x}_0\|_2^2]}{\sigma^2}. \quad (1.4)$$

NMSE (1.4) is a function of variance  $\sigma^2$ , vector  $\mathbf{x}_0$  and function  $f(\cdot)$ . We find a formula for the highest value of NMSE over  $\sigma > 0$ , which is only a function of  $f(\cdot)$  and  $\mathbf{x}_0$ . To state our result, we need to introduce some notation.

- *Subdifferential*: The set of subgradients of a convex function  $f(\cdot)$  at  $\mathbf{x}_0$  will be denoted by  $\partial f(\mathbf{x}_0)$ .
- *Distance*: Given a nonempty set  $\mathcal{C}$ , the distance of a vector  $\mathbf{v}$  to  $\mathcal{C}$  is  $\text{dist}(\mathbf{v}, \mathcal{C}) := \inf_{\mathbf{u} \in \mathcal{C}} \|\mathbf{v} - \mathbf{u}\|_2$ .
- *Mean-squared-distance (MSD)*: Let  $\mathbf{g} \in \mathbb{R}^n$  be a vector with standard normal entries and  $\mathcal{C}$  be a nonempty set. Define  $\mathbf{D}(\mathcal{C}) = \mathbb{E}[\text{dist}(\mathbf{g}, \mathcal{C})^2]$ .

The following theorem gives a sample result.

**Theorem 1.1** (Worst Case NMSE). *Assume  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function,  $\mathbf{x}_0 \in \mathbb{R}^n$ ,  $\lambda \geq 0$  and  $\mathbf{v}$  has independent standard normal entries. Then,*

$$\max_{\sigma > 0} \frac{\mathbb{E}[\|\text{prox}_f(\mathbf{x}_0 + \sigma\mathbf{v}, \sigma\lambda) - \mathbf{x}_0\|_2^2]}{\sigma^2} = \mathbf{D}(\lambda\partial f(\mathbf{x}_0)). \quad (1.5)$$

Furthermore, the worst case NMSE is achieved as  $\sigma \rightarrow 0$ .

**Remark 1**: Observe that, the result is not interesting if the function is differentiable at  $\mathbf{x}_0$ , in which case, subdifferential is a singleton. It becomes useful when the subdifferential is large; which decreases the distance term  $\mathbf{D}(\lambda\partial f(\mathbf{x}_0))$ .

**Remark 2**: The fact that worst case NMSE is achieved as  $\sigma \rightarrow 0$  has been observed in relevant problems which will be discussed later on [33–35, 40]

The quantity  $\mathbf{D}(\lambda\partial f(\mathbf{x}_0))$  may seem abstract at first sight. However, for the problems of interest, the subdifferential  $\partial f(\mathbf{x}_0)$  is a highly structured and well-studied set, [1, 23, 28]. Several works, [8, 25, 34, 61], provide useful upper bounds on this quantity for certain structured signal classes.

	$k$ -sparse, $\mathbf{x}_0 \in \mathbb{R}^n$	Rank $r$ , $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$	$k$ -block sparse, $\mathbf{x}_0 \in \mathbb{R}^{tb}$
$\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$	$(\lambda^2 + 3)k$ for $\lambda \geq \sqrt{2 \log \frac{n}{k}}$	$\lambda^2 r + 2d(r + 1)$ for $\lambda \geq 2\sqrt{d}$	$(\lambda^2 + b + 2)k$ for $\lambda \geq \sqrt{b} + \sqrt{2 \log \frac{t}{k}}$

Table 1: Closed form upper bounds for  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$  corresponding to (1.6), (1.7) and (1.8) (from [34, 61]).

## 1.1 Examples

We will now list some specific examples which are applications of Theorem 1.1 in combination with Table 1.

1. **Sparse signal estimation:** Assume  $\mathbf{x}_0 \in \mathbb{R}^n$  has  $k$  nonzero entries and  $\mathbf{v}$  has independent standard normal entries and  $\mathbf{y} = \mathbf{x}_0 + \sigma \mathbf{v}$ . Pick  $f(\cdot)$  as the  $\ell_1$  norm. Then, for  $\lambda \geq \sqrt{2 \log \frac{n}{k}}$ ,

$$\frac{\mathbb{E}[\|\text{prox}_{\ell_1}(\mathbf{y}, \sigma \lambda) - \mathbf{x}_0\|_2^2]}{\sigma^2} \leq (\lambda^2 + 3)k. \quad (1.6)$$

2. **Low-rank matrix estimation:** Assume  $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$  is a rank  $r$  matrix,  $n = d \times d$ . This time,  $\mathbf{x}_0 \in \mathbb{R}^n$  corresponds to vectorization of  $\mathbf{X}_0$  and  $f(\cdot)$  is chosen as the nuclear norm  $\|\cdot\|_*$  (sum of the singular values of a matrix) [45, 48]. Hence, we observe  $\mathbf{y} = \text{vec}(\mathbf{X}_0) + \sigma \mathbf{v}$  and estimate  $\mathbf{X}_0$  as,

$$\text{prox}_*(\mathbf{y}, \sigma \lambda) = \arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{y} - \text{vec}(\mathbf{X})\|_2^2 + \sigma \lambda \|\mathbf{X}\|_*.$$

Then, whenever  $\lambda \geq 2\sqrt{d}$ , normalized mean-squared-error satisfies,

$$\frac{\mathbb{E}[\|\text{prox}_*(\mathbf{y}, \sigma \lambda) - \mathbf{X}_0\|_F^2]}{\sigma^2} \leq (\lambda^2 + 2d)r + 2d. \quad (1.7)$$

3. **Block sparse estimation:** Let  $n = t \times b$  and assume the entries of  $\mathbf{x}_0 \in \mathbb{R}^n$  can be grouped into  $t$  known blocks of size  $b$  so that only  $k$  of these  $t$  blocks are nonzero. To induce the structure, the standard approach is to use the  $\ell_{1,2}$  norm which sums up the  $\ell_2$  norms of the blocks, [7, 50, 51]. In particular, denoting the subvector corresponding to  $i$ 'th block of a vector  $\mathbf{x}$  by  $\mathbf{x}_i$ , the  $\ell_{1,2}$  norm is equal to  $\|\mathbf{x}\|_{1,2} = \sum_{i=1}^t \|\mathbf{x}_i\|_2$ . Pick  $f(\cdot) = \|\cdot\|_{1,2}$  and assume  $\lambda \geq \sqrt{b} + \sqrt{2 \log \frac{t}{k}}$ ,  $\mathbf{v}$  has independent standard normal entries and  $\mathbf{y} = \mathbf{x}_0 + \sigma \mathbf{v}$ .

$$\frac{\mathbb{E}[\|\text{prox}_{\ell_{1,2}}(\mathbf{y}, \sigma \lambda) - \mathbf{x}_0\|_2^2]}{\sigma^2} \leq (\lambda^2 + b + 2)k. \quad (1.8)$$

## 1.2 Relevant literature

Proximity operator with  $\ell_1$  minimization has been subject of various works and results of type (1.6) is known [4, 8, 27, 29]. Block sparse signals have been studied in [40, 50]. NMSE properties of low-rank matrices has been analyzed in detail by [35–37]. Our main contribution is giving the *exact* worst case NMSE for arbitrary convex functions. As a side result, we provide closed form upper bounds such as (1.6), (1.7) and (1.8). Closer to our results, in [19], Bhaskar et al. considers denoising via structure inducing “atomic norms” with a focus on line spectral estimation. However, their error bounds are looser than what Theorem 1.1 gives. For example, in (1.7), we are able to bound the error in terms of the sparsity of the vector; whereas [19] bounds in terms of the  $\ell_1$  norm of the vector which can be *substantially* larger. In [20], Chandrasekaran and Jordan consider the “constrained denoising” problem which enforces the constraint  $f(\mathbf{x}) \leq f(\mathbf{x}_0)$  rather than penalization  $\sigma \lambda f(\mathbf{x})$  in (1.3). As it will be discussed in Section 3, their results are consistent with ours; however, we additionally show that NMSE bounds are sharp as  $\sigma \rightarrow 0$ .

A more challenging problem occurs when  $\mathbf{y}$  is arising from noisy *linear* observations  $\mathbf{A}\mathbf{x}_0$  and we have to solve an underdetermined linear inverse problem. In this case, we can consider a variant of (1.1), namely,

$$\arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda f(\mathbf{x}). \quad (1.9)$$

When  $f(\cdot)$  is the  $\ell_1$  norm and  $\mathbf{x}_0$  is a sparse signal (1.9) is known as LASSO. LASSO has been the subject of great interest as it is a natural and powerful approach to noise robust compressed sensing [43, 73–78]. Section

3 provides certain results on LASSO and connection of our results to the linear inverse problems in more detail.

Some of the advantages of our results are as follows.

- In various compressed sensing or estimation results, guarantees are orderwise rather than the exact quantities. For example, while NMSE of  $\mathcal{O}(dr)$  is orderwise optimal for nuclear norm minimization (recall (1.7)), it is often critical to know the actual constant term that multiplies  $dr$  for real life applications. For Gaussian noise, it is known that, this constant is as small as 6 rather than being, say, 1000 (set  $\lambda = 2\sqrt{d}$  in (1.7)). Consequently, a *formula* that gives the exact performance is desirable.
- Instead of finding case specific results for various structured signal classes, we are stating our results for a generic convex function  $f(\cdot)$ ; which allows us to systematically obtain the specific instances. The only required information is the subdifferential  $\partial f(\mathbf{x}_0)$ .
- As it will be discussed in Section 3, we will relate Theorem 1.1 to the linear inverse problem (1.9) and in fact, we will find a general relation between them, in connection to the recent results of [8, 25]. In particular, for both problems, the subdifferential based quantity,  $\mathbf{D}(\lambda\partial f(\mathbf{x}_0))$ , plays a critical role.

To provide a better intuition about our results, we will now exemplify other possible signal forms and the associated structure inducing functions.

## 2 Common structures and the associated functions

We consider the following low dimensional signal models and for each case, we explain the signal structure and provide the convex function  $f(\cdot)$  that exploits the structure.

• **Sum of a low rank and a sparse matrix:** The matrix of interest  $\mathbf{X}_0$  can be decomposed into a low rank piece  $\mathbf{L}_0$  and a sparse piece  $\mathbf{S}_0$ , hence it is a “mixture” of the low rank and sparse structures. This model is useful in applications such as video surveillance and face recognition, [57, 58]. In this case, function  $f(\cdot)$  should be chosen so that it simultaneously emphasizes sparse and low rank pieces in the given matrix.

$$f(\mathbf{X}) = \inf_{\mathbf{L}+\mathbf{S}=\mathbf{X}} \|\mathbf{L}\|_* + \gamma\|\mathbf{S}\|_1 \quad (2.1)$$

(2.1) is known as the infimal convolution of the  $\ell_1$  norm and the nuclear norm, [56, 58].

• **Discrete total variation:** In many imaging applications, [64, 66, 67], the signal of interest  $\mathbf{x}_0$  rarely changes as a function of time. Consequently, letting  $\mathbf{d}_i = \mathbf{x}_{0,i+1} - \mathbf{x}_{0,i}$  for  $1 \leq i \leq n-1$ ,  $\mathbf{d} \in \mathbb{R}^{n-1}$  becomes a sparse vector. To induce this structure, one may minimize the total variation of  $\mathbf{x}_0$ , namely  $\|\mathbf{x}_0\|_{TV} = \|\mathbf{d}\|_1$ .

• **Nonuniformly sparse signals and weighted  $\ell_1$  minimization:** In many cases, we might have prior information regarding the sparsity pattern of the signal, [52–55]. In particular, the signal  $\mathbf{x}_0$  might be relatively sparse over a certain region and dense over another. To exploit this additional information we can use a modified  $\ell_1$  minimization where different weights are assigned to different regions. More rigorously, assume the set of entries  $\{1, 2, \dots, n\}$  is divided into  $t$  disjoint sets  $S_1, \dots, S_t$  that correspond to regions with different sparsity levels. Then, given a nonnegative weight vector  $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_t]$ , weighted  $\ell_1$  norm  $\|\mathbf{x}\|_w$  can be given as:

$$\|\mathbf{x}\|_w = \sum_{i=1}^t w_i \sum_{j \in S_i} |x_j|$$

• **Other models:** We can include various other models: signals that are sparse and positive, [41, 42]; positive semidefinite constraints, [45, 46]; simultaneously sparse and low rank matrices, [68–70]; permutation matrices, binary vectors, cut matrices, [8]; etc.

## 3 Main Contributions

### 3.1 Notation

Before stating our main results, we will introduce the relevant notation.  $\mathcal{N}(0, \sigma^2 \mathbf{I}_n)$  is used to denote the distribution of a vector in  $\mathbb{R}^n$  with independent Gaussian entries with variance  $\sigma^2$  and mean zero. For a

convex function  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ , the set of subgradients of  $f(\cdot)$  at  $\mathbf{x}$  is denoted by  $\partial f(\mathbf{x})$ .  $\partial f(\mathbf{x})$  is a nonempty, convex and compact set, [88]. Given a set  $A$  and  $\lambda \in \mathbb{R}$ ,  $\lambda A$  will denote the set obtained by scaling elements of  $A$  by  $\lambda$ . The cone obtained by  $A$  is given as,

$$\text{cone}(A) = \{\lambda \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \in A, \lambda \geq 0\}.$$

When  $A, B$  are two sets,  $A + B$  will denote the Minkowski sum  $\{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A \text{ and } \mathbf{b} \in B\}$ . The closure of a set  $A$  will be denoted by  $\text{Cl}(A)$ . We will now briefly describe our notation on convex geometry. The distance between a set nonempty set  $\mathcal{C}$  and a point  $\mathbf{x} \in \mathbb{R}^n$  is given as:  $\text{dist}(\mathbf{x}, \mathcal{C}) = \min_{\mathbf{s} \in \mathcal{C}} \|\mathbf{x} - \mathbf{s}\|_2$ . When  $\mathcal{C}$  is closed and convex, there exists a unique point in  $\mathcal{C}$  that is closest to  $\mathbf{x}$  called the ‘‘projection of  $\mathbf{x}$  on  $\mathcal{C}$ ’’. This point will be denoted by  $\text{Proj}(\mathbf{x}, \mathcal{C})$ . Basically,  $\text{Proj}(\mathbf{x}, \mathcal{C}) \in \mathcal{C}$  and,

$$\|\mathbf{x} - \text{Proj}(\mathbf{x}, \mathcal{C})\|_2 = \text{dist}(\mathbf{x}, \mathcal{C}).$$

The polar cone of  $\mathcal{C}$ , which is always closed and convex, will be denoted by  $\mathcal{C}^*$  and is given as:

$$\mathcal{C}^* = \{\mathbf{v} \mid \langle \mathbf{v}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in \mathcal{C}\}.$$

### 3.1.1 Useful concepts

**Mean-squared-distance** Let  $\mathcal{C}$  be a nonempty subset of  $\mathbb{R}^n$ . The mean-squared-distance (MSD) to  $\mathcal{C}$  will be denoted by  $\mathbf{D}(\mathcal{C})$  and is defined as,

$$\mathbf{D}(\mathcal{C}) = \mathbb{E}[\text{dist}(\mathbf{g}, \mathcal{C})^2], \quad (3.1)$$

where  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I})$ . The relevant definitions have been used in [34, 61]. This definition is closely related to the concepts like ‘‘statistical dimension’’ of [25] and ‘‘Gaussian width’’ of [8].

**Feasible sets and tangent cones:** Given a convex function  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ , the descent set at  $\mathbf{x}_0$  is given as  $F_f(\mathbf{x}_0) = \{\mathbf{z} \mid f(\mathbf{x}_0 + \mathbf{z}) \leq f(\mathbf{x}_0)\}$ . Then, the tangent cone at  $\mathbf{x}_0$  is defined as  $T_f(\mathbf{x}_0) = \text{Cl}(\text{cone}(F_f(\mathbf{x}_0)))$ . In particular, when  $\mathbf{x}_0$  is not a minimizer of  $f(\cdot)$ , we have  $T_f(\mathbf{x}_0) = \text{cone}(\partial f(\mathbf{x}_0))^*$ . In fact, this will be the only assumption we will be making besides the convexity of  $f(\cdot)$ . The details of this result can be found in [84, 86]. In a similar manner, for a convex set  $\mathcal{C}$  and  $\mathbf{x}_0 \in \mathcal{C}$ , the set of feasible directions at  $\mathbf{x}_0$  is denoted by  $F_{\mathcal{C}}(\mathbf{x}_0)$  and is given by  $F_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{z} \mid \mathbf{x}_0 + \mathbf{z} \in \mathcal{C}\}$ . Then, the tangent cone at  $\mathbf{x}_0$  becomes  $T_{\mathcal{C}}(\mathbf{x}_0) := \text{Cl}(\text{cone}(F_{\mathcal{C}}(\mathbf{x}_0)))$ .

## 3.2 Results on denoising

Throughout the discussion below, we assume  $\mathbf{x}_0$  is *not* a minimizer of  $f(\cdot)$ .

**Theorem 3.1** (Constrained denoising). *Let  $\mathcal{C}$  be a nonempty, convex and closed set and let  $\mathbf{x}_0$  be an arbitrary vector in  $\mathcal{C}$ . Let,*

$$\text{prox}_{\mathcal{C}}(\mathbf{y}) = \arg \min_{\mathbf{x} \in \mathcal{C}} \|\mathbf{y} - \mathbf{x}\|_2. \quad (3.2)$$

Assume  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I}_n)$ . Then,

$$\max_{\sigma > 0} \frac{\mathbb{E}[\|\text{prox}_{\mathcal{C}}(\mathbf{x}_0 + \sigma \mathbf{v}) - \mathbf{x}_0\|_2^2]}{\sigma^2} = \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*). \quad (3.3)$$

Furthermore, the equality above is achieved as  $\sigma \rightarrow 0$ . Now, let  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function and choose  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ . For this choice, we have,

$$\max_{\sigma > 0} \frac{\mathbb{E}[\|\text{prox}_{\mathcal{C}}(\mathbf{x}_0 + \sigma \mathbf{v}) - \mathbf{x}_0\|_2^2]}{\sigma^2} = \mathbf{D}(T_f(\mathbf{x}_0)^*) = \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0))). \quad (3.4)$$

**Remark:** The constrained and the regularized denoising problems are closely related. To relate these, consider the right hand sides of (3.4) and (1.5), in particular, the quantities,  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  and  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$ . Since  $\lambda \partial f(\mathbf{x}_0)$  is subset of  $\text{cone}(\partial f(\mathbf{x}_0))$ , for all  $\lambda \geq 0$ , one has,

$$\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0))) \leq \mathbf{D}(\lambda \partial f(\mathbf{x}_0)).$$

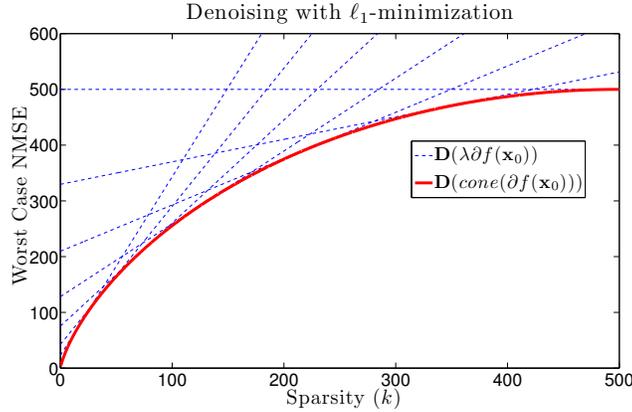


Figure 1: We assumed  $\mathbf{x}_0$  is a  $k$  sparse vector in  $\mathbb{R}^{500}$  and picked  $f(\cdot)$  to be the  $\ell_1$ -norm. The solid red line corresponds to  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  and dashed blue lines correspond to  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$  for  $\lambda$  from 0 to 1.5. In case of  $\ell_1$  minimization, both quantities are solely functions of the sparsity. Observe that  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  is always upper bounded by  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$ .

On the other hand, as it will be discussed later on, when  $f(\cdot)$  is a norm, making use of the results of [25], we can actually obtain,

$$0 \leq \min_{\lambda \geq 0} \mathbf{D}(\lambda \partial f(\mathbf{x}_0)) - \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0))) \leq 2 \frac{\sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2}{f(\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2})}. \quad (3.5)$$

(3.5) can be simplified for structure inducing functions. For example, when  $\mathbf{x}_0$  is a  $k$  sparse vector in  $\mathbb{R}^n$ , right hand side can be replaced by  $2\sqrt{\frac{n}{k}}$ . For the low-rank and block-sparse signals described in Section 1.1, we can use  $2\sqrt{\frac{d}{r}}$  and  $2\sqrt{\frac{t}{k}}$  respectively.

Our next result considers a mixture of constrained and regularized estimators.

**Theorem 3.2** (General upper bound). *Assume  $\mathcal{C}$  is a nonempty, convex and closed set and  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function. Assume  $\mathbf{x}_0 \in \mathcal{C}$ ,  $\lambda \geq 0$  and  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I}_n)$ . Consider the following estimator,*

$$\text{prox}_{f,\mathcal{C}}(\mathbf{y}, \sigma\lambda) = \arg \min_{\mathbf{x} \in \mathcal{C}} \sigma\lambda f(\mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2. \quad (3.6)$$

For any  $\sigma > 0$ , we have,

$$\frac{\mathbb{E}[\|\text{prox}_{f,\mathcal{C}}(\mathbf{x}_0 + \sigma\mathbf{v}, \sigma\lambda) - \mathbf{x}_0\|_2^2]}{\sigma^2} \leq \mathbf{D}(\lambda \partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*). \quad (3.7)$$

### 3.3 Connecting the NMSE to the linear inverse problem

We now turn our attention to the connection between the noiseless linear inverse problem and normalized MSE. Linear inverse problem is the problem of recovering a signal of size  $n$  from  $m$  noiseless linear observations  $\mathbf{y} = \mathbf{A}\mathbf{x}_0$ . To tackle this problem, when  $m < n$ , in a generic setup, a common approach is to make use of a structure inducing function  $f(\cdot)$  and solving,

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (3.8)$$

It is known that, for structured signals, (3.8) exhibits a *phase transition* from failure to success as the number of observations increases. Works by Chandrasekaran et al. [8] and Amelunxen et al. [25] showed that, this phase transition occurs around  $m \approx \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ . This is impressive as  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  not only corresponds to the worst case NMSE of the constrained denoising (3.2); but also to a seemingly unrelated problem (3.8). This observation was first made by Donoho et al. in [40]. Authors in [40] effectively claimed that, the worst case NMSE of optimally tuned (1.3) corresponds to the phase transition point of (3.8). We show that, this is indeed the case as  $\min_{\lambda \geq 0} \mathbf{D}(\lambda \partial f(\mathbf{x}_0)) \approx \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  (recall (3.5)). Hence, our results combined with [8, 25] rigorously justifies the claims in [40].

### 3.4 Relation to LASSO

As a next step, we will now briefly extend our results to the estimation of  $\mathbf{x}_0$  from noisy linear observations of  $\mathbf{x}_0$ . While many generic compressed sensing problems deal with matrices with independent identically distributed entries, we will be stating our results for a random partial unitary matrix  $\mathbf{A}$ . In other words,  $\mathbf{A}$  is generated uniformly at random among the matrices satisfying  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{A}\mathbf{A}^T = \mathbf{I}_m$ .

Matrices with independent standard normal entries and uniformly random unitary matrices are clearly not identical; however, they are closely related. For instance, both matrices have uniformly distributed null spaces with respect to the Haar measure. This ensures that, the two matrices have statistically identical behavior for the purposes of the noiseless linear inverse problem (3.8), [25].

We observe  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \sigma\mathbf{v}$  where  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ . This is the standard setup for noisy linear inverse problems and it is used in numerous papers including [29, 30, 43, 73–75]. In order to estimate  $\mathbf{x}_0$  from  $\mathbf{y}$ , we will be using the following generalized setup,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \mathbf{x} \in \mathcal{C}. \quad (3.9)$$

We can again relate this problem to the classic LASSO when  $\mathbf{x}_0$  is sparse and when we choose  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 \leq \|\mathbf{x}_0\|_1\}$ . When  $f(\cdot)$  is an arbitrary function, we can use  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ .

We have the following result that relates optimal LASSO cost and the LASSO error to  $\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$  and also  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ .

**Theorem 3.3.** *Assume  $\mathcal{C}$  is a nonempty, closed and convex set and  $\mathbf{x}_0 \in \mathcal{C}$ . Assume  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is a partial unitary matrix generated uniformly at random. Let the noise vector  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$  be independent of  $\mathbf{A}$  and  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \sigma\mathbf{v}$ . Denote the minimizer of (3.9) by  $\mathbf{x}^* = \mathbf{x}^*(\mathbf{y}, \mathbf{A})$ . Conditioned on  $\mathbf{A}$ , define the normalized LASSO cost and “projected LASSO error” as follows,*

$$F_{LASSO}(\mathbf{A}) = \lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{y} - \mathbf{A}\mathbf{x}^*\|_2^2]}{\sigma^2}, \quad \eta_{LASSO}(\mathbf{A}) = \max_{\sigma > 0} \frac{\mathbb{E}[\|\mathbf{A}\mathbf{x}^* - \mathbf{A}\mathbf{x}_0\|_2^2]}{\sigma^2},$$

where the expectation is over  $\mathbf{v}$ . Then, there exists constants  $c_1, c_2 > 0$  such that,

- Whenever  $m < \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , with probability  $1 - c_1 \exp(-c_2 \frac{(m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*))^2}{n})$ ,

$$\eta_{LASSO}(\mathbf{A}) = m \quad \text{and} \quad F_{LASSO}(\mathbf{A}) = 0.$$

- Whenever  $m > \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , with probability  $1 - c_1 \exp(-c_2 t^2)$ ,

$$\begin{aligned} \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - t\sqrt{n} &\leq \eta_{LASSO}(\mathbf{A}) \leq \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) + t\sqrt{n}, \\ m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - t\sqrt{n} &\leq F_{LASSO}(\mathbf{A}) \leq m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) + t\sqrt{n} \end{aligned}$$

Here, the probabilities are over the random measurement matrix  $\mathbf{A}$ .

This theorem suggests that,  $\eta_{LASSO}$  and  $F_{LASSO}$  has a phase transition around the point  $m \approx \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ ; which not only shows up in (3.4) but also corresponds to the phase transition point of (3.8) (setting  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ ). When  $m < \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , one cannot recover  $\mathbf{x}_0$ , even from the noiseless observations  $\mathbf{A}\mathbf{x}_0$ ; as a result, it is futile to expect noise robustness. In this regime, our result suggests,  $\eta_{LASSO} = m$ . When the number of measurements are more than the phase transition point, interestingly,  $\eta_{LASSO}$  stops growing proportionally to  $m$  and takes a value around  $\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$  depending on the particular realization of  $\mathbf{A}$ .

While this is an interesting phenomenon, we should emphasize that, the more critical question is  $\frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2}$  rather than the projection of the error term on  $\mathbf{A}$ . This problem has been investigated for Gaussian measurement matrices and for  $\ell_1$ -minimization in a series of work, [29, 30, 76, 77]. However, to the best of our knowledge, Theorem 3.3 is the first result, that relates the LASSO cost and error to the convex geometry of the problem in a general framework.

### 3.5 Organization of the paper

The rest of this work is dedicated to the technical aspects of the theorems stated in this section. A summary of the consequent sections are as follows.

- **Section 4, Upper bounding the error:** We will start by introducing the optimality conditions for the problem (3.6). Then, we will find a tight upper bound to the resulting MSE. This will prove Theorem 3.2.
- **Section 5, Lower bounding the error:** We will restrict our attention to the analysis of the proximity operator (1.3). As  $\sigma \rightarrow 0$ , we will find a lower bound; which is arbitrarily close to the upper bound. This will prove Theorem 1.1. We use a similar argument for the constrained problem described in Theorem 3.1.
- **Section 6, Further discussion:** Following from (3.5), we discuss the relation between *optimally tuned* (over  $\lambda$ ) “regularized estimator” (1.3) and the “constrained estimator” (3.2).
- **Section 7, Connection to LASSO:** The proof of Theorem 3.3 will be provided. This proof requires some technical details related to the intersection of a random subspace and a cone. In particular, we will characterize certain properties of the intersection by using a modification of the results provided in [25]. We will also give examples to illustrate generality of our results.

## 4 Upper Bound and Its Interpretation

For the rest of the paper, we will make the following assumptions.

- $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function.
- $\mathcal{C} \subset \mathbb{R}^n$  is a nonempty, convex and closed set.
- $\mathbf{x}_0 \in \mathcal{C}$ .

**Notation:** When it is clear from context, the minimizer of an optimization over  $\mathbf{x}$  will be denoted by  $\mathbf{x}^*$ . Following from Section 3.1, for a closed and convex set  $\mathcal{C}$ , the “distance vector” from  $\mathcal{C}$  to  $\mathbf{x}$  is defined as  $\mathbf{x} - \text{Proj}(\mathbf{x}, \mathcal{C})$  and is denoted by  $\Pi(\mathbf{x}, \mathcal{C})$ . Since projection is the nearest point,  $\|\Pi(\mathbf{x}, \mathcal{C})\|_2 = \text{dist}(\mathbf{x}, \mathcal{C})$ .

To distinguish the deterministic analysis from stochastic analysis, we will allocate  $\tau$  and solve,

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{C}} \tau f(\mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 \quad (4.1)$$

when the noise is deterministic. When noise is distributed as  $\mathcal{N}(0, \sigma^2 \mathbf{I})$ , we will set  $\tau = \sigma \lambda$  and  $\lambda$  will correspond to the penalty for stochastic noise.

### 4.1 Upper bounding the error

**Lemma 4.1** (Optimality conditions). *Let  $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$  and assume  $f(\cdot)$  is a convex function and  $\mathcal{C}$  is a convex and closed set containing  $\mathbf{x}_0$ . Let  $\tau \geq 0$ . Consider the problem (4.1).  $\mathbf{x}^*$  is the unique optimal solution if and only if there exists  $\mathbf{s} \in \partial f(\mathbf{x}^*)$  and  $\mathbf{u} \in T_{\mathcal{C}}(\mathbf{x}^*)^*$  such that:*

$$\tau \mathbf{s} + \mathbf{u} + \mathbf{x}^* = \mathbf{y} \quad (4.2)$$

*Alternatively, we can state (4.2) in terms of the error vector  $\mathbf{w}^* = \mathbf{x}^* - \mathbf{x}_0$  and the noise  $\mathbf{z}$ , as follows.*

$$\tau \mathbf{s} + \mathbf{u} + \mathbf{w}^* = \mathbf{z} \quad (4.3)$$

*Proof.* The fact that “ $\mathbf{x}^*$  is the optimal solution if and only if (4.2) holds” follows from Proposition 4.7.2 of [86] and considering the subdifferential of the objective  $\tau f(\mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ . On the other hand, uniqueness follows from the strict convexity of the objective function of the proximity operator (4.1).  $\square$

The next lemma provides an upper bound on the  $\ell_2$  norm of the error term  $\mathbf{w}^* = \mathbf{x}^* - \mathbf{x}_0$ .

**Lemma 4.2.** *Consider the problem (4.1). Let  $\mathbf{w}^* = \mathbf{x}^* - \mathbf{x}_0$ . Assuming the setup of Lemma 4.1, we have:*

$$\|\mathbf{w}^*\|_2 \leq \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*).$$

*Proof.* Before starting the proof we emphasize that  $\tau\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*$  is closed due to the fact that  $\tau\partial f(\mathbf{x}_0)$  is compact and  $T_{\mathcal{C}}(\mathbf{x}_0)^*$  is closed. From the optimality conditions, there exists  $\mathbf{s} \in \partial f(\mathbf{x}^*)$  and  $\mathbf{u} \in T_{\mathcal{C}}(\mathbf{x}^*)^*$  such that (4.3) holds. Now, let  $\mathcal{K} := \mathcal{K}(\mathcal{C}, f, \mathbf{x}_0) = \tau\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*$ . Further, let  $\mathbf{s}_0 \in \partial f(\mathbf{x}_0)$ ,  $\mathbf{u}_0 \in T_{\mathcal{C}}(\mathbf{x}_0)^*$  and  $\mathbf{w}_0 = \mathbf{z} - \tau\mathbf{s}_0 - \mathbf{u}_0$ . We will first explore the relation between the dual vectors and  $\mathbf{w}^*$ . The following inequality follows from the standard properties of the subgradients.

$$(\mathbf{w}^*)^T \mathbf{s} \geq f(\mathbf{x}_0 + \mathbf{w}^*) - f(\mathbf{x}_0) \geq (\mathbf{w}^*)^T \mathbf{s}_0 \implies (\mathbf{w}^*)^T (\mathbf{s} - \mathbf{s}_0) \geq 0 \quad (4.4)$$

Since  $\mathbf{u} \in T_{\mathcal{C}}(\mathbf{x}^*)^*$  and  $-\mathbf{w}^* \in T_{\mathcal{C}}(\mathbf{x}^*)$ , we have:

$$\langle \mathbf{u}, \mathbf{w}^* \rangle \geq 0 \quad (4.5)$$

Similarly,  $\mathbf{u}_0 \in T_{\mathcal{C}}(\mathbf{x}_0)^*$  and  $\mathbf{w}^* \in T_{\mathcal{C}}(\mathbf{x}_0)$ , hence:

$$\langle \mathbf{u}_0, \mathbf{w}^* \rangle \leq 0 \quad (4.6)$$

Overall, combining (4.4), (4.5) and (4.6), we find:

$$\langle \mathbf{w}^*, (\mathbf{u} + \tau\mathbf{s}) - (\mathbf{u}_0 + \tau\mathbf{s}_0) \rangle \geq 0 \quad (4.7)$$

From (4.7), we will conclude that  $\|\mathbf{w}^*\|_2 \leq \|\mathbf{w}_0\|_2$ . (4.7) is equivalent to:

$$\langle \mathbf{w}^*, (\mathbf{z} - \mathbf{w}^*) - (\mathbf{z} - \mathbf{w}_0) \rangle = \langle \mathbf{w}^*, \mathbf{w}_0 - \mathbf{w}^* \rangle \geq 0 \implies \|\mathbf{w}^*\|_2^2 \leq \langle \mathbf{w}^*, \mathbf{w}_0 \rangle \leq \|\mathbf{w}^*\|_2 \|\mathbf{w}_0\|_2 \quad (4.8)$$

Hence, we indeed have:  $\|\mathbf{w}_0\|_2 \geq \|\mathbf{w}^*\|_2$ . Since this is true for all  $\mathbf{s}_0, \mathbf{u}_0$ , we can choose the  $\mathbf{w}_0$  with the shortest length, namely, choose  $\mathbf{w}_0$  to be the distance vector  $\mathbf{w}_0 = \Pi(\mathbf{z}, T_{\mathcal{C}}(\mathbf{x}_0)^* + \tau\partial f(\mathbf{x}_0))$  to conclude.  $\square$

As a direct application of Lemma 4.2, we can upper bound the expected errors as follows when  $\mathbf{z}$  is stochastic.

**Proposition 4.1.** *Let  $\sigma > 0, \lambda \geq 0$ . In problem (4.1), let  $\mathbf{z} = \sigma\mathbf{v}$ , where  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ . Then,*

- Denote the minimizer of (4.1) by  $\mathbf{x}^* = \mathbf{x}^*(\tau, \mathbf{z})$ . Letting  $\tau = \sigma\lambda$ ,

$$\frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2} \leq \mathbf{D}(\lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*). \quad (4.9)$$

- Let  $\lambda^* = \arg \min_{\lambda \geq 0} \mathbf{D}(\lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*)$ . By choosing  $\tau = \lambda^*\sigma$  in (4.1) and using (4.9) we have,

$$\frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2} \leq \min_{\lambda \geq 0} \mathbf{D}(\lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*) \quad (4.10)$$

- Setting  $\mathcal{C} = \mathbb{R}^n$  in (4.9), we show  $\leq$ -direction of Theorem 1.1.
- Setting  $f(\cdot) = 0$  in (4.9), we show  $\leq$ -direction of Theorem 3.1.

*Proof.* Let  $\tau = \sigma\lambda$  and  $\mathbf{w}^*(\tau, \mathbf{z}) = \mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0$ . From Lemma 4.2, the noise in (4.1) satisfies:

$$\|\mathbf{w}^*(\tau, \mathbf{z})\|_2 \leq \text{dist}(\mathbf{z}, \sigma\lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*) = \text{dist}(\sigma\mathbf{v}, \sigma\lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*).$$

Taking the squares of both sides and normalizing by  $\sigma^2$ , we find,

$$\frac{\|\mathbf{w}^*(\tau, \mathbf{z})\|_2^2}{\sigma^2} \leq \text{dist}(\mathbf{v}, \lambda\partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*)^2.$$

Now, taking the expectations of both sides and using  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ , we end up with the first statement (4.9). Second statement follows immediately. For third statement, observe that if  $\mathcal{C} = \mathbb{R}^n$  (i.e. no constraint) then  $F_{\mathcal{C}}(\mathbf{x}_0) = \mathbb{R}^n$  and  $T_{\mathcal{C}}(\mathbf{x}_0)^* = \{0\}$ . Finally, if  $f(\cdot) = 0$  (i.e. no penalization),  $\partial f(\mathbf{x}_0) = \lambda\partial f(\mathbf{x}_0) = \{0\}$ .  $\square$

## 4.2 First order approximation

In this section, we will consider the first order approximation of the proximal denoising problem around  $\mathbf{x}_0$  and we will argue that, the error term of the first order problem is exactly same as the upper bound provided in Lemma 4.2. This will provide a good intuition for our consecutive results, which argue that the least favorable noise distribution that maximizes normalized MSE occurs as  $\sigma \rightarrow 0$ .

**Notation:** Given  $\mathbf{w}, \mathbf{x}_0, f(\cdot), \epsilon > 0$ , the ratio  $\frac{f(\mathbf{x}_0 + \epsilon \mathbf{w}) - f(\mathbf{x}_0)}{\epsilon}$  is nondecreasing as a function of  $\epsilon$ . Furthermore, the “directional derivative” along  $\mathbf{w}$  is defined as [84, 88],

$$\lim_{\epsilon \rightarrow 0} \frac{f(\mathbf{x}_0 + \epsilon \mathbf{w}) - f(\mathbf{x}_0)}{\epsilon} = f'(\mathbf{x}_0, \mathbf{w}) \quad (4.11)$$

It is known that, [88], there is a trivial relation between the directional derivative and the subgradients,

$$f'(\mathbf{x}_0, \mathbf{w}) = \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s} \rangle. \quad (4.12)$$

We will extensively make use of this equality later on. Finally, let us denote the “set of maximizing subgradients” of  $f(\cdot)$  at  $\mathbf{x}_0$  along  $\mathbf{w}$ , via  $\partial f(\mathbf{x}_0, \mathbf{w})$ . In particular,

$$\partial f(\mathbf{x}_0, \mathbf{w}) := \{\mathbf{s} \in \partial f(\mathbf{x}_0) \mid \mathbf{s}^T \mathbf{w} = \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0)} (\mathbf{s}')^T \mathbf{w}\} \quad (4.13)$$

$\partial f(\mathbf{x}_0, \mathbf{w})$  basically corresponds to a face of the subdifferential along  $\mathbf{w}$ . Observe that, the supremum on the right hand side is the directional derivative of  $f(\cdot)$  at  $\mathbf{x}_0$  along  $\mathbf{w}$ .

### 4.2.1 Approximated problem

- **Approximating the function  $f(\cdot)$  at  $\mathbf{x}_0$ :** We will use the first order approximation  $\hat{f}_{\mathbf{x}_0}(\cdot)$  of  $f(\cdot)$  at  $\mathbf{x}_0$ , which is achieved via directional derivatives. In particular, we will let:

$$\hat{f}_{\mathbf{x}_0}(\mathbf{x}) = f(\mathbf{x}_0) + \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \langle \mathbf{s}, \mathbf{x} - \mathbf{x}_0 \rangle \quad (4.14)$$

It is known that,  $f(\cdot)$  behaves like  $\hat{f}_{\mathbf{x}_0}(\cdot)$  in the close proximity of  $\mathbf{x}_0$ , [86]. (4.14) can be viewed as a generalization of the first order approximation of differentiable functions. Furthermore, observe that,  $\hat{f}_{\mathbf{x}_0}(\mathbf{x}) \leq f(\mathbf{x})$  for all  $\mathbf{x}$  due to convexity and the subgradient property.

- **Approximating the set  $\mathcal{C}$  at  $\mathbf{x}_0$ :** We will find an approximation of  $\mathcal{C}$  by considering the tangent cone of  $\mathcal{C}$  at  $\mathbf{x}_0$ . This follows from the fact that  $T_{\mathcal{C}}(\mathbf{x}_0)$  provides a good approximation of the set of feasible perturbations  $F_{\mathcal{C}}(\mathbf{x}_0)$  around 0 as we have  $T_{\mathcal{C}}(\mathbf{x}_0) = \text{Cl}(\text{cone}(F_{\mathcal{C}}(\mathbf{x}_0)))$ . Hence, the approximation to the set  $\mathcal{C}$  will be:

$$\hat{\mathcal{C}}_{\mathbf{x}_0} = \mathbf{x}_0 + T_{\mathcal{C}}(\mathbf{x}_0) \quad (4.15)$$

- **The approximated problem around  $\mathbf{x}_0$ :** Approximated problem is obtained by approximating the function and the set simultaneously and considering (4.1) with the new function and set  $\hat{f}, \hat{\mathcal{C}}$ .

$$\min_{\mathbf{x} \in \hat{\mathcal{C}}_{\mathbf{x}_0}} \tau \hat{f}_{\mathbf{x}_0}(\mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 \quad (4.16)$$

As it is argued in Lemma B.1, the first order approximation of  $f(\cdot)$  is a convex function and hence (4.16) is a convex problem.

The nice property of the problem (4.16) is the fact that, it can be solved exactly. It also provides an insight about what to expect as the minimax risk. We will now show that, the error in the approximated problem is exactly equal to the upper bound given in Lemma 4.2. Hence, the error in the original problem is always upper bounded by the error in the approximated problem. However, intuitively, the approximation will be tight when the noise variance is sufficiently small, hence when the variance  $\sigma$  is small, the error in the original problem will be approximately equal to the upper bound.

Before analysis of the approximated problem, we will provide a lemma which shows that the subgradients of  $\hat{f}_{\mathbf{x}_0}$  can be found exactly. The proof can be found in the Appendix B.

**Lemma 4.3.** Let  $\hat{f}_{\mathbf{x}_0}$  be same as in (4.14). For any  $\mathbf{x}$ , we have,

$$\partial \hat{f}_{\mathbf{x}_0}(\mathbf{x}) = \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) \quad (4.17)$$

Basically, the subgradients of  $\hat{f}_{\mathbf{x}_0}$  are the maximizing subgradients of  $f(\cdot)$ , i.e., the set of subgradients that maximizes the inner product with  $\mathbf{x} - \mathbf{x}_0$ .

The next result, characterizes a useful property of the projection on a convex set, [86].

**Lemma 4.4.** Let  $\mathcal{K}$  be a closed and convex set and  $\mathbf{v}$  be an arbitrary vector. We have,

$$\langle \text{Proj}(\mathbf{v}, \mathcal{K}), \Pi(\mathbf{v}, \mathcal{K}) \rangle = \sup_{\mathbf{u} \in \mathcal{K}} \langle \mathbf{u}, \Pi(\mathbf{v}, \mathcal{K}) \rangle$$

In words, the projection vector maximizes the inner product with the distance vector over  $\mathcal{K}$ .

We have the following result regarding the first order approximated problem (4.16).

**Proposition 4.2** (Solution for the approximated problem). Let  $\hat{f}_{\mathbf{x}_0}$  and  $\hat{\mathcal{C}}$  be same as (4.14) and (4.15) respectively. Let  $\hat{\mathbf{x}} = \hat{\mathbf{x}}(\tau, \mathbf{z})$  denote the optimal solution of (4.16). Then, we have:

$$\hat{\mathbf{x}} - \mathbf{x}_0 = \Pi(\mathbf{z}, \tau \partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*) \quad (4.18)$$

*Proof.* Let  $\hat{\mathbf{w}} = \hat{\mathbf{x}} - \mathbf{x}_0$ . From optimality conditions,  $\hat{\mathbf{x}}$  is the unique minimizer of (4.16) if and only if, it satisfies,

$$\tau \mathbf{s} + \mathbf{u} + \hat{\mathbf{w}} = \mathbf{z} \quad (4.19)$$

where  $\mathbf{s} \in \partial \hat{f}_{\mathbf{x}_0}(\hat{\mathbf{x}})$  and  $\mathbf{u} \in T_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})^*$ . Using Lemma 4.3,  $\mathbf{s} \in \partial \hat{f}_{\mathbf{x}_0}(\hat{\mathbf{x}})$  is identical to  $\mathbf{s} \in \partial f(\mathbf{x}_0, \hat{\mathbf{x}} - \mathbf{x}_0)$ .

Assume (4.18) holds, then  $\mathbf{z} - \hat{\mathbf{w}} = \tau \mathbf{s} + \mathbf{u}$  for some  $\mathbf{u} \in T_{\mathcal{C}}(\mathbf{x}_0)^*$  and  $\mathbf{s} \in \partial f(\mathbf{x}_0)$ . We will show that, (4.19) holds for this particular  $\hat{\mathbf{w}}, \mathbf{s}, \mathbf{u}$  and  $\hat{\mathbf{x}}$  is indeed the minimizer of (4.16). Since  $\mathbf{z} - \mathbf{u}$  is the closest point to  $\tau \partial f(\mathbf{x}_0)$ , we have:

$$\hat{\mathbf{w}} = \Pi(\mathbf{z} - \mathbf{u}, \tau \partial f(\mathbf{x}_0)), \quad \tau \mathbf{s} = \text{Proj}(\mathbf{z} - \mathbf{u}, \tau \partial f(\mathbf{x}_0))$$

In conjunction with Lemma 4.4, this implies  $\langle \hat{\mathbf{w}}, \mathbf{s} \rangle = \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0)} \langle \mathbf{s}', \hat{\mathbf{w}} \rangle$  which in turn implies  $\mathbf{s} \in \partial f(\mathbf{x}_0, \hat{\mathbf{x}} - \mathbf{x}_0) = \partial \hat{f}_{\mathbf{x}_0}(\hat{\mathbf{x}})$  based on Lemma 4.3. In a similar manner, we have:

$$\hat{\mathbf{w}} = \Pi(\mathbf{z} - \tau \mathbf{s}, T_{\mathcal{C}}(\mathbf{x}_0)^*), \quad \mathbf{u} = \text{Proj}(\mathbf{z} - \tau \mathbf{s}, T_{\mathcal{C}}(\mathbf{x}_0)^*)$$

Using Moreau's decomposition theorem (see Lemma A.3), this implies that  $\hat{\mathbf{w}} \in T_{\mathcal{C}}(\mathbf{x}_0)$  and  $\langle \hat{\mathbf{w}}, \mathbf{u} \rangle = 0$ . Now, we will argue that  $\mathbf{u} \in T_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})^*$  to conclude. Let  $\mathbf{v} \in F_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})$ . This implies  $\mathbf{v} + \mathbf{x}_0 + \hat{\mathbf{w}} \in \hat{\mathcal{C}}_{\mathbf{x}_0}$  or alternatively,  $\mathbf{v} + \hat{\mathbf{w}} \in T_{\mathcal{C}}(\mathbf{x}_0)$ . Hence,

$$\langle \mathbf{v} + \hat{\mathbf{w}}, \mathbf{u} \rangle \leq 0$$

Using  $\langle \hat{\mathbf{w}}, \mathbf{u} \rangle = 0$ , we find for all  $\mathbf{v} \in F_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})$ ,  $\langle \mathbf{v}, \mathbf{u} \rangle \leq 0$ . This will still be true when we take the closure and the cone of the set  $F_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}}(\tau, \mathbf{z}))$  which gives  $\langle \mathbf{v}, \mathbf{u} \rangle \leq 0$  for all  $\mathbf{v} \in T_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})$  and implies  $\mathbf{u}$  is in the polar cone of  $T_{\hat{\mathcal{C}}_{\mathbf{x}_0}}(\hat{\mathbf{x}})$ .  $\square$

This shows that the approximated problem (4.16) can be solved exactly and its error exactly matches with the upper bound found in Lemma 4.2. This observation provides a good interpretation of Lemma 4.2.

The next corollary follows immediately from Proposition 4.2 and gives MSE results under stochastic noise. The proof is based on taking the expectation of (4.18) over  $\mathbf{z}$ .

**Corollary 4.1** (Approximated MSE). Assume  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ . Let  $\mathbf{z} = \sigma \mathbf{v}$  in the approximated problem (4.16) and let  $\tau = \sigma \lambda$ . Denote the minimizer by  $\hat{\mathbf{x}}(\lambda, \mathbf{z})$ . Then,

$$\mathbb{E}[\|\hat{\mathbf{x}}(\lambda, \mathbf{z}) - \mathbf{x}_0\|_2^2] = \sigma^2 \mathbb{E}[\text{dist}(\mathbf{v}, \lambda \partial f(\mathbf{x}_0) + T_{\mathcal{C}}(\mathbf{x}_0)^*)^2].$$

The next section is dedicated to showing that the bounds we have provided so far are in fact sharp. In particular, as  $\sigma \rightarrow 0$ , the original problem (4.1) behaves same as the approximated problem (4.16).

## 5 Tight lower bound when $\sigma \rightarrow 0$

We will now focus our attention to the  $\geq$ -direction of Theorem 1.1 and hence the regularized estimator (no constraint),

$$\mathbf{x}^*(\tau, \mathbf{z}) = \arg \min \tau f(\mathbf{x}) + \frac{1}{2} \|(\mathbf{x}_0 + \mathbf{z}) - \mathbf{x}\|_2^2 \quad (5.1)$$

We will show that, one can find a good lower bound for the estimation error  $\|\mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0\|_2$  in (5.1). The following proposition states that  $f(\cdot)$  can be approximated by  $\hat{f}_{\mathbf{x}_0}$  (recall (4.14)) around  $\mathbf{x}_0$  along all directions simultaneously. This result is a stronger version of (4.12) and is due to Lemma 2.1.1 of Chapter VI of [88].

**Proposition 5.1.** *Assume  $f(\cdot)$  is a convex function on  $\mathbb{R}^n$ . Then, for any given  $\delta_0 > 0$ , there exists some  $\epsilon_0 > 0$  such that, for all  $\mathbf{w} \in \mathbb{R}^n$ , with  $\|\mathbf{w}\|_2 \leq \epsilon_0$ , the following holds:*

$$0 \leq f(\mathbf{x}_0 + \mathbf{w}) - \hat{f}_{\mathbf{x}_0}(\mathbf{x}_0 + \mathbf{w}) \leq \delta_0 \|\mathbf{w}\|_2 \quad (5.2)$$

While it is clear that  $\hat{f}_{\mathbf{x}_0}(\cdot)$  is an under estimator of  $f(\cdot)$  at all points, upper bounding  $f(\cdot)$  in terms of  $\hat{f}_{\mathbf{x}_0}(\cdot)$  and a small first order term will be quite helpful for the perturbation analysis.

### 5.1 Deterministic lower bound

The following lemma provides a deterministic lower bound on the error  $\mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0$ . Unlike the previous sections, there will be no set constrained (i.e.  $\mathcal{C}$ ) in the consequent analysis.

**Lemma 5.1.** *Consider the problem (5.1). Let  $\mathbf{w}_0 := \Pi(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$ . Assume  $\tau \leq C_0 \|\mathbf{w}_0\|_2$  for some constant  $C_0 > 0$ . Using Proposition 5.1, choose  $\delta_0, \epsilon_0 > 0$  ensuring (5.2). Then, whenever  $\|\mathbf{w}_0\|_2 \leq \epsilon_0$ ,*

$$\|\mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0\|_2 \geq (1 - \sqrt{2\delta_0 C_0}) \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0)).$$

*Proof.* To begin with, observe that,  $\|\mathbf{w}_0\|_2 = \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$ . Let  $\mathbf{s}_0 = \frac{\text{Proj}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))}{\tau}$ . From Lemma 4.4,  $\mathbf{s}_0 \in \partial f(\mathbf{x}_0, \mathbf{w}_0)$  where  $\partial f(\mathbf{x}_0, \mathbf{w}_0)$  is given by (4.13). Let us rewrite the objective function of (5.1) as a function of the perturbation  $\mathbf{w} = \mathbf{x} - \mathbf{x}_0$ .

$$g(\mathbf{w}) = \tau(f(\mathbf{x}_0 + \mathbf{w}) - f(\mathbf{x}_0)) + \frac{1}{2} \|\mathbf{z} - \mathbf{w}\|_2^2$$

Using the facts that  $\|\mathbf{w}_0\|_2 \leq \epsilon_0$ ,  $\tau \leq C_0 \|\mathbf{w}_0\|_2$  and  $\mathbf{w}_0 + \tau \mathbf{s}_0 = \mathbf{z}$ , we have the following,

$$\begin{aligned} g(\mathbf{w}_0) &= \tau(f(\mathbf{x}_0 + \mathbf{w}_0) - f(\mathbf{x}_0)) + \frac{1}{2} \|\mathbf{z} - \mathbf{w}_0\|_2^2 \\ &= \tau(\langle \mathbf{s}_0, \mathbf{w}_0 \rangle + (f(\mathbf{x}_0 + \mathbf{w}_0) - \hat{f}_{\mathbf{x}_0}(\mathbf{x}_0 + \mathbf{w}_0))) + \frac{1}{2} \|\mathbf{z} - \mathbf{w}_0\|_2^2 \\ &\leq \tau(\langle \mathbf{s}_0, \mathbf{w}_0 \rangle + \delta_0 \|\mathbf{w}_0\|_2) + \frac{1}{2} \|\mathbf{w}_0\|_2^2 - \langle \mathbf{z}, \mathbf{w}_0 \rangle + \frac{1}{2} \|\mathbf{z}\|_2^2 \\ &\leq C_0 \delta_0 \|\mathbf{w}_0\|_2^2 + \frac{1}{2} \|\mathbf{w}_0\|_2^2 - \langle \mathbf{z} - \tau \mathbf{s}_0, \mathbf{w}_0 \rangle + \frac{1}{2} \|\mathbf{z}\|_2^2 \\ &= (C_0 \delta_0 - \frac{1}{2}) \|\mathbf{w}_0\|_2^2 + \frac{1}{2} \|\mathbf{z}\|_2^2 \\ &= (C_0 \delta_0 - \frac{1}{2}) \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))^2 + \frac{1}{2} \|\mathbf{z}\|_2^2. \end{aligned} \quad (5.3)$$

Letting  $\mathbf{w}^*(\tau, \mathbf{z}) = \mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0$ ,  $g(\mathbf{w}^*(\tau, \mathbf{z}))$  should be smaller than the right hand side of (5.3) as  $\mathbf{x}^*(\tau, \mathbf{z})$  should outperform  $\mathbf{x}_0 + \mathbf{w}_0$ . Choose  $\mathbf{s}^* \in \partial f(\mathbf{x}_0, \mathbf{x}^*(\tau, \mathbf{z}) - \mathbf{x}_0)$ . Then,  $g(\mathbf{w}^*(\tau, \mathbf{z}))$  can be lower bounded as follows:

$$\begin{aligned} g(\mathbf{w}^*(\tau, \mathbf{z})) &\geq \tau \langle \mathbf{s}^*, \mathbf{w}^*(\tau, \mathbf{z}) \rangle + \frac{1}{2} \|\mathbf{z} - \mathbf{w}^*(\tau, \mathbf{z})\|_2^2 \\ &= \tau \langle \mathbf{s}^*, \mathbf{w}^*(\tau, \mathbf{z}) \rangle - \langle \mathbf{z}, \mathbf{w}^*(\tau, \mathbf{z}) \rangle + \frac{1}{2} \|\mathbf{w}^*(\tau, \mathbf{z})\|_2^2 + \frac{1}{2} \|\mathbf{z}\|_2^2 \\ &= -\langle \mathbf{w}^*(\tau, \mathbf{z}), \mathbf{z} - \tau \mathbf{s}^* \rangle + \frac{1}{2} \|\mathbf{w}^*(\tau, \mathbf{z})\|_2^2 + \frac{1}{2} \|\mathbf{z}\|_2^2 \end{aligned} \quad (5.4)$$

Now, using  $\mathbf{s}^* \in \partial f(\mathbf{x}_0, \mathbf{w}^*(\tau, \mathbf{z}))$ ,

$$\langle \mathbf{w}^*(\tau, \mathbf{z}), \mathbf{s}^* \rangle \geq \langle \mathbf{w}^*(\tau, \mathbf{z}), \mathbf{s}_0 \rangle \implies -\langle \mathbf{w}^*(\tau, \mathbf{z}), \mathbf{z} - \tau \mathbf{s}^* \rangle \geq -\langle \mathbf{w}^*(\tau, \mathbf{z}), \mathbf{z} - \tau \mathbf{s}_0 \rangle$$

Combining this and (5.4),  $g(\mathbf{w}^*(\tau, \mathbf{z}))$  satisfies:

$$g(\mathbf{w}^*(\tau, \mathbf{z})) \geq \frac{1}{2} \|\mathbf{w}^*(\tau, \mathbf{z})\|_2^2 - \|\mathbf{w}^*(\tau, \mathbf{z})\|_2 \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0)) + \frac{1}{2} \|\mathbf{z}\|_2^2 \quad (5.5)$$

From Lemma 4.2, it is known that  $\|\mathbf{w}^*(\tau, \mathbf{z})\|_2 \leq \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$ . Now, calling  $a = \|\mathbf{w}^*(\tau, \mathbf{z})\|_2$ ,  $b = \text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$ , and  $t = 2C_0\delta_0$ ,  $g(\mathbf{w}^*(\tau, \mathbf{z})) \leq g(\mathbf{w}_0)$  implies,

$$\frac{a^2}{2} - ab \leq \left(\frac{t-1}{2}\right)b^2.$$

Setting  $u = a/b$  and using  $1 \geq u, t \geq 0$ , we have:

$$ub^2 - u^2b^2/2 \geq (1-t)b^2/2 \implies 2u - u^2 \geq (1-t) \implies u^2 - 2u + (1-t) \leq 0$$

Overall,  $1 \geq u \geq 1 - \sqrt{t}$ . Equivalently,  $\|\mathbf{w}^*(\tau, \mathbf{z})\|_2 \geq (1 - \sqrt{2C_0\delta_0})\text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$ .  $\square$

This lemma shows that, when  $\mathbf{z}$  is sufficiently small, then the error  $\|\mathbf{w}^*\|_2$  will be close to its upper bound given in Lemma 4.2. Observe that, we assumed  $\tau \leq C_0\text{dist}(\mathbf{z}, \tau \partial f(\mathbf{x}_0))$  for some  $C_0 > 0$ . The setup of Theorem 1.1 already prepares the background for this assumption as it chooses  $\tau = \sigma\lambda$  and  $\mathbf{z} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ . Together, these will ensure  $\mathbf{z}, \tau$  is approximately proportional and Lemma 4.2 is applicable, with high probability, for sufficiently large  $C_0$ . The choice  $\tau = \sigma\lambda$  for a fixed  $\lambda$ , is in fact a well-known property of the soft thresholding operator, [4, 30, 35, 40, 73, 74].

## 5.2 Stochastic lower bound for regularized NMSE

To finish the proof of Theorem 1.1, we need to show that the upper bound can be achieved as  $\sigma \rightarrow 0$ . This will be the topic of this section.

**Proposition 5.2.** *Consider program (5.1) with  $\tau = \sigma\lambda$  and  $\mathbf{z} = \sigma\mathbf{v}$  where  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I}_n)$ . Then, we have:*

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{x}^*(\sigma\lambda, \sigma\mathbf{v}) - \mathbf{x}\|_2^2]}{\sigma^2} = \mathbf{D}(\lambda \partial f(\mathbf{x}_0))$$

*Proof.* To begin with, observe that  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0)) > 0$  as  $\lambda \partial f(\mathbf{x}_0)$  is a compact set. Let  $\mathbf{w}^* = \mathbf{x}^*(\sigma\lambda, \sigma\mathbf{v}) - \mathbf{x}$ . Denote the probability density function of an  $\mathcal{N}(0, c\mathbf{I})$  distributed vector by  $p_c(\cdot)$ . We can write:

$$\mathbb{E}[\|\mathbf{w}^*\|_2^2] = \int_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{w}^*\|_2^2 p_\sigma(\mathbf{z}) d\mathbf{z}$$

We will split the error term into three parts where the integration is performed over the following sets.

- $S_1 = \{\mathbf{a} \in \mathbb{R}^n \mid \|\mathbf{a}\|_2 \geq C_1\sigma\}$ .
- $S_2 = \{\mathbf{a} \in \mathbb{R}^n \mid \|\mathbf{a}\|_2 \leq C_1\sigma, \text{dist}(\mathbf{a}, \lambda\sigma \partial f(\mathbf{x}_0)) \geq \epsilon_1\sigma\}$ .
- $S_3 = \{\mathbf{a} \in \mathbb{R}^n \mid \|\mathbf{a}\|_2 \leq C_1\sigma, \text{dist}(\mathbf{a}, \lambda\sigma \partial f(\mathbf{x}_0)) \leq \epsilon_1\sigma\}$ .

where  $C_1, \epsilon_1$  are positive scalars to be determined later on. Recall that  $\mathbf{v} = \frac{\mathbf{z}}{\sigma}$  and define the associated restricted integrations  $I_i, \hat{I}_i$ , which are given as:

$$I_i = \int_{\mathbf{z} \in S_i} \|\mathbf{w}^*\|_2^2 p_\sigma(\mathbf{z}) d\mathbf{z}$$

$$\hat{I}_i = \int_{\sigma\mathbf{v} \in S_i} \text{dist}(\mathbf{v}, \lambda \partial f(\mathbf{x}_0))^2 p_1(\mathbf{v}) d\mathbf{z}$$

From Proposition 4.1,  $\mathbb{E}[\|\mathbf{w}^*(\sigma\lambda, \sigma\mathbf{v})\|_2^2] = I_1 + I_2 + I_3 \leq \sigma^2(\hat{I}_1 + \hat{I}_2 + \hat{I}_3)$ . To proceed,

- We will first argue that  $\hat{I}_1$  and  $\hat{I}_3$  are smaller compared to  $\hat{I}_2$ .
- Then, we will argue that  $I_2$  is close to  $\hat{I}_2$ .
- These will yield:

$$I_1 + I_2 + I_3 \geq I_2 \approx \hat{I}_2 \approx \hat{I}_1 + \hat{I}_2 + \hat{I}_3$$

**Claim:** We have,

$$\hat{I}_2 \geq \mathbf{D}(\lambda \partial f(\mathbf{x}_0)) - [\epsilon_1^2 + Q(C_1)] \quad (5.6)$$

where  $Q(C_1) \rightarrow 0$  as  $C_1 \rightarrow \infty$ .

*Proof.* Using the boundedness of  $\partial f(\mathbf{x}_0)$ , let  $R = R_{\mathbf{x}_0} = \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2$ .  $\text{dist}(\mathbf{v}, \lambda \partial f(\mathbf{x}_0))^2$  can be upper bounded as,

$$\text{dist}(\mathbf{v}, \lambda \partial f(\mathbf{x}_0))^2 \leq (\|\mathbf{v}\|_2 + \lambda R)^2 \leq 2(\|\mathbf{v}\|_2^2 + \lambda^2 R^2).$$

Observe that,  $\hat{I}_1$  is calculated over the region  $\|\mathbf{v}\|_2 \geq C_1$ . As  $C_1 \rightarrow \infty$ , since the multivariate Gaussian distribution have finite moments, we have  $\hat{I}_1 := Q(C_1) \rightarrow 0$ .

$\hat{I}_3$  can be upper bounded by the integration of  $\epsilon_1^2$  over  $\mathbb{R}^n$  which yields  $\hat{I}_3 \leq \epsilon_1^2$ . Finally, using,

$$\sum_{i=1}^3 \hat{I}_i = \mathbf{D}(\lambda \partial f(\mathbf{x}_0))$$

we have  $\hat{I}_2 \geq \mathbf{D}(\lambda \partial f(\mathbf{x}_0)) - \epsilon_1^2 - Q(C_1)$ . □

The following lemma finishes the proof by using Lemma 5.1.

**Lemma 5.2.** *Let  $R = \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2$ . Based on Proposition 5.1, for any  $\delta_0 > 0$ , there exists  $\epsilon_0 > 0$  so that,*

$$f(\mathbf{x}_0 + \mathbf{w}) - \hat{f}_{\mathbf{x}_0}(\mathbf{x}) \leq \delta_0 \|\mathbf{w}\|_2 \text{ for all } \|\mathbf{w}\|_2 \leq \epsilon_0. \quad (5.7)$$

Now, for any given (fixed)  $\epsilon_1, C_1, \delta_0, \epsilon_0$ , choose  $\sigma$  sufficiently small to ensure  $(C_1 + R\lambda)\sigma < \epsilon_0$ . Then, for all such  $\sigma$ 's,

$$I_2 \geq \sigma^2 (1 - \sqrt{2\delta_0 \lambda \epsilon_1^{-1}}) \hat{I}_2. \quad (5.8)$$

Following from (5.8) and letting  $\delta_0 \rightarrow 0$  (by choosing  $\epsilon_0$  properly) and keeping  $\sigma$  accordingly small, we obtain,

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{w}^*(\lambda, \mathbf{z})\|_2^2]}{\sigma^2 \hat{I}_2} \geq \lim_{\sigma \rightarrow 0} \frac{I_2}{\sigma^2 \hat{I}_2} \geq 1. \quad (5.9)$$

Since this is true for arbitrary  $\epsilon_1, C_1$  letting  $C_1 \rightarrow \infty$  and  $\epsilon_1 \rightarrow 0$  and using the relation (5.6), we can conclude,

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{w}^*(\lambda, \mathbf{z})\|_2^2]}{\sigma^2} \geq \mathbf{D}(\lambda \partial f(\mathbf{x}_0)). \quad (5.10)$$

*Proof.* To show the result, we will use Lemma 5.1. To apply Lemma 5.1 with  $\mathbf{w}_0 = \text{dist}(\mathbf{z}, \sigma \lambda \partial f(\mathbf{x}_0))$ , we will show that, all  $\mathbf{z} \in S_2$  satisfy the requirements.

- First, observe that, for any  $\mathbf{z} \in S_2$ , from triangle inequality, we have,

$$\text{dist}(\mathbf{z}, \sigma \lambda \partial f(\mathbf{x}_0)) \leq \sigma \lambda R + \sigma C_1 \leq (\lambda R + C_1) \sigma \leq \epsilon_0.$$

- Secondly, we additionally have,

$$\tau = \sigma \lambda \leq \frac{\lambda}{\epsilon_1} \text{dist}(\mathbf{z}, \sigma \lambda \partial f(\mathbf{x}_0)).$$

- Using these and  $(C_1 + \lambda R)\sigma < \epsilon_0$ , we can apply Lemma 5.1 for all  $\mathbf{z} \in S_2$  to obtain,

$$\|\mathbf{w}^*(\lambda, \mathbf{z})\|_2 \geq (1 - \sqrt{2\delta_0 \lambda \epsilon_1^{-1}}) \text{dist}(\mathbf{z}, \sigma \lambda \partial f(\mathbf{x}_0)) \quad (5.11)$$

Integrating over  $S_2$ , we find (5.8). Recall that,  $\sigma$  is under our control and all  $\mathbf{z} \in S_2 = S_2(\sigma)$  with  $\sigma < \frac{\epsilon_0}{\lambda R + C_1}$  satisfies (5.11); which is only a function of  $\delta_0$  when  $C_1, \epsilon_1$  are fixed. We can let  $\delta_0 \rightarrow 0$ , while keeping  $\sigma$  sufficiently small compared to  $\epsilon_0$ , to obtain, (5.9) after an integration over  $S_2$ .

(5.9) is true for any constants  $C_1, \epsilon_1 > 0$  as  $\sigma \rightarrow 0$ . Letting  $C_1 \rightarrow \infty, \epsilon_1 \rightarrow 0$  and using (5.6), we obtain,

$$\lim_{C_1 \rightarrow \infty, \epsilon_1 \rightarrow 0} \frac{\hat{I}_2}{\mathbf{D}(\lambda \partial f(\mathbf{x}_0))} = 1.$$

Finally, combining this with (5.9), and letting  $\sigma \rightarrow 0$  as a function of  $C_1 \rightarrow \infty, \epsilon_1 \rightarrow 0$ , we get the desired result (5.10). Hence, this shows the  $\geq$ -direction in Theorem 1.1 as  $\sigma \rightarrow 0$ .  $\square$

$\square$

## 6 Further discussion on proximal denoising

The remaining two things are the proofs for Theorem 3.1 and discussion of (3.5) which relates Theorem 3.1 and 1.1. We should remark that, for Theorem 3.1, Proposition 4.1 already provides the upper bound on the estimation error. In order to show the exact equality in (3.3), we use arguments which are similar to Proposition 5.2 in nature. This time, instead of arguing the function  $f(\cdot)$  can be well approximated by the directional derivatives in a small neighborhood, we consider the first order approximation of  $F_C(\mathbf{x}_0)$  by the tangent cone  $T_C(\mathbf{x}_0)$ . The reader is referred to Appendix E for the proof.

We will now discuss the relation between the terms  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  and  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$ ; which relates constrained and regularized estimation results and also the noiseless linear inverse problem (3.8) to each other. From Propositions 5.2 and 4.1, recall that  $\mathbf{D}(\lambda \partial f(\mathbf{x}_0))$  corresponds to the worst case NMSE of the  $\lambda$ -regularized proximity operator (1.3). When we tune  $\lambda$  optimally, we can achieve an error as small as  $\inf_{\lambda \geq 0} \mathbf{D}(\lambda \partial f(\mathbf{x}_0))$ . We related this to  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  via (3.5). The right hand side is due to Theorem 4.5 of [25] which gives the following bound,

$$\inf_{\lambda \geq 0} \mathbf{D}(\lambda \partial f(\mathbf{x}_0)) \leq \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0))) + \frac{2 \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2}{f\left(\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}\right)} \quad (6.1)$$

when  $f(\cdot)$  is a norm. This bound can be enhanced by choosing an  $\mathbf{x}$  that maximizes  $f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|_2}\right)$  while ensuring  $\partial f(\mathbf{x}) = \partial f(\mathbf{x}_0)$ . We will now provide some observations on (6.1). Later on, we will exemplify how this upper bound is quite simple and powerful for well-known structure inducing functions. We will also propose an alternative upper bound of our own that only depends on the subdifferential  $\partial f(\mathbf{x}_0)$  and does not require  $f(\cdot)$  to be norm.

### 6.1 Observations on the upper bound

Upper bounding  $\inf_{\tau \geq 0} \mathbf{D}(\mathbf{x}_0, \tau f)$  in terms of  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  requires technical argument involving Gaussian concentration. While (6.1) looks simple, magnitudes of  $\sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2$  and  $f\left(\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}\right)$  might not always be clear.

For the consequent discussion, we will use the following notation. Let,

$$R_{\mathbf{x}_0} = \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \|\mathbf{s}\|_2 \quad \text{and} \quad f_{\max}(\mathbf{x}_0) = \sup_{\partial f(\mathbf{x}) = \partial f(\mathbf{x}_0)} f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|_2}\right). \quad (6.2)$$

$f_{\max}(\mathbf{x}_0)$  corresponds to the largest value of  $f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|_2}\right)$  while keeping subdifferential fixed. Since  $\inf_{\tau \geq 0} \mathbf{D}(\mathbf{x}_0, \tau f)$  and  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  depends only on the set subdifferential we can use  $f_{\max}(\mathbf{x}_0)$  to get a better bound on the right hand side of (6.1).

We will now briefly investigate this quantity for the classical sparsity inducing functions and argue that it has a simple interpretation in general.

- $\ell_1$  norm: If  $\mathbf{x}_0$  is a  $k$  sparse signal, then, subgradient is given as:

$$\partial \|\mathbf{x}_0\|_1 = \left\{ \mathbf{s} \in \mathbb{R}^n \mid \begin{cases} s_i = \text{sgn}(x_{0,i}) & \text{when } x_{0,i} \neq 0 \\ s_i \in [-1, 1] & \text{else} \end{cases} \right\} \quad (6.3)$$

Hence, subgradients only depend on the sign of the signal. This way, by keeping  $\text{sgn}(\mathbf{x}_0)$  same and normalizing the magnitudes, we can make  $f(\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2})$  as large as  $\sqrt{k}$ . On the other hand,  $R_{\mathbf{x}_0}$  is trivially equal to  $\sqrt{n}$  and it is achieved by choosing  $|s_i| = 1$  for all  $1 \leq i \leq n$ . Consequently, we find:

$$\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)} = \sqrt{\frac{n}{k}} \quad (6.4)$$

- Nuclear norm: We now assume  $\mathbf{x}_0$  is a vectorized form of a rank  $r$  matrix  $\mathbf{X}_0 \in \mathbb{R}^{d \times d}$  where  $n = d^2$ . Assume,  $\mathbf{X}_0$  has singular value decomposition  $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  where  $\mathbf{\Sigma} \in \mathbb{R}^{r \times r}$  has positive diagonals. Set of subgradients of the nuclear norm at  $\mathbf{X}_0$  can be given as [23],

$$\partial\|\mathbf{X}_0\|_* = \{\mathbf{S}|\mathbf{U}^T\mathbf{S}\mathbf{V} = \mathbf{I}_r \quad \text{and} \quad \|(\mathbf{I}_d - \mathbf{U}\mathbf{U}^T)\mathbf{S}(\mathbf{I}_d - \mathbf{V}\mathbf{V}^T)\|_* \leq 1\} \quad (6.5)$$

In a very similar manner to the  $\ell_1$  norm, subgradients of the nuclear norm depends only on the singular vectors of  $\mathbf{X}_0$  and not on its singular values. Consequently, by normalizing singular values of  $\mathbf{X}_0$  without modifying  $\mathbf{S}$  we can make  $f(\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2})$  as high as  $\sqrt{r}$ . On the other hand,  $\sup_{\mathbf{S} \in \partial\|\mathbf{x}_0\|_*} \|\mathbf{S}\|_F$  is  $\sqrt{d}$  and it is achieved when all singular values of  $\mathbf{S}$  are equal to 1. This yields,

$$\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)} = \sqrt{\frac{d}{r}} \quad (6.6)$$

- $\ell_{1,2}$  norm: Let  $\mathbf{x}_0 \in \mathbb{R}^n$  for  $n = tb$  as described in (1.8). Subdifferential of  $\ell_{1,2}$  yields  $R_{\mathbf{x}_0} = \sqrt{t}$ , [23]. Similar to  $\ell_1$  norm, normalizing  $\ell_2$  norms of the individual blocks of  $\mathbf{x}_0$ , it can be shown that  $f_{\max}(\mathbf{x}_0) = \sqrt{k}$ . Hence,

$$\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)} = \sqrt{\frac{t}{k}} \quad (6.7)$$

The reason behind these examples is to provide the reader with an intuition about the quantity  $\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)}$ . From these examples, we observe the following relation,

$$\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)} \approx \left( \frac{\text{Ambient dimension}}{\text{Degrees of freedom}} \right)^{1/2} \quad (6.8)$$

For  $\ell_1$  minimization, this is fairly clear as a sparse signal has  $k$  degrees of freedom. For a  $d \times d$  rank  $r$  matrix, degrees of freedom is  $r(2d - r)$ , [47], which lies between  $dr$  and  $2dr$  and the ambient dimension that the matrix belongs to is  $d^2$ . From (6.6), we indeed have:  $\sqrt{\frac{d}{r}} = \sqrt{\frac{d^2}{dr}}$ .

### 6.1.1 Simple bounds on optimal tuning for common functions

From (6.4) and (6.6), we can state the following results as straightforward observations.

**Proposition 6.1** (Sandwiching optimally tuned NMSE). *Let  $\mathbf{x}_0 \in \mathbb{R}^n$ .*

- Assume,  $\mathbf{x}_0$  is a  $k$  sparse signal (recall (1.6)). Then, using  $\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_1)) \geq k$ , we have:

$$1 \leq \frac{\min_{\lambda \geq 0} \mathbf{D}(\lambda \partial\|\mathbf{x}_0\|_1)}{\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_1))} \leq 1 + 2\sqrt{\frac{n}{k^3}} \quad (6.9)$$

- Assume  $\mathbf{x}_0$  corresponds to a rank  $r$  matrix with size  $d \times d$  where  $n = d^2$  (recall (1.7)). Using  $\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_*)) \geq rd$ ,

$$1 \leq \frac{\min_{\lambda \geq 0} \mathbf{D}(\lambda \partial\|\mathbf{x}_0\|_*)}{\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_*))} \leq 1 + 2\sqrt{\frac{1}{r^3d}} \quad (6.10)$$

- Assume  $\mathbf{x}_0$  is a block sparse signals (recall (1.8)) where  $n = tb$  has  $t$  blocks of size  $b$  and  $k$  of the blocks are nonzero. Using  $\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_{1,2})) \geq kb$ ,

$$1 \leq \frac{\min_{\lambda \geq 0} \mathbf{D}(\lambda \partial\|\mathbf{x}_0\|_{1,2})}{\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_{1,2}))} \leq 1 + 2\sqrt{\frac{t}{b^2k^3}} \quad (6.11)$$

*Proof.* For the proof, we combine (6.4), (6.6), (6.7) with the facts that,

$$\mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_1)) \geq k, \quad \mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_*)) \geq rd, \quad \mathbf{D}(\text{cone}(\partial\|\mathbf{x}_0\|_{1,2})) \geq bk. \quad (6.12)$$

Then, we apply (6.1). These bounds trivially arise from the specific structure of the sub-differentials of the  $\ell_1$ , nuclear norm and  $\ell_{1,2}$  norm. For (6.12), the reader is referred to [23, 24, 68] for a detailed discussion. Note that  $k$ ,  $2dr - r^2$  and  $bk$  are the “degrees of freedom” for the  $k$  sparse signal, rank  $r$  matrix and  $k$  block-sparse signal hence (6.12) is intuitive.  $\square$

### 6.1.2 Alternative upper bounds

We should emphasize that, different bounds can be developed. We will next consider a bound that does not involve the term  $f(\mathbf{x})$ ; however requires the following assumption.

**Assumption 6.1.** *There exists a nontrivial subspace  $T = T(f, \mathbf{x}_0)$  and a vector  $\mathbf{e} = \mathbf{e}(f, \mathbf{x}_0)$  so that, for all  $\mathbf{s} \in \partial f(\mathbf{x}_0)$ , we have:*

$$\text{Proj}(\mathbf{s}, T) = \mathbf{e} \quad (6.13)$$

This assumption is related to the concept of decomposable norms, [23, 24, 56, 68], and is known to be true for sparsity inducing functions. If there are many such subspaces, we will choose the one with the largest dimension. In general,  $T$  is called the *support* of the sparse signal and  $\mathbf{e}$  is known as the “sign” vector. For example, for  $\ell_1$  norm,  $T$  is the set of vectors, whose location of the nonzero entries are same as that of  $\mathbf{x}_0$ . On the other hand,  $\mathbf{e}$  is simply the vector  $\text{sgn}(\mathbf{x}_0)$ .

For the nuclear norm described in (6.5),  $T$  is the set of matrices  $\mathbf{X}$  satisfying  $(\mathbf{I}_d - \mathbf{U}\mathbf{U}^T)\mathbf{S}(\mathbf{I}_d - \mathbf{V}\mathbf{V}^T) = 0$  and the sign vector is  $\mathbf{U}\mathbf{V}^T$ .

To be able to analyze  $\min_{\tau \geq 0} \mathbf{D}(\mathbf{x}_0, \tau f)$ , for a given vector  $\mathbf{v}$ , we are interested in the optimal regularizer  $\tau(\mathbf{v})$  that minimizes  $\text{dist}(\mathbf{v}, \tau \partial f(\mathbf{x}_0))$  over all  $\tau \geq 0$ . In general,  $\tau(\mathbf{v})$  may not be unique. In this case, we will use the smallest value. Assumption 6.1 implies an interesting property for  $\tau(\mathbf{v})$ , namely,  $\tau(\mathbf{v})$  is a  $\|\mathbf{e}\|_2^{-1}$ -Lipschitz function of  $\mathbf{v}$ . In other words, the optimal regularizer cannot change too much with a small change in  $\mathbf{v}$ . More rigorously, for all  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ ,

$$\|\tau(\mathbf{v}_1) - \tau(\mathbf{v}_2)\|_2 \leq \frac{\|\mathbf{v}_1 - \mathbf{v}_2\|_2}{\|\mathbf{e}\|_2} \quad (6.14)$$

Based on this observation, we have the following result. The proof can be found in Lemma C.4 of Appendix C.

**Proposition 6.2.** *Suppose Assumption 6.1 holds. Let  $L_{\mathbf{x}_0} = \|\mathbf{e}\|_2^{-1}$ . Then,*

$$\inf_{\tau \geq 0} \mathbf{D}(\mathbf{x}_0, \tau f) \leq \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0))) + 2\pi(R_{\mathbf{x}_0}^2 L_{\mathbf{x}_0}^2 + R_{\mathbf{x}_0} L_{\mathbf{x}_0} \sqrt{\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))} + 1) \quad (6.15)$$

The right hand side of (6.15) involves the quantity  $R_{\mathbf{x}_0} L_{\mathbf{x}_0}$ . For the norms we considered, this quantity is actually equal to  $\frac{R_{\mathbf{x}_0}}{f_{\max}(\mathbf{x}_0)}$  as we have  $\|\mathbf{e}\|_2 = f_{\max}(\mathbf{x}_0)$ . Consequently, there is a close relation between the two upper bounds (6.1) and (6.15). Observe that, unlike (6.1), (6.15) only depends on the properties of the subdifferential set  $\partial f(\mathbf{x}_0)$  rather than the particular value of  $f(\mathbf{x}_0)$ .

Finally, we should emphasize that  $\|\mathbf{e}\|_2^2$  is closely related to the “generalized sparsity” of the signal. The reader is referred to [23, 24, 56, 68] for a comprehensive discussion.

## 7 Calculation of the LASSO Cost

This section will make use of the results of [25]. Let us introduce the concept of statistical dimension and its relation to Gaussian squared-distance.

**Definition 7.1** (Statistical dimension, [25]). *Let  $\mathcal{C} \in \mathbb{R}^n$  be a closed and convex cone. Statistical dimension of  $\mathcal{C}$  is denoted by  $\delta(\mathcal{C})$  and is defined as,*

$$\delta(\mathcal{C}) = \mathbb{E}[\|\text{Proj}(\mathbf{g}, \mathcal{C})\|_2^2],$$

where  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_n)$ .

When  $\mathcal{C}$  is a linear subspace, the statistical dimension reduces to the regular notion of dimension. In general,  $\delta(\mathcal{C})$  is a good measure of the *size* of  $\mathcal{C}$  and  $\mathcal{C}$  behaves like a  $\delta(\mathcal{C})$  dimensional linear subspace, [25]. It is also closely related to the Gaussian width that has been the topics of papers [8, 20]. When  $\mathcal{C}$  is a convex and closed cone,  $\delta(\mathcal{C})$  and  $\mathbf{D}(\mathcal{C})$  are related as follows,

$$\delta(\mathcal{C}^*) = \mathbf{D}(\mathcal{C})$$

This follows from Moreau's decomposition (Fact A.3) as follows,

$$\mathbf{D}(\mathcal{C}) = \mathbb{E}[\text{dist}(\mathbf{g}, \mathcal{C})^2] = \mathbb{E}[\|\Pi(\mathbf{g}, \mathcal{C})\|_2^2] = \mathbb{E}[\|\text{Proj}(\mathbf{g}, \mathcal{C}^*)\|_2^2] = \delta(\mathcal{C}^*)$$

Again using Moreau's decomposition, it can be shown that  $\delta(\mathcal{C}) + \delta(\mathcal{C}^*) = \mathbb{E}[\|\mathbf{g}\|_2^2] = n$ .

## 7.1 Proof of Theorem 3.3

We first present the proof of Theorem 3.3. To do this, we will make use of Theorem I of [25]; which characterizes the probability of nontrivial intersection for two randomly oriented cones. To rotate a cone  $\mathcal{C}$  randomly, we will multiply its elements by a unitary matrix  $\mathbf{U}$ ; which is drawn uniformly at random.

$$\mathbf{UC} := \{\mathbf{U}\mathbf{a} \mid \mathbf{a} \in \mathcal{C}\}$$

**Theorem 7.1** (Kinematic Formula, [25]). *Let  $A, B$  be two arbitrary closed and convex cones, one of which is not a subspace. Let  $\mathbf{U}$  be a unitary matrix chosen uniformly at random with respect to Haar measure. Then, for any  $\epsilon > 0$ , the followings hold:*

$$\begin{aligned} \delta(A) + \delta(B) \leq n - \epsilon\sqrt{n} &\implies \mathbb{P}(A \cap \mathbf{U}B = \emptyset) \geq 1 - 4 \exp\left(-\frac{\epsilon^2}{16}\right) \\ \delta(A) + \delta(B) \geq n + \epsilon\sqrt{n} &\implies \mathbb{P}(A \cap \mathbf{U}B = \emptyset) \leq 4 \exp\left(-\frac{\epsilon^2}{16}\right) \end{aligned}$$

Theorem 7.1 shows that, two cones will intersect with high probability when  $\delta(A) + \delta(B) > n$ . In order to prove Theorem 3.3, we introduce a useful variation of Theorem 7.1, which probabilistically characterizes the statistical dimension of the intersection when two cones do intersect. Basically, we show that, when  $\delta(A) + \delta(B)$ , with high probability,  $\delta(A \cap \mathbf{U}B)$  is around  $\delta(A) + \delta(B) - n$ . The detailed discussion of this can be found in Appendix D. We now proceed with the proof of Theorem 3.3 and as a first step, we will prove the results on the projected LASSO error  $\eta_{LASSO}$ .

## 7.2 Finding $\eta_{LASSO}$

*Proof of Theorem 3.3, Calculation of  $\eta_{LASSO}$ .* We will reduce (3.9) to a problem that is equivalent to the regular constrained denoising (3.2). Consider the set,

$$\mathbf{AC} = \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathcal{C}\}$$

Hence, finding  $\mathbf{A}\mathbf{x}^*$  is equivalent to finding the solution of:

$$\min_{\mathbf{u} \in \mathbf{AC}} \|\mathbf{y} - \mathbf{u}\|_2 \tag{7.1}$$

where  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \sigma\mathbf{v}$ . From Theorem 3.1, we know that, for fixed  $\mathbf{A}$  and Gaussian  $\mathbf{v}$ , the worst case NMSE  $\mathbf{A}(\mathbf{x} - \mathbf{x}_0)$  satisfies:

$$\eta_{LASSO} = \max_{\sigma > 0} \frac{\mathbb{E}[\|\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_2^2]}{\sigma^2} = \mathbf{D}(T_{\mathbf{AC}}(\mathbf{A}\mathbf{x}_0)^*) \tag{7.2}$$

Hence, we simply need to characterize  $\mathbf{D}(T_{\mathbf{AC}}(\mathbf{A}\mathbf{x}_0)^*)$  to conclude. To do this, we will make use of the fact that  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is a uniformly random partial unitary matrix and we will first characterize  $T_{\mathbf{AC}}(\mathbf{A}\mathbf{x}_0)^*$ . The following lemma, provides this characterization.

**Lemma 7.1.** *Assume  $\mathbf{A}\mathbf{A}^T = \mathbf{I}_m$  and let  $\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T) = T_{\mathcal{C}}(\mathbf{x}_0)^* \cap \text{Range}(\mathbf{A}^T)$ . Then,*

$$T_{\mathbf{AC}}(\mathbf{A}\mathbf{x}_0)^* = \mathbf{AC}(\mathbf{x}_0, \mathbf{A}^T) \tag{7.3}$$

*Proof.* Assume  $\mathbf{s} \in \mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T)$  and  $\mathbf{u} \in F_{\mathcal{AC}}(\mathbf{Ax}_0)$ . Then, there exists  $\mathbf{x} \in \mathcal{C}$  such that,  $\mathbf{u} = \mathbf{A}(\mathbf{x} - \mathbf{x}_0)$  and  $\mathbf{s}' \in \mathcal{C}(\mathbf{x}_0, \mathbf{A}^T)$  such that  $\mathbf{As}' = \mathbf{s}$ . Then,

$$\langle \mathbf{s}, \mathbf{u} \rangle = \langle \mathbf{As}', \mathbf{A}(\mathbf{x} - \mathbf{x}_0) \rangle = \langle \mathbf{A}^T \mathbf{As}', \mathbf{A}^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) \rangle = \langle \mathbf{s}', \mathbf{A}^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) \rangle = \langle \mathbf{s}', \mathbf{x} - \mathbf{x}_0 \rangle \leq 0$$

Hence  $\mathbf{s} \in T_{\mathcal{AC}}(\mathbf{Ax}_0)^*$ . Conversely, assume  $\mathbf{s} \notin \mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T)$  and let  $\mathbf{s}' = \mathbf{A}^T \mathbf{s}$ .  $\mathbf{s}' \notin T_{\mathcal{C}}(\mathbf{x}_0)^*$  because  $\mathbf{s}' \in \text{Range}(\mathbf{A}^T)$  however  $\mathbf{As}' \notin \mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T)$ . Consequently, there exists  $\mathbf{x} \in \mathcal{C}$  such that  $\langle \mathbf{x} - \mathbf{x}_0, \mathbf{s}' \rangle > 0$ . It follows,

$$0 < \langle \mathbf{x} - \mathbf{x}_0, \mathbf{s}' \rangle = \langle \mathbf{x} - \mathbf{x}_0, \mathbf{A}^T \mathbf{As}' \rangle = \langle \mathbf{A}(\mathbf{x} - \mathbf{x}_0), \mathbf{As}' \rangle = \langle \mathbf{A}(\mathbf{x} - \mathbf{x}_0), \mathbf{s} \rangle.$$

where  $\mathbf{Ax} - \mathbf{Ax}_0 \in T_{\mathcal{AC}}(\mathbf{Ax}_0)$ . This implies  $\mathbf{s} \notin T_{\mathcal{AC}}(\mathbf{Ax}_0)^*$ . Overall, we find (7.3).  $\square$

Based on Lemma 7.1, we can write,

$$\eta_{LASSO} = \mathbf{D}(T_{\mathcal{AC}}(\mathbf{Ax}_0)^*) = m - \delta(T_{\mathcal{AC}}(\mathbf{Ax}_0)^*) = m - \delta(\mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T)),$$

and calculate  $\delta(\mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T))$  with high probability. Recall from Lemma 7.1 that,  $\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T)$  is the intersection of a cone with the uniformly random subspace  $\text{Range}(\mathbf{A}^T)$ . Hence Theorem 7.1 is applicable. We will now split the problem in two cases.

- **Case 1**  $m < \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ : In this case, using Theorem 7.1 with  $t = \frac{\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - m}{\sqrt{n}}$ , we find, with probability  $1 - c_1 \exp(-c_2 \frac{(m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*))^2}{n})$ ,  $T_{\mathcal{C}}(\mathbf{x}_0)^* \cap \text{Range}(\mathbf{A}^T) = \{0\}$ . Consequently,  $T_{\mathcal{AC}}(\mathbf{Ax}_0)^* = \mathbf{A}\{0\} = \{0\}$  which gives  $\eta_{LASSO} = m$ .
- **Case 2**  $m > \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ : In this case, we use a modification of Theorem 7.1 which characterizes the statistical dimension of the intersection of a random subspace and a cone. This is, in fact, the topic of Appendix D. From Proposition D.1, there exists constants  $c_1, c_2 > 0$  such that, with probability  $1 - c_1 \exp(-c_2 t^2)$ , we have:

$$|\delta(T_{\mathcal{C}}(\mathbf{x}_0)^* \cap \text{Range}(\mathbf{A}^T)) - (m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*))| \leq t\sqrt{n}$$

Equivalently, we have:

$$m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - t\sqrt{n} \leq \delta(\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T)) \leq m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) + t\sqrt{n} \quad (7.4)$$

On the other hand, since  $\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T) \subset \text{Range}(\mathbf{A}^T)$ , multiplication with partial unitary  $\mathbf{A}$  preserves the distances over  $\text{Range}(\mathbf{A}^T)$  and is isometric, hence it will preserve the statistical dimension (see properties of statistical dimension in [25]) and will yield,

$$\delta(\mathcal{AC}(\mathbf{x}_0, \mathbf{A}^T)) = \delta(\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T))$$

Overall, using (7.2) and applying Lemma 7.1, we have  $\eta_{LASSO} = m - \delta(\mathcal{C}(\mathbf{x}_0, \mathbf{A}^T))$ . Now, using (7.4), with the same probability, we find:

$$\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - t\sqrt{n} \leq \eta_{LASSO} \leq \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) + t\sqrt{n}$$

$\square$

### 7.3 Pinpointing the LASSO Cost

The next result directly follows from characterization of  $\eta_{LASSO}$  in Section 7.2 and Proposition E.1 and gives the result on the LASSO cost  $F_{LASSO}$ . Basically, we use the fact that  $F_{LASSO} + \eta_{LASSO} = m$ .

**Proposition 7.1** (LASSO cost). *Assume  $\mathbf{A}, \mathbf{v}, \mathbf{x}_0, f$  is same as in Theorem 3.3. Consider the cost function  $f_{obj}(\mathbf{x}_0, \mathbf{A}, \mathbf{v}) = \min_{\mathbf{x} \in \mathcal{C}} \|\mathbf{y} - \mathbf{Ax}\|_2^2$  in (3.9). Conditioned on  $\mathbf{A}$ , define the asymptotic LASSO cost as:*

$$F_{LASSO}(\mathbf{A}) = \lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[f_{obj}(\mathbf{x}_0, \mathbf{A}, \mathbf{v})]}{\sigma^2}$$

where the expectation is over  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I}_m)$ . Then, there exists constants  $c_1, c_2 > 0$  such that,

- Whenever  $m < \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , with probability  $1 - c_1 \exp(-c_2 \frac{(m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*))^2}{n})$ ,  $F_{LASSO} = 0$ .
- Whenever  $m > \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , with probability  $1 - c_1 \exp(-c_2 t^2)$ ,

$$m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - t\sqrt{n} \leq F_{LASSO} \leq m - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) + t\sqrt{n}.$$

## 7.4 Numerical results

**Remark:** To relate Theorem 3.3 to the level-set constrained problem  $\min_{f(\mathbf{x}) \leq f(\mathbf{x}_0)} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$ , we will make use of the fact that  $T_f(\mathbf{x}_0)^* = \text{cone}(\partial f(\mathbf{x}_0))$  when  $\mathbf{x}_0$  is not a minimizer of  $f(\cdot)$  (see [84]).

### 7.4.1 LASSO with the $\ell_1$ norm

We considered the following  $\ell_1$  constrained optimization,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq \|\mathbf{x}_0\|_1. \quad (7.5)$$

We let  $\mathbf{x}_0$  to be a  $k$  sparse vector and chose  $k$  to be 20, 40 and 60 while varying number of measurements  $m$  from 20 to 400. The ambient dimension is  $n = 500$ . We have performed 50 realization of the problem in which  $\mathbf{A}$  is generated as a random unitary matrix and  $\sigma\mathbf{v}$  is the noise vector, where  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$  and  $\sigma$  is sufficiently small.

We have estimated the quantities  $\eta_{LASSO}$  and  $F_{LASSO}$  by averaging  $\|\mathbf{y} - \mathbf{A}\mathbf{x}^*\|_2^2$  and  $\|\mathbf{A}\mathbf{x}^* - \mathbf{A}\mathbf{x}_0\|_2^2$  over 50 realizations.

Finally, in order to verify our results, we estimate the term  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ . This is done by making use of the classical results on  $\ell_1$  phase transitions, (see [3, 26, 28]). In particular, see Theorem 4 of [26]. For example, when  $n = 500$  and  $\mathbf{x}_0$  is 20 sparse, we find  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1)) \approx 89$ . Similarly, 40 and 60 gives  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1)) \approx 142$  and 186 respectively.

Figure 2 illustrates our results. In Figure 2(a), we observe that we can accurately predict projected LASSO error based on Theorem 3.3. The dashed red line is what we theoretically expect and the blue, green and black markers are the experimental results for  $k = 20, 40$  and 60 respectively.

Figure 2(b) demonstrates that, LASSO cost can be similarly predicted. We should note that, in both figures there is an apparent phase transition. When the number of measurements are not sufficient, ( $m < \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ ), we observe that  $\eta_{LASSO}$  increases linearly in  $m$  actually it is equal to  $m$ . On the other hand, when the number of measurements are sufficient,  $\eta_{LASSO}$  stays same and is simply equal to  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ .

Recall that,  $m > \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$  regime is the regime where the noiseless compressed sensing problem succeeds. For  $\ell_1$  minimization, [33] argues that, there is in fact a phase transition for noise sensitivity, and when  $m > \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ , the LASSO will recover  $\mathbf{x}_0$  robustly and when  $m < \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ , the normalized error  $\frac{\|\mathbf{x}^* - \mathbf{x}_0\|_2^2}{\sigma^2}$  will be unbounded as  $\sigma \rightarrow 0$ . Overall, we observe that the phase transition for  $\eta_{LASSO}$  and  $F_{LASSO}$  occurs exactly at  $m \approx \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1))$ ; which is also known to be the noise sensitivity threshold [33].

It should be emphasized that the more popular formulation of LASSO is,

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1.$$

This has been studied in a series of papers [29–31, 33] and [29, 30, 74] give analytical characterization of the asymptotic LASSO error  $\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]$ . On the other hand, our results hold for arbitrary convex functions and are not limited to  $\ell_1$  minimization. Now, we will illustrate this with an example on nuclear norm.

### 7.4.2 LASSO with the nuclear norm

We next considered a scenario where  $\mathbf{x}_0$  is a vectorized form of a  $30 \times 30$ , rank 4 matrix. We observe  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \sigma\mathbf{v}$  where  $\mathbf{A}$  is an  $m \times 900$  partial unitary and  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ . We solve,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_{\star} \leq \|\mathbf{x}_0\|_{\star} \quad (7.6)$$

where  $\|\cdot\|_{\star}$  returns the nuclear norm of the  $30 \times 30$  matrix form of  $\mathbf{x}_0$ .

To predict the LASSO cost, we made use of the results of [46], which approximates  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_{\star}))$  based on the asymptotic singular value distribution of the  $n \times n$  i.i.d. Gaussian matrices in a similar manner to the  $\ell_1$  phase transition calculations.

Using results of [46], we estimate  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_{\star})) \approx 389$ . We then repeat the same experiment that is described in the previous section where we vary  $m$  from 30 to 900 in the intervals of 30 and average the results to estimate  $\eta_{LASSO}$  and  $F_{LASSO}$ .

This time, we additionally approximated the actual asymptotic LASSO error  $E_{LASSO} = \lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2}$ . Figure 3 shows the projected error and the actual error as a function of  $m$ . We observe that, projected error

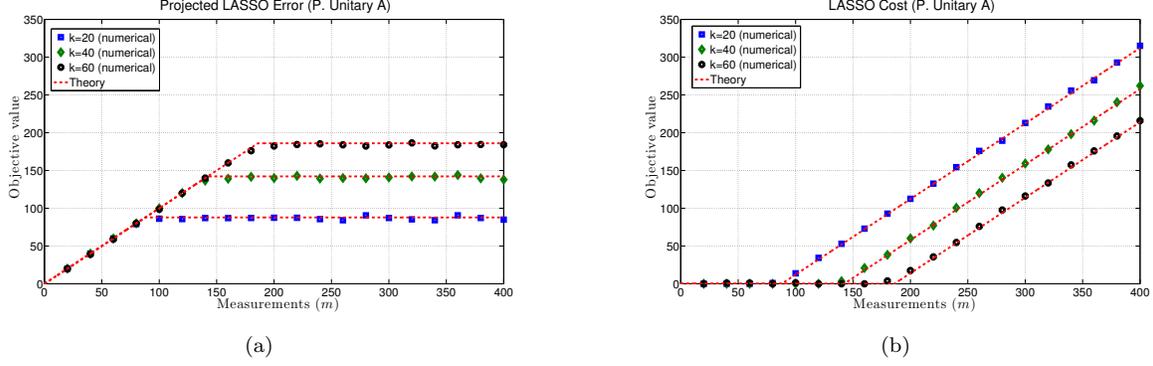


Figure 2: *a)* represents the expected projected LASSO error  $\|\mathbf{A}(\mathbf{x}^* - \mathbf{x}_0)\|_2^2$  when  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is partial unitary. *b)* represents the LASSO cost i.e.  $\mathbb{E}[\|\mathbf{y} - \mathbf{A}\mathbf{x}^*\|_2^2]$ . Results are for  $\ell_1$  minimization, and for three different sparsity levels of  $k = 20, 40$  and  $60$ . Vector length is  $n = 500$ . Measurements run from  $m = 20$  to  $400$ .

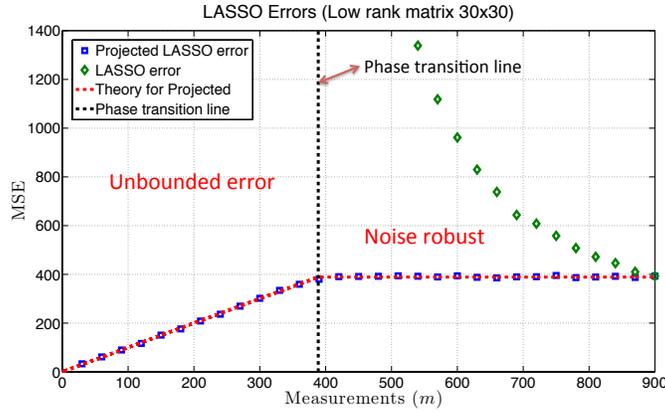


Figure 3: Simulation is performed for a  $30 \times 30$  matrix of rank 4. The blue markers are the projected LASSO error; which matches with theory. Green markers are the (total) normalized LASSO MSE; which is unbounded on the left side of the phase transition line and a decreasing function of  $m$  on the right side of the PT line  $m = \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  (recall Section 3.3).

can again be accurately predicted from our theoretical results. The LASSO is not robust for the regime  $m < \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_*))$  and  $E_{LASSO}$  is unbounded. Conversely, when  $m > \mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_*))$  the error becomes finite and decreases as a function of  $m$ . At  $m = n = 900$ , we see that  $E_{LASSO} = \eta_{LASSO}$ . This is not surprising due to the fact that  $\|\mathbf{A}(\mathbf{x}^* - \mathbf{x}_0)\|_2 = \|\mathbf{x}^* - \mathbf{x}_0\|_2$  when  $m = 900$  as  $\mathbf{A}$  becomes a full (square) unitary matrix and preserves the  $\ell_2$  norm. Figure 4(a) provides the LASSO cost for the same simulation. The cost satisfies  $F_{LASSO} = m - \eta_{LASSO}$  as predicted by our theory. Figure 4(a) also illustrates how  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  corresponds to the stability phase transition of (7.6) (recall Section 3.3).

### 7.4.3 LASSO for Gaussian matrices

We additionally performed the same simulations in Section 7.4.1 and solved (7.5) for i.i.d. Gaussian  $\mathbf{A}$  rather than partial unitary. Estimation of the quantities  $\frac{\mathbb{E}[\|\mathbf{A}(\mathbf{x}^* - \mathbf{x}_0)\|_2^2]}{\sigma^2}$  and  $\frac{\mathbb{E}[\|\mathbf{y} - \mathbf{A}\mathbf{x}^*\|_2^2]}{\sigma^2}$  are carried out in the in the exact same manner. In Figure 4(b) the dashed red line is the theoretical prediction for  $\eta_{LASSO}$  when  $\mathbf{A}$  is unitary and the blue, green and black markers are the results for i.i.d. Gaussian  $\mathbf{A}$ . Hence, Figure 4(b) indicates that, our results for partial unitary compression matrices also holds for Gaussian compression matrices and the results might be universal. This would not be surprising given the recent advances of the universality of the phase transitions, [71, 72].

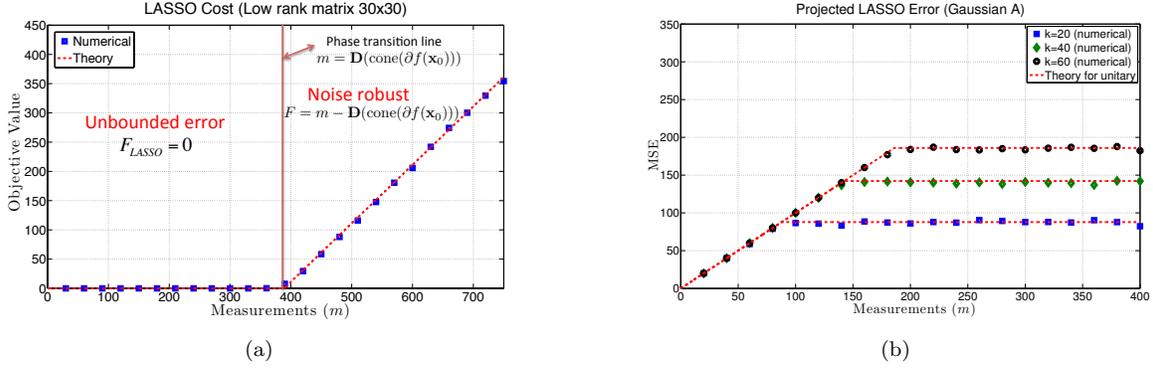


Figure 4: a) The LASSO cost as a function of measurements for  $30 \times 30$ , rank 4 matrix. b) This figure is same as Figure 2(a) except i.i.d. Gaussian measurement is used rather than partial unitary.

## 8 Discussion of the results

We have considered the proximal denoising problem and provided sharp MSE upper bounds that are achievable in the small noise regime. Our bounds depend on the convex geometry of the problem (1.1) and can be captured by distance to the scaled subdifferential  $\lambda \partial f(\mathbf{x}_0)$  or to the subdifferential cone  $\text{cone}(\partial f(\mathbf{x}_0))$ . These are meaningful quantities when  $\mathbf{x}_0$  is a structured signal and  $f(\cdot)$  is properly chosen to induce structure. Surprisingly, our estimation bounds are closely related to the recovery phase transition of the linear inverse problem (3.8).

We also showed an interesting phase transition for the generalized LASSO problem. When the number of measurements  $m$  are smaller than  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ , there is no noise robustness and the cost value is equal to 0. On the other hand, when the number of measurements are more than  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ , the cost is around  $m - \mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ . This indicates that behavior of LASSO is closely related to the same quantity  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ .

We should again emphasize that, while  $\ell_1$ -minimization is often the primary interest, our results apply to all convex functions.

### 8.1 Future directions

- Analysis of Generalized LASSO: While we considered some basic properties of the generalized LASSO problem (3.9), the most critical one is yet to be explored. We believe there is a simple formula that predicts the LASSO error in,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda f(\mathbf{x}) \quad (8.1)$$

which depends on the number of measurements  $m$ , the subdifferential  $\partial f(\mathbf{x}_0)$  and the regularizer  $\lambda$ . The generalized LASSO analysis will be a unification of the results on noiseless compressed sensing (3.8) and compression-less denoising problem (1.1). Results of Chandrasekaran et al. and Amelunxen et al. [8,25] apply to noiseless linear inverse problem (3.8) and we have mostly focused on estimation via convex functions without linear measurements. With generalized LASSO analysis, one will hopefully be able to predict the behavior of the noisy linear inverse problem and extend the results of [29,30,74], from  $\ell_1$ -minimization to arbitrary functions.

- Universality in denoising: Assume the noise vector  $\mathbf{z}$  has independent and identically distributed (i.i.d.) entries but the distribution is not normal. In this case, unfortunately our MSE bounds are no longer accurate. A simple example is the scenario where  $\mathbf{x}_0$  is a sparse signal, and the entries of  $\mathbf{z}$  are equally likely to be  $\sigma$  and  $-\sigma$ . If  $\mathbf{x}_0$  is  $k$  sparse,  $\mathbf{D}(\text{cone}(\|\mathbf{x}_0\|_1)) \sim k \log \frac{2n}{k}$ , [28].

On the other hand, assuming  $\sigma$  is sufficiently small and setting  $\lambda = 1$  in (1.3), soft-thresholding reveals that the estimate  $\mathbf{x}^*$  satisfies,

$$x_i^* = \begin{cases} 0 & \text{if } x_{0,i} = 0 \\ x_{0,i} - \text{sgn}(x_{0,i})\sigma + z_i & \text{else} \end{cases} \quad (8.2)$$

Overall, this gives  $\mathbb{E}[\frac{\|\mathbf{x}^* - \mathbf{x}_0\|_2^2}{\sigma^2}] = 2k$ ; which is smaller than  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$  for the regime  $k \ll n$ .

While this shows that estimation error may depend on the distribution, one can actually introduce further randomization to the noise. Given  $\mathbf{z}$ , assume we observe  $\mathbf{x}_0 + \mathbf{U}\mathbf{z}$  where  $\mathbf{U}$  is a uniformly random unitary matrix. We believe that, if one first generates and fixes  $\mathbf{U}$  and then takes the expectation over  $\mathbf{z}$ , the worst case normalized MSE will in fact be around  $\mathbf{D}(\text{cone}(\partial f(\mathbf{x}_0)))$ , under mild assumptions. Possible such assumptions are i.i.d.'ness and subgaussianity of the entries of  $\mathbf{z}$ ; which are used in [71] to show the universality of the compressed sensing phase transitions for  $\ell_1$  minimization.

## Acknowledgments

Authors would like to thank Michael McCoy and Joel Tropp for stimulating discussions and helpful comments. Michael McCoy pointed out Lemma C.1 and informed us of various recent results most importantly Theorem 7.1. S.O. would also like to thank his colleagues Kishore Jaganathan and Christos Thrampoulidis for their support.

## References

- [1] E. J. Candès, J. Romberg, and T. Tao. “Stable signal recovery from incomplete and inaccurate measurements”. *Comm. Pure Appl. Math.*, 59:1207–1223, 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao. “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information”. *IEEE Trans. Inform. Theory*, 52 489–509.
- [3] D. L. Donoho, “Compressed Sensing,” *IEEE Trans. on Information Theory*, 52(4), pp. 1289 - 1306, April 2006.
- [4] D. L. Donoho, “De-noising by soft-thresholding,” *Information Theory, IEEE Transactions on*, vol. 41, no. 3, pp. 613-627, 1995.
- [5] J. Cai, E. J. Candès, and Z. Shen. “A Singular Value Thresholding Algorithm for Matrix Completion.” *SIAM Journal on Optimization*, 20 1956–1982, (2008).
- [6] E. J. Candès and B. Recht. “Exact matrix completion via convex optimization”. *Found. of Comput. Math.*, 9 717–772.
- [7] M. Stojnic, F. Parvaresh, and B. Hassibi, “On the reconstruction of block-sparse signals with an optimal number of measurements”. *IEEE Trans. on Signal Processing*, vol.57, no.8, pp.3075-3085, Aug. 2009
- [8] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, “The Convex Geometry of Linear Inverse Problems”, *Foundations of Computational Mathematics*. Online First, October 2012.
- [9] F. Bach “Structured sparsity-inducing norms through submodular functions”, NIPS 2010.
- [10] J. J. Moreau. “Fonctions convexes duales et points proximaux dans un espace hilbertien.” *C.R. Acad. Sci. Paris Sér. A Math.*, 255:1897–1899, 1962.
- [11] Jenatton, Rodolphe, et al. “Proximal methods for sparse hierarchical dictionary learning.” *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 2010.
- [12] Güler, Osman. “On the convergence of the proximal point algorithm for convex minimization.” *SIAM Journal on Control and Optimization* 29.2 (1991): 403-419.
- [13] Rockafellar, R. Tyrrell. “Monotone operators and the proximal point algorithm.” *SIAM Journal on Control and Optimization* 14.5 (1976): 877-898.
- [14] T. Hale, W. Yin, and Y. Zhang, “A fixed-point continuation method for  $\ell_1$  regularized minimization: Methodology and convergence.” (2008). *SIAM J. Optim.*,19(3), 1107–1130.
- [15] P. Combettes, and V. Wajs, “Signal recovery by proximal forward-backward splitting.” *Multiscale Model. Simul.* Vol. 4, No. 4, pp. 1168–1200
- [16] P.L. Combettes and J.C. Pesquet, “Proximal splitting methods in signal processing”. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212, 2011.
- [17] N. Parikh and S. Boyd, “Proximal algorithms”. *Foundations and Trends in Optimization*, 1-96, (2013).
- [18] M. Figueiredo and R. Nowak, “An EM Algorithm for Wavelet-Based Image Restoration.” *IEEE Transactions on Image Processing* (2003),12 906–916.

- [19] B. N. Bhaskar, G. Tang, and B. Recht, “Atomic norm denoising with applications to line spectral estimation”, arXiv:1204.0562.
- [20] V. Chandrasekaran, and M. I. Jordan, “Computational and statistical tradeoffs via convex relaxation”, Proceedings of the National Academy of Sciences 110.13 (2013): E1181-E1190.
- [21] Becker, Stephen, and M. Jalal Fadili. ”A quasi-Newton proximal splitting method.” arXiv preprint arXiv:1206.1156 (2012).
- [22] Jason D. Lee, Yuekai Sun, and Michael A. Saunders, “Proximal Newton-type methods for minimizing convex objective functions in composite form”, arXiv:1206.1623.
- [23] E. J. Candès and B. Recht. “Simple bounds for low-complexity model reconstruction”. To appear in Mathematical Programming Series A.
- [24] S. N. Negahban, P. Ravikumar, M. J. Wainwright, B. Yu, “A Unified Framework for High-Dimensional Analysis of M-Estimators with Decomposable Regularizers”, Statist. Sci. Volume 27, Number 4 (2012), 538–557.
- [25] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, “Living on the edge: A geometric theory of phase transitions in convex optimization”. arXiv:1303.6672.
- [26] M. Stojnic, “Various thresholds for  $\ell_1$  - optimization in compressed sensing”, arXiv:0907.3666v1.
- [27] D. L. Donoho and J. Tanner, “Neighborliness of randomly-projected simplices in high dimensions,” Proc. National Academy of Sciences, 102(27), pp. 9452-9457, 2005.
- [28] D. L. Donoho, “High-dimensional centrally-symmetric polytopes with neighborliness proportional to dimension”, Comput. Geometry, (online) Dec. 2005.
- [29] M. Bayati and A. Montanari. “The dynamics of message passing on dense graphs, with applications to compressed sensing”, IEEE Transactions on Information Theory, Vol. 57, No. 2, 2011.
- [30] M. Bayati and A. Montanari, “The LASSO risk for gaussian matrices”, IEEE Transactions on Information Theory, Vol. 58, No. 4, 2012.
- [31] D. L. Donoho, A. Maleki, and A. Montanari, “Message Passing Algorithms for Compressed Sensing”, PNAS November 10, 2009 vol. 106 no. 45 18914–18919.
- [32] A. Maleki, “Analysis of approximate message passing algorithm”, CISS 2010.
- [33] D. L. Donoho, A. Maleki, A. Montanari, “The Noise-Sensitivity Phase Transition in Compressed Sensing”, IEEE Trans. Inform. Theory, 57 6920–6941 .
- [34] S. Oymak, C. Thrampoulidis, and B. Hassibi, “The Squared-Error of Generalized LASSO: A Precise Analysis”, arXiv:1311.0830.
- [35] D. L. Donoho, M. Gavish, A. Montanari, “The Phase Transition of Matrix Recovery from Gaussian Measurements Matches the Minimax MSE of Matrix Denoising”, arXiv:1302.2331.
- [36] D. L. Donoho, M. Gavish, “Minimax Risk of Matrix Denoising by Singular Value Thresholding”, arXiv:1304.2085.
- [37] A. A. Shabalin, and A. B. Nobel. “Reconstruction of a low-rank matrix in the presence of Gaussian noise.” Journal of Multivariate Analysis (2013).
- [38] D. L. Donoho and I. M. Johnstone. “Minimax risk over  $\ell_p$ -balls for  $\ell_q$ -error”. Probability Theory and Related Fields, 303:277–303, 1994
- [39] S. Oymak and B. Hassibi, “On a relation between the minimax risk and the phase transitions of compressed recovery,” Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on , vol., no., pp.1018,1025, 1–5 Oct. 2012.
- [40] D. L. Donoho, I. Johnstone, and A. Montanari, “Accurate Prediction of Phase Transitions in Compressed Sensing via a Connection to Minimax Denoising”, IEEE Trans. on Information Theory, vol.59, no.6, pp.3396,3433, June 2013.
- [41] D. L. Donoho, J. Tanner, “Thresholds for the recovery of sparse solutions via  $\ell_1$  minimization”, Conf. on Information Sciences and Systems, 2006.
- [42] M. A. Khajehnejad and B. Hassibi, “Sparse Recovery of Positive Signals with Minimal Expansion”, In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP ’09).
- [43] V. Koltchinskii, K. Lounici, and A. Tsybakov, “Nuclear norm penalization and optimal rates for noisy matrix completion”, Annals of Statistics, 2011.
- [44] E. J. Candès and Y. Plan. “Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements.” IEEE Transactions on Information Theory 57(4), 2342-2359.
- [45] M. Fazel, “Matrix Rank Minimization with Applications”. Elec. Eng. Dept, Stanford University, March 2002.

- [46] S. Oymak and B. Hassibi, “New Null Space Results and Recovery Thresholds for Matrix Rank Minimization”, arXiv:1011.6326.
- [47] Recht, Benjamin, Weiyu Xu, and Babak Hassibi. “Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization.” *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on.* IEEE, 2008.
- [48] B. Recht, M. Fazel, P. Parrilo, “Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization”. *SIAM Review*, Vol 52, no 3, pages 471–501, 2010.
- [49] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model based compressive sensing,” *IEEE Trans. on Information Theory*.
- [50] Y. C. Eldar, P. Kuppinger, and H. Blcskei, “Block-Sparse Signals: Uncertainty Relations and Efficient Recovery”, *IEEE Trans. on Signal Proc.*, Vol. 58, No. 6, June 2010.
- [51] N. Rao, B. Recht, and R. Nowak, “Tight Measurement Bounds for Exact Recovery of Structured Sparse Signals”. In *Proceedings of AISTATS*, 2012.
- [52] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, “Weighted  $\ell_1$  minimization for sparse recovery with prior information”, *ISIT* 2009.
- [53] E. J. Candès, M. Wakin and S. Boyd. “Enhancing sparsity by reweighted  $\ell_1$  minimization”. *J. Fourier Anal. Appl.*, 14 877–905.
- [54] S. Oymak, M. A. Khajehnejad, and B. Hassibi, “Recovery threshold for optimal weight  $\ell_1$  minimization”, *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, vol., no., pp.2032,2036, 1–6 July 2012.
- [55] N. Vaswani and W. Lu, Modified-CS: Modifying Compressive Sensing for Problems with Partially Known Support, *IEEE Trans. Signal Processing*, September 2010.
- [56] J. Wright, A. Ganesh, K. Min, and Y. Ma, “Compressive Principal Component Pursuit”, arXiv:1202.4596.
- [57] Z. Zhou, X. Li, J. Wright, E. Candès, and Y. Ma. “Stable principal component pursuit.” In *Proc. of International Symposium on Information Theory*, 2010.
- [58] E. J. Candès, X. Li, Y. Ma, and J. Wright. “Robust Principal Component Analysis?” *Journal of ACM* 58(1), 1–37.
- [59] A. Agarwal, S. N. Negahban, M. J. Wainwright, “Noisy matrix decomposition via convex relaxation: Optimal rates in high dimensions”. *Annals of Statistics* 2012, Vol. 40, No. 2, 1171–1197.
- [60] M. Tao and X. Yuan, “Recovering low-rank and sparse components of matrices from incomplete and noisy observations”, *SIAM J. Optim.*, 21(1), 57–81.
- [61] R. Foygel and L. Mackey, “Corrupted Sensing: Novel Guarantees for Separating Structured Signals”, arXiv:1305.2524.
- [62] S. Ma, D. Goldfarb, and L. Chen, “Fixed point and Bregman iterative methods for matrix rank minimization.” *Mathematical Programming*, 128 321–353 (2011).
- [63] K. Toh and S. Yun, (2010). “An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems.” *Pacific Journal on Optimization*, 6 615–640
- [64] L. I. Rudin, S Osher, E Fatemi, “Nonlinear total variation based noise removal algorithms”, *Physica D: Nonlinear Phenomena*, Volume 60, Issues 1–4, 1 November 1992, Pages 259–268.
- [65] S. L. Keeling, “Total variation based convex filters for medical imaging”, *Appl. Math. Comput.*, 139 (2003), pp. 101–119.
- [66] D. Needell and R. Ward, “Stable image reconstruction using total variation minimization”, arXiv preprint arXiv:1202.6429, (2012).
- [67] J-F Cai and W. Xu, “Guarantees of Total Variation Minimization for Signal Recovery”, arXiv:1301.6791.
- [68] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi, “Simultaneously Structured Models with Application to Sparse and Low-rank Matrices”, arXiv:1212.3753.
- [69] E. Richard, P. Savalle, and N. Vayatis, “Estimation of Simultaneously Sparse and Low Rank Matrices”, in *Proceedings of the 29th International Conference on Machine Learning (ICML 2012)*.
- [70] E. Richard, F. Bach, and J-P Vert. “Intersecting singularities for multi-structured estimation.”
- [71] M. Bayati, M. Lelarge, and A. Montanari. “Universality in polytope phase transitions and message passing algorithms”. Available at [arxiv.org/abs/1207.7321](http://arxiv.org/abs/1207.7321), 2012.
- [72] D. L. Donoho and J. Tanner. “Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing.” *Phil. Trans. R. Soc. A* 13 November 2009 vol. 367 no. 1906 4273–4293.
- [73] R. Tibshirani, “Regression shrinkage and selection via the lasso.” *Journal of the Royal Statistical Society*, 58:267–288, 1996.

- [74] P. J. Bickel, Y. Ritov and A. Tsybakov “Simultaneous analysis of LASSO and Dantzig Selector”. *The Annals of Statistics*, 37(4):1705–1732, 2009.
- [75] F. Bunea, A. B. Tsybakov, and M. H. Wegkamp. Sparsity oracle inequalities for the lasso. *Electronic Journal of Statistics*, 1:169–194, 2007.
- [76] M. Stojnic, “A framework to characterize performance of LASSO algorithms”, arXiv:1303.7291.
- [77] M. Stojnic, “A performance analysis framework for SOCP algorithms in noisy compressed sensing”, arXiv:1304.0002.
- [78] S.S. Chen and D. Donoho. “Examples of basis pursuit”. *Proceeding of wavelet applications in signal and image processing III*, 1995.
- [79] CVX Research, Inc. “CVX: Matlab software for disciplined convex programming”, version 2.0 beta. <http://cvxr.com/cvx>, September 2012.
- [80] J.F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for Optimization over Symmetric Cones”, *Optimization Methods and Software* 11(12), pp. 625–653, 1999.
- [81] M. Ledoux, M. Talagrand, “Probability in Banach Spaces: Isoperimetry and Processes”. Springer, 1991.
- [82] M. Ledoux. “The concentration of measure phenomenon”, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI 2001.
- [83] V. I. Bogachev. “Gaussian measures”, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998
- [84] R. T. Rockafellar, “Convex Analysis”, Vol. 28. Princeton university press, 1997.
- [85] Y. Nesterov, “Introductory Lectures on Convex Optimization. A Basic Course”, 2004.
- [86] D. Bertsekas with A. Nedic and A.E. Ozdaglar, “Convex Analysis and Optimization”. Athena Scientific, 2003.
- [87] S. Boyd and L. Vandenberghe, “Convex Optimization” Cambridge University Press, 2004.
- [88] JB Hiriart-Urruty and C Lemarchal. “Convex Analysis and Minimization Algorithms: Part 1: Fundamentals”. Vol. 1. Springer, 1996.

# APPENDIX

## A Auxiliary results

**Fact A.1** (Hyperplane separation theorem, [86]). *Assume  $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathbb{R}^n$  are disjoint closed sets at least one of which is compact. Then, there exists a hyperplane  $H$  such that  $\mathcal{C}_1, \mathcal{C}_2$  lies on different open half planes induced by  $H$ .*

**Fact A.2** (Properties of the projection, [86, 87]). *Assume  $\mathcal{C} \subseteq \mathbb{R}^n$  is a nonempty, closed and convex set and  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$  are arbitrary points. Then:*

$$\|\text{Proj}(\mathbf{a}) - \text{Proj}(\mathbf{b})\|_2 \leq \|\mathbf{a} - \mathbf{b}\|_2$$

The projection  $\text{Proj}(\mathbf{a}, \mathcal{C})$  is the unique vector satisfying,

$$\text{Proj}(\mathbf{a}, \mathcal{C}) = \arg \min_{\mathbf{v} \in \mathcal{C}} \|\mathbf{a} - \mathbf{v}\|_2 \tag{A.1}$$

The projection  $\text{Proj}(\mathbf{a}, \mathcal{C})$  is also the unique vector  $\mathbf{s}_0$  that satisfies,

$$\langle \mathbf{s}_0, \mathbf{a} - \mathbf{s}_0 \rangle = \sup_{\mathbf{s} \in \mathcal{C}} \langle \mathbf{s}, \mathbf{a} - \mathbf{s}_0 \rangle \tag{A.2}$$

In other words,  $\mathbf{a}$  and  $\mathcal{C}$  lies on different half planes induced by the hyperplane that goes through  $\text{Proj}(\mathbf{a}, \mathcal{C})$  and that is orthogonal to  $\mathbf{a} - \text{Proj}(\mathbf{a}, \mathcal{C})$ .

**Fact A.3** (Moreau's decomposition theorem, [10]). *Let  $\mathcal{C}$  be a closed and convex cone in  $\mathbb{R}^n$ . For any  $\mathbf{v} \in \mathbb{R}^n$ , the followings are equivalent:*

- $\mathbf{v} = \mathbf{a} + \mathbf{b}$ ,  $\mathbf{a} \in \mathcal{C}$ ,  $\mathbf{b} \in \mathcal{C}^*$  and  $\mathbf{a}^T \mathbf{b} = 0$ .
- $\mathbf{a} = \text{Proj}(\mathbf{v}, \mathcal{C})$ ,  $\mathbf{b} = \text{Proj}(\mathbf{v}, \mathcal{C}^*)$ .

**Definition A.1** (Lipschitz function).  $h(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  is called  $L$ -Lipschitz if for all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $|h(\mathbf{x}) - h(\mathbf{y})| \leq L \|\mathbf{x} - \mathbf{y}\|_2$ .

The next lemma provides a concentration inequality for Lipschitz functions of Gaussian vectors, [81].

**Fact A.4.** *Let  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I})$  and  $h(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  be an  $L$ -Lipschitz function. Then for all  $t \geq 0$ :*

$$\mathbb{P}(|h(\mathbf{g}) - \mathbb{E}[h(\mathbf{g})]| \geq t) \leq 2 \exp\left(-\frac{t^2}{2L^2}\right)$$

**Lemma A.1.** *For any  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I})$ ,  $c > 1$ , we have:*

$$\mathbb{P}(\|\mathbf{g}\|_2 \geq c\sqrt{n}) \leq 2 \exp\left(-\frac{(c-1)^2 n}{2}\right)$$

*Proof.*  $\mathbb{E}[\|\mathbf{g}\|_2] \leq \sqrt{\mathbb{E}[\|\mathbf{g}\|_2^2]} = \sqrt{n}$ . Secondly  $\ell_2$  norm is a 1-Lipschitz function due to the triangle inequality. Hence:

$$\mathbb{P}(\|\mathbf{g}\|_2 \geq c\sqrt{n}) \leq \mathbb{P}(\|\mathbf{g}\|_2 \geq (c-1)\sqrt{n} + \mathbb{E}[\|\mathbf{g}\|_2]) \leq 2 \exp\left(-\frac{(c-1)^2 n}{2}\right)$$

□

**Lemma A.2.** *Let  $\mathcal{C}$  be a closed and convex cone in  $\mathbb{R}^n$ . Then,  $\mathbf{D}(\mathcal{C}) + \mathbf{D}(\mathcal{C}^*) = n$ .*

*Proof.* Using Fact A.3, any  $\mathbf{v} \in \mathbb{R}^n$  can be written as  $\text{Proj}(\mathbf{v}, \mathcal{C}) + \text{Proj}(\mathbf{v}, \mathcal{C}^*) = \mathbf{v}$  and  $\langle \text{Proj}(\mathbf{v}, \mathcal{C}), \text{Proj}(\mathbf{v}, \mathcal{C}^*) \rangle = 0$ . Hence,

$$\|\mathbf{v}\|^2 = \|\text{Proj}(\mathbf{v}, \mathcal{C}^*)\|^2 + \|\text{Proj}(\mathbf{v}, \mathcal{C})\|^2 = \text{dist}(\mathbf{v}, \mathcal{C})^2 + \text{dist}(\mathbf{v}, \mathcal{C}^*)^2$$

Letting  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$  and taking the expectations, we can conclude.

□

## B Subdifferential of the approximation

*Proof of Lemma 4.3.* Recall that  $\hat{f}_{\mathbf{x}_0}(\mathbf{x}_0 + \mathbf{v}) - f(\mathbf{x}_0)$  is equal to the directional derivative  $f'(\mathbf{x}_0, \mathbf{v}) = \sup_{\mathbf{s} \in \partial f(\mathbf{x}_0)} \langle \mathbf{s}, \mathbf{v} \rangle$ . Also recall the “set of maximizing subgradients” from (4.13). Clearly,  $\partial f'(\mathbf{x}_0, \mathbf{v}) = \partial \hat{f}_{\mathbf{x}_0}(\mathbf{x}_0 + \mathbf{v})$ . We will let  $\mathbf{x} = \mathbf{w} + \mathbf{x}_0$  and investigate  $\partial f'(\mathbf{x}_0, \mathbf{w})$  as a function of  $\mathbf{w}$ .

**If  $\mathbf{w} = 0$ :** For any  $\mathbf{s} \in \partial f(\mathbf{x}_0)$  and any  $\mathbf{v}$  by definition, we have:

$$f'(\mathbf{x}_0, \mathbf{v}) - f'(\mathbf{x}_0, 0) = f'(\mathbf{x}_0, \mathbf{v}) = \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0)} \langle \mathbf{v}, \mathbf{s}' \rangle \geq \langle \mathbf{v}, \mathbf{s} \rangle$$

hence  $\mathbf{s} \in \partial f'(\mathbf{x}_0, 0)$ . Conversely, assume  $\mathbf{s} \notin \partial f(\mathbf{x}_0)$ , then there exists  $\mathbf{v}$  such that:

$$f(\mathbf{v} + \mathbf{x}_0) < f(\mathbf{x}_0) + \langle \mathbf{v}, \mathbf{s} \rangle$$

By convexity for any  $\epsilon > 0$ :

$$\frac{f(\epsilon \mathbf{v} + \mathbf{x}_0) - f(\mathbf{x}_0)}{\epsilon} \leq f(\mathbf{v} + \mathbf{x}_0) - f(\mathbf{x}_0) < \langle \mathbf{v}, \mathbf{s} \rangle$$

Taking  $\epsilon \rightarrow 0$  on the left hand side we find:

$$f'(\mathbf{x}_0, \mathbf{v}) - f'(\mathbf{x}_0, 0) = f'(\mathbf{x}_0, \mathbf{v}) < \langle \mathbf{v}, \mathbf{s} \rangle$$

which implies  $\mathbf{s} \notin \partial f'(\mathbf{x}_0, 0)$ .

**If  $\mathbf{w} \neq 0$ :** Now, consider the case  $\mathbf{w} \neq 0$ . Assume,  $\mathbf{s} \in \partial f(\mathbf{x}_0, \mathbf{w})$ . Then, for any  $\mathbf{v}$ , we have:

$$f'(\mathbf{x}_0, \mathbf{w} + \mathbf{v}) - f'(\mathbf{x}_0, \mathbf{w}) = \sup_{\mathbf{s}_1 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w} + \mathbf{v}, \mathbf{s}_1 \rangle - \sup_{\mathbf{s}_2 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s}_2 \rangle \quad (\text{B.1})$$

$$= \sup_{\mathbf{s}_1 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w} + \mathbf{v}, \mathbf{s}_1 \rangle - \langle \mathbf{w}, \mathbf{s} \rangle \geq \langle \mathbf{v}, \mathbf{s} \rangle \quad (\text{B.2})$$

Hence,  $\mathbf{s} \in \partial f'(\mathbf{x}_0, \mathbf{w})$ . Conversely, assume  $\mathbf{s} \notin \partial f(\mathbf{x}_0, \mathbf{w})$ . Then, we'll argue that  $\mathbf{s} \notin \partial f'(\mathbf{x}_0, \mathbf{w})$ .

Assume  $f'(\mathbf{x}_0, \mathbf{w}) = c\|\mathbf{w}\|_2^2$  for some scalar  $c = c(\mathbf{w})$ . We can write  $\mathbf{s} = a\mathbf{w} + \mathbf{u}$  where  $\mathbf{u}^T \mathbf{w} = 0$ . Choose  $\mathbf{v} = \epsilon \mathbf{w}$  with  $|\epsilon| < 1$ . We end up with:

$$f'(\mathbf{x}_0, \mathbf{w} + \mathbf{v}) - f'(\mathbf{x}_0, \mathbf{w}) = \epsilon \sup_{\mathbf{s}_1 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s}_1 \rangle = c\epsilon\|\mathbf{w}\|_2^2 \geq \langle \mathbf{s}, \mathbf{v} \rangle = a\epsilon\|\mathbf{w}\|_2^2$$

Consequently, we have  $c\epsilon \geq a\epsilon$  for all  $|\epsilon| < 1$  which implies  $a = c$ . Hence,  $\mathbf{s}$  can be written as  $c\mathbf{w} + \mathbf{u}$ . Now, if  $\mathbf{s} \in \partial f(\mathbf{x}_0)$  then  $\mathbf{s} \in \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)$  as it maximizes  $\langle \mathbf{s}', \mathbf{w} \rangle$  over  $\mathbf{s}' \in \partial f(\mathbf{x}_0)$ . However we assumed  $\mathbf{s} \notin \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)$ . Observe that  $\mathbf{u} = \mathbf{s} - c\mathbf{w}$  and  $\partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}$  lies on  $n - 1$  dimensional subspace  $H$  that is perpendicular to  $\mathbf{w}$ . By assumption  $\mathbf{u} \notin \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}$ . We'll argue that this leads to a contradiction. By making use of convexity of  $\partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}$  and invoking Hyperplane separation theorem (Fact A.1), we can find a direction  $\mathbf{h} \in H$  such that:

$$\langle \mathbf{h}, \mathbf{u} \rangle > \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}} \langle \mathbf{h}, \mathbf{s}' \rangle \quad (\text{B.3})$$

Next, considering  $\epsilon \mathbf{h}$  perturbation, we have:

$$f'(\mathbf{x}_0, \mathbf{w} + \epsilon \mathbf{h}) - f'(\mathbf{x}_0, \mathbf{w}) = \sup_{\mathbf{s}_1 \in \partial f(\mathbf{x}_0)} (\epsilon \langle \mathbf{h}, \mathbf{s}_1 \rangle - \sup_{\mathbf{s}_2 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s}_2 - \mathbf{s}_1 \rangle) \quad (\text{B.4})$$

Denote the  $\mathbf{s}_1$  that establish equality by  $\mathbf{s}_1^*$ .

**Claim:** As  $\epsilon \rightarrow 0$ ,  $\langle \mathbf{s}_1^*, \mathbf{w} \rangle \rightarrow c\|\mathbf{w}\|_2^2$ .

*Proof.* Recall that  $\partial f(\mathbf{x}_0)$  is bounded. Let  $R = \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0)} \|\mathbf{s}'\|_2$ . Choosing  $\mathbf{s}_1 \in \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)$ , we always have:

$$f'(\mathbf{x}_0, \mathbf{w} + \epsilon \mathbf{h}) - f'(\mathbf{x}_0, \mathbf{w}) \geq \epsilon \langle \mathbf{s}_1, \mathbf{h} \rangle \geq -\epsilon R \|\mathbf{h}\|_2$$

On the other hand, for any  $\mathbf{s}_1$  we may write:

$$\epsilon \langle \mathbf{h}, \mathbf{s}_1 \rangle - \sup_{\mathbf{s}_2 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s}_2 - \mathbf{s}_1 \rangle \leq \epsilon R \|\mathbf{h}\|_2 + \langle \mathbf{s}_1, \mathbf{w} \rangle - c\|\mathbf{w}\|_2^2$$

Hence, for  $\mathbf{s}_1^*$ , we obtain:

$$\epsilon R \|\mathbf{h}\|_2 + \langle \mathbf{s}_1^*, \mathbf{w} \rangle - c \|\mathbf{w}\|_2^2 \geq -\epsilon R \|\mathbf{h}\|_2 \implies \langle \mathbf{s}_1^*, \mathbf{w} \rangle \geq c \|\mathbf{w}\|_2^2 - 2\epsilon R \|\mathbf{h}\|_2$$

Letting  $\epsilon \rightarrow 0$ , we obtain the desired result.  $\square$

**Claim:** Given  $\partial f(\mathbf{x}_0)$ , for any  $\epsilon' > 0$  there exists a  $\delta > 0$  such that for all  $\mathbf{s}_1 \in \partial f(\mathbf{x}_0)$  satisfying  $\langle \mathbf{s}_1, \mathbf{w} \rangle > c \|\mathbf{w}\|_2^2 - \delta$  we have  $\text{dist}(\mathbf{s}_1, \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)) < \epsilon'$ .

*Proof.* Assume for some  $\epsilon' > 0$ , claim is false. Then, we can construct a sequence  $\mathbf{s}(i)$  such that  $\text{dist}(\mathbf{s}(i), \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)) \geq \epsilon'$  but  $\langle \mathbf{s}(i), \mathbf{w} \rangle \rightarrow c \|\mathbf{w}\|_2^2$ . From the well-known Bolzano-Weierstrass Theorem and the compactness of  $\partial f(\mathbf{x}_0) \subseteq \mathbb{R}^n$ ,  $\mathbf{s}(i)$  will have a convergent subsequence whose limit  $\mathbf{s}(\infty)$  will be inside  $\partial f(\mathbf{x}_0)$  and will satisfy  $\langle \mathbf{s}(\infty), \mathbf{w} \rangle = c \|\mathbf{w}\|_2^2 = f'(\mathbf{x}_0, \mathbf{w})$ . On the other hand,  $\text{dist}(\mathbf{s}(\infty), \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)) \geq \epsilon' \implies \mathbf{s}(\infty) \notin \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)$  which is a contradiction.  $\square$

Going back to what we have, using the first claim, as  $\epsilon \mathbf{h} \rightarrow 0$ ,  $\langle \mathbf{s}_1^*, \mathbf{w} \rangle \rightarrow c \|\mathbf{w}\|_2^2$ . Using the second claim, this implies for some  $\delta$  which approaches to 0 as  $\epsilon \rightarrow 0$ , we have:

$$\sup_{\mathbf{s}_1 \in \partial f(\mathbf{x}_0)} (\epsilon \langle \mathbf{h}, \mathbf{s}_1 \rangle) - \sup_{\mathbf{s}_2 \in \partial f(\mathbf{x}_0)} \langle \mathbf{w}, \mathbf{s}_2 - \mathbf{s}_1 \rangle \leq \epsilon (\delta \|\mathbf{h}\|_2) + \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}} \langle \mathbf{s}', \mathbf{h} \rangle$$

Finally, based on (B.3), whenever  $\epsilon$  is chosen to ensure  $\delta \|\mathbf{h}\|_2 < \langle \mathbf{h}, \mathbf{u} \rangle - \sup_{\mathbf{s}' \in \partial f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) - c\mathbf{w}} \langle \mathbf{s}', \mathbf{h} \rangle$  we have,

$$f'(\mathbf{x}_0, \mathbf{w} + \epsilon \mathbf{h}) - f'(\mathbf{x}_0, \mathbf{w}) < \epsilon \langle \mathbf{h}, \mathbf{u} \rangle,$$

which contradicts with the initial assumption that  $\mathbf{s}$  is a subgradient of  $f'(\mathbf{x}_0, \cdot)$  at  $\mathbf{w}$ , since,

$$f'(\mathbf{x}_0, \mathbf{w} + \epsilon \mathbf{h}) - f'(\mathbf{x}_0, \mathbf{w}) \geq \langle \mathbf{s}, \epsilon \mathbf{h} \rangle = \epsilon \langle \mathbf{u}, \mathbf{h} \rangle.$$

$\square$

**Lemma B.1.**  $\hat{f}_{\mathbf{x}_0}(\mathbf{x})$  is a convex function of  $\mathbf{x}$ .

*Proof.* To show convexity, we need to argue that the function  $f'(\mathbf{x}_0, \mathbf{w})$  is a convex function of  $\mathbf{w} = \mathbf{x} - \mathbf{x}_0$ . Observe that  $g(\mathbf{w}) = f(\mathbf{x}_0 + \mathbf{w}) - f(\mathbf{x}_0)$  is a convex function of  $\mathbf{w}$  and behaves same as the directional derivative  $f'(\mathbf{x}_0, \mathbf{w})$  for sufficiently small  $\mathbf{w}$ . More rigorously, from (4.11), for any  $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^n$  and  $\delta > 0$  there exists  $\epsilon > 0$  such that, we have:

$$g(\epsilon \mathbf{w}_1) \leq f'(\mathbf{x}_0, \epsilon \mathbf{w}_1) + \delta \epsilon, \quad g(\epsilon \mathbf{w}_2) \leq f'(\mathbf{x}_0, \epsilon \mathbf{w}_2) + \delta \epsilon$$

Hence, for any  $0 \leq c \leq 1$ :

$$\begin{aligned} f'(\mathbf{x}_0, \epsilon(c\mathbf{w}_1 + (1-c)\mathbf{w}_2)) &\leq g(\epsilon(c\mathbf{w}_1 + (1-c)\mathbf{w}_2)) \\ &\leq cg(\epsilon\mathbf{w}_1) + (1-c)g(\epsilon\mathbf{w}_2) \\ &\leq cf'(\mathbf{x}_0, \epsilon\mathbf{w}_1) + (1-c)f'(\mathbf{x}_0, \epsilon\mathbf{w}_2) + \epsilon\delta \end{aligned}$$

Making use of the fact that  $f'(\mathbf{x}_0, \epsilon \mathbf{s}) = \epsilon f'(\mathbf{x}_0, \mathbf{s})$  for any direction  $\mathbf{s}$ , we obtain:

$$f'(\mathbf{x}_0, c\mathbf{w}_1 + (1-c)\mathbf{w}_2) \leq cf'(\mathbf{x}_0, \mathbf{w}_1) + (1-c)f'(\mathbf{x}_0, \mathbf{w}_2) + \delta$$

Letting  $\delta \rightarrow 0$ , we may conclude with the convexity of  $f'(\mathbf{x}_0, \cdot)$  and problem (4.16).  $\square$

## C Swapping the minimization over $\tau$ and the expectation

**Lemma C.1** ([82, 83]). Assume  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_n)$  and let  $h(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  be an  $L$ -Lipschitz function. Then, we have:

$$\text{Var}(h(\mathbf{g})) \leq L^2$$

We next show a closely related result.

**Lemma C.2.** Assume  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_n)$  and let  $h(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  be an  $L$ -Lipschitz function. Then, we have:

$$|h(\mathbf{g}) - \mathbb{E}[h(\mathbf{g})]| \leq \sqrt{2\pi}L$$

*Proof.* From Lipschitzness of  $h(\cdot)$ , letting  $\mathbf{a} = h(\mathbf{g}) - \mathbb{E}[h(\mathbf{g})]$  and invoking Lemma A.4 for all  $t \geq 0$ , we have:

$$\mathbb{P}(|\mathbf{a} - \mathbb{E}[\mathbf{a}]| \geq t) = \mathbb{P}(|\mathbf{a}| \geq t) \leq 2 \exp\left(-\frac{t^2}{2L^2}\right)$$

Denote the probability density function of  $|\mathbf{a}|$  by  $p(\cdot)$  and let  $Q(u) = \mathbb{P}(|\mathbf{a}| \geq u)$ . We may write:

$$\mathbb{E}[|\mathbf{a}|] = \int_0^\infty up(u)du = \int_0^\infty u dQ(u) = [uQ(u)]_\infty^0 + \int_0^\infty Q(u)du$$

Using  $Q(u) \leq 2 \exp(-\frac{u^2}{2L^2})$  for  $u \geq 0$ , we have:

$$[uQ(u)]_\infty^0 = [2u \exp(-\frac{u^2}{2L^2})]_\infty^0 = 0$$

Next,

$$\int_0^\infty Q(u)du \leq \int_0^\infty 2 \exp(-\frac{u^2}{2L^2})du = \sqrt{2\pi}L$$

□

**Lemma C.3.** Suppose Assumption 6.1 holds. Recall that  $\tau(\mathbf{v}) = \arg \min_{\tau \geq 0} \text{dist}(\mathbf{v}, \tau \partial f(\mathbf{x}_0))$ . Then, for all  $\mathbf{v}_1, \mathbf{v}_2$ ,

$$|\tau(\mathbf{v}_1) - \tau(\mathbf{v}_2)| \leq \frac{\|\mathbf{v}_1 - \mathbf{v}_2\|_2}{\|\mathbf{e}\|_2} \quad (\text{C.1})$$

Hence,  $\tau(\mathbf{v})$  is  $\|\mathbf{e}\|_2^{-1}$ -Lipschitz function of  $\mathbf{v}$ .

*Proof.* Let  $\mathbf{a}_i = \text{Proj}(\mathbf{v}_i, \text{cone}(\mathcal{C}))$  for  $1 \leq i \leq 2$ . Using Lemma A.2, we have  $\|\mathbf{a}_1 - \mathbf{a}_2\|_2 \leq \|\mathbf{v}_1 - \mathbf{v}_2\|_2$  as  $\partial f(\mathbf{x}_0)$  is convex. Now, we will further lower bound  $\|\mathbf{v}_1 - \mathbf{v}_2\|_2$  as follows:

$$\|\text{Proj}(\mathbf{a}_1 - \mathbf{a}_2, T)\|_2 \leq \|\mathbf{a}_1 - \mathbf{a}_2\|_2 \leq \|\mathbf{v}_1 - \mathbf{v}_2\|_2$$

Now, observe that  $\|\mathcal{P}_T(\mathbf{a}_1 - \mathbf{a}_2)\|_2 = \|\tau(\mathbf{v}_1)\mathbf{e} - \tau(\mathbf{v}_2)\mathbf{e}\|_2$ . Hence, we may conclude with (C.1). □

**Lemma C.4.** Let  $\mathcal{C}$  be a convex and closed set. Define the set of  $\tau$  that minimizes  $\text{dist}(\mathbf{v}, \tau\mathcal{C})$ ,

$$\mathbf{T}(\mathbf{v}) = \{\tau \geq 0 \mid \arg \min_{\tau \geq 0} \text{dist}(\mathbf{v}, \tau\mathcal{C})\}$$

and let  $\tau(\mathbf{v}) = \inf_{\tau \in \mathbf{T}(\mathbf{v})} \tau$ .  $\tau(\mathbf{v})$  is uniquely determined, given  $\mathcal{C}$  and  $\mathbf{v}$ . Further, assume  $\tau(\mathbf{v})$  is an  $L$  Lipschitz function of  $\mathbf{v}$  and let  $R := R(\mathcal{C}) = \max_{\mathbf{u} \in \mathcal{C}} \|\mathbf{u}\|_2$ . Then,

$$\min_{\tau \geq 0} \mathbb{E}[\text{dist}(\mathbf{g}, \tau\mathcal{C})^2] \leq \mathbf{D}(\text{cone}(\mathcal{C})) + 2\pi(R^2L^2 + RL\sqrt{\mathbf{D}(\text{cone}(\mathcal{C}))}) + 1$$

*Proof.* Let  $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I})$  and let  $\tau^* = \mathbb{E}[\tau(\mathbf{g})]$ . Now, from triangle inequality:

$$|\tau(\mathbf{v}) - \tau^*| \leq t \implies \text{dist}(\mathbf{v}, \tau^*\mathcal{C}) \leq \text{dist}(\mathbf{v}, \tau(\mathbf{v})\mathcal{C}) + Rt$$

Consequently,

$$\mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C})] \leq \min_{\tau \geq 0} \mathbb{E}[\text{dist}(\mathbf{g}, \tau\mathcal{C})] \leq \mathbb{E}[\text{dist}(\mathbf{g}, \tau^*\mathcal{C})] \leq \mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C}) + R|\tau(\mathbf{g}) - \tau^*|]$$

This gives:

$$\mathbb{E}[\text{dist}(\mathbf{g}, \tau^*\mathcal{C})] - \mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C})] \leq R\mathbb{E}[|\tau(\mathbf{g}) - \tau^*|]$$

Observing  $\mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C})] = \mathbb{E}[\text{dist}(\mathbf{g}, \text{cone}(\mathcal{C}))] \leq \sqrt{\mathbf{D}(\text{cone}(\mathcal{C}))}$ , and using Lemma C.2 we find:

$$\mathbb{E}[\text{dist}(\mathbf{g}, \tau^*\mathcal{C})] - \sqrt{\mathbf{D}(\text{cone}(\mathcal{C}))} \leq \sqrt{2\pi}RL$$

This yields:

$$\mathbb{E}[\text{dist}(\mathbf{g}, \tau^*\mathcal{C})]^2 - \mathbf{D}(\text{cone}(\mathcal{C})) \leq \sqrt{2\pi}RL(2\sqrt{\mathbf{D}(\text{cone}(\mathcal{C}))} + \sqrt{2\pi}RL)$$

Using Lemma C.1 we have  $\mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C})^2] \geq \mathbb{E}[\text{dist}(\mathbf{g}, \tau(\mathbf{g})\mathcal{C})]^2 - 1$ , which gives:

$$\min_{\tau \geq 0} \mathbb{E}[\text{dist}(\mathbf{g}, \tau\mathcal{C})^2] - \mathbf{D}(\text{cone}(\mathcal{C})) \leq 2\pi R^2L^2 + 2\sqrt{2\pi}RL\sqrt{\mathbf{D}(\text{cone}(\mathcal{C}))} + 1$$

□

## D Intersection of a cone and a subspace

### D.1 Intersections of randomly oriented cones

Based on Kinematic formula (Theorem 7.1), one may find the following result on the intersection of the two cones. We first consider the scenario in which one of the cones is a subspace.

**Proposition D.1** (Intersection with a subspace). *Let  $A$  be a closed and convex cone and let  $B$  be a linear subspace. Denote  $\delta(A) + \delta(B) - n$  by  $\delta(A, B)$ . Assume the unitary  $\mathbf{U}$  is generated uniformly at random. Given  $\epsilon > 0$ , we have the following:*

- If  $\delta(A) + \delta(B) + \epsilon\sqrt{n} > n$ ,

$$\mathbb{P}(\delta(A \cap \mathbf{U}B) \geq \delta(A, B) + \epsilon\sqrt{n}) \leq 8 \exp\left(-\frac{\epsilon^2}{64}\right)$$

- $\mathbb{P}(\delta(A \cap \mathbf{U}B) \leq \delta(A, B) - \epsilon\sqrt{n}) \leq 8 \exp\left(-\frac{\epsilon^2}{64}\right)$ .

*Proof.* Denote  $A \cap \mathbf{U}B$  by  $C$ . Let  $H$  be a subspace with dimension  $n - d$  chosen uniformly at random independent of  $\mathbf{U}$ . Observe that,  $\mathbf{U}B \cap H$  is a  $\delta(B) - d$  dimensional random subspace for  $d < \delta(B)$ . Hence, using Theorem 7.1 with  $A$  and  $\mathbf{U}B \cap H$  yields:

$$\delta(A) + \delta(B) - d \leq n - t\sqrt{n} \implies \mathbb{P}(A \cap \mathbf{U}B \cap H = \{0\}) \geq 1 - 4 \exp\left(-\frac{t^2}{16}\right) \quad (\text{D.1})$$

$$\delta(A) + \delta(B) - d \geq n + t\sqrt{n} \implies \mathbb{P}(A \cap \mathbf{U}B \cap H = \{0\}) \leq 4 \exp\left(-\frac{t^2}{16}\right) \quad (\text{D.2})$$

Observe that (D.1) is true even when  $d \geq \delta(B)$  since if  $d \geq \delta(B)$ ,  $\mathbf{U}B \cap H = \{0\}$  with probability 1.

**Proving the first statement:** Let  $\gamma = \delta(A) + \delta(B) - n$ ,  $\gamma_\epsilon = \gamma + \epsilon\sqrt{n}$  and  $\gamma_{\epsilon/2} = \gamma + \frac{\epsilon}{2}\sqrt{n}$ . We assume  $\gamma_\epsilon > 0$ . Observing  $A \cap \mathbf{U}B \cap H = C \cap H$ , we may write:

$$\mathbb{P}(C \cap H = \{0\}) \leq \mathbb{P}(C \cap H = \{0\} | \delta(C) \geq \gamma_\epsilon) + \mathbb{P}(\delta(C) \leq \gamma_\epsilon) \quad (\text{D.3})$$

$$\text{and } \mathbb{P}(\delta(C) \leq \gamma_\epsilon) \geq \mathbb{P}(C \cap H = \{0\}) - \mathbb{P}(C \cap H = \{0\} | \delta(C) \geq \gamma_\epsilon) \quad (\text{D.4})$$

If,  $\gamma_\epsilon > n$ ,  $\mathbb{P}(\delta(C) \leq \gamma_\epsilon) = 1$ . Otherwise, choose  $d = \max\{\gamma_{\epsilon/2}, 0\}$ .

**Case 1:** If  $d = 0$ , then,  $\gamma_{\epsilon/2} \leq 0$  and  $H = \mathbb{R}^n$ . This gives,

$$\mathbb{P}(C \cap H = \{0\} | \delta(C) \geq \gamma_\epsilon) = \mathbb{P}(C = \{0\} | \delta(C) \geq \gamma_\epsilon) = 0 \quad (\text{D.5})$$

Also, choosing  $t = \frac{\epsilon}{2}\sqrt{n}$  in (D.1) and using  $\gamma \leq -\frac{\epsilon}{2}\sqrt{n}$ , we obtain:

$$\mathbb{P}(C \cap H = \{0\}) = \mathbb{P}(C = \{0\}) \geq 1 - 4 \exp\left(-\frac{\epsilon^2}{64}\right) \quad (\text{D.6})$$

**Case 2:** Otherwise,  $d = \gamma_{\epsilon/2} > 0$ . Applying Theorem 7.1, we find:

$$\mathbb{P}(C \cap H = \{0\} | \delta(C) \geq \gamma_\epsilon) \leq 4 \exp\left(-\frac{\epsilon^2}{64}\right) \quad (\text{D.7})$$

Next, choosing  $t = \frac{\epsilon}{2}\sqrt{n}$  in (D.1), we obtain:

$$\mathbb{P}(C \cap H = \{0\}) \geq 1 - 4 \exp\left(-\frac{\epsilon^2}{64}\right) \quad (\text{D.8})$$

Overall, combining (D.4), (D.5), (D.6), (D.7) and (D.8), we obtain:

$$\mathbb{P}(\delta(C) \leq \gamma_\epsilon) \geq 1 - 8 \exp\left(-\frac{\epsilon^2}{64}\right)$$

**Proving the second statement:** In the exact same manner, this time, let  $\gamma_{-\epsilon} = \gamma - \epsilon\sqrt{n}$ ,  $\gamma_{-\epsilon/2} = \gamma - \frac{\epsilon}{2}\sqrt{n}$ . If  $\gamma_{-\epsilon} < 0$ ,

$$\mathbb{P}(\delta(C) \leq \gamma_{-\epsilon}) \leq \mathbb{P}(\delta(C) < 0) = 0$$

Otherwise, let  $d = \gamma_{-\epsilon/2}$ , we may write,

$$\mathbb{P}(\delta(C) \geq \gamma_{-\epsilon}) \geq \mathbb{P}(C \cap H \neq \{0\}) - \mathbb{P}(C \cap H \neq \{0\} | \delta(C) \leq \gamma_{-\epsilon}) \quad (\text{D.9})$$

in an identical way to (D.4). Repeating the previous argument and using (D.2), we may first obtain,

$$\mathbb{P}(C \cap H \neq \{0\}) \geq 1 - 4 \exp\left(-\frac{\epsilon^2}{64}\right)$$

and using Theorem 7.1,

$$\mathbb{P}(C \cap H \neq \{0\} | \delta(C) \leq \gamma_{-\epsilon}) \leq 4 \exp\left(-\frac{\epsilon^2}{64}\right)$$

Combining these, gives the desired result.

$$\mathbb{P}(\delta(C) \geq \gamma_{-\epsilon}) \geq 1 - 8 \exp\left(-\frac{\epsilon^2}{64}\right)$$

□

## E Proof of Theorem 3.1 – Lower bound

**Theorem E.1.** *Let  $C$  be a closed and convex set,  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$  and let  $\mathbf{x}^*(\sigma\mathbf{v}) = \arg \min_{\mathbf{x} \in C} \|\mathbf{x}_0 + \sigma\mathbf{v} - \mathbf{x}\|_2$ . Then, we have,*

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{x}^*(\sigma\mathbf{v}) - \mathbf{x}_0\|_2^2]}{\sigma^2} = \mathbf{D}(T_C(\mathbf{x}_0)^*)$$

*Proof.* Let  $1 \geq \alpha, \epsilon > 0$  be numbers to be determined. Denote probability density function of a  $\mathcal{N}(0, c\mathbf{I})$  distributed vector by  $p_c(\cdot)$ . From Lemma F.1, the expected error  $\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]$  is simply,

$$\int_{\mathbf{v} \in \mathbb{R}^n} \|\text{Proj}(\sigma\mathbf{v}, F_C(\mathbf{x}_0))\|_2^2 p_1(\mathbf{v}) d\mathbf{v}$$

Let  $S_\alpha$  be the set satisfying:

$$S_\alpha = \{\mathbf{u} \in \mathbb{R}^n \mid \frac{\|\text{Proj}(\mathbf{u}, T_C(\mathbf{x}_0))\|_2}{\|\mathbf{u}\|_2} \geq \alpha\}.$$

Let  $\bar{S}_\alpha = \mathbb{R}^n - S_\alpha$ . Using Proposition F.1, given  $\epsilon > 0$ , choose  $\epsilon_0 > 0$  such that, for all  $\|\mathbf{u}\|_2 \leq \epsilon_0$  and  $\mathbf{u} \in S_\alpha$ , we have,

$$\|\text{Proj}(\mathbf{u}, F_C(\mathbf{x}_0))\|_2 \geq (1 - \epsilon) \|\text{Proj}(\mathbf{u}, T_C(\mathbf{x}_0))\|_2. \quad (\text{E.1})$$

Now, let  $\mathbf{z} = \sigma\mathbf{v}$ . Split the error into three groups, namely:

- $F_1 = \int_{\|\mathbf{z}\|_2 \leq \epsilon_0, \mathbf{z} \in S_\alpha} \|\text{Proj}(\mathbf{z}, F_C(\mathbf{x}_0))\|_2^2 p_\sigma(\mathbf{z}) d\mathbf{z}, \quad T_1 = \int_{\|\mathbf{v}\|_2 \leq \frac{\epsilon_0}{\sigma}, \mathbf{v} \in S_\alpha} \|\text{Proj}(\mathbf{v}, T_C(\mathbf{x}_0))\|_2^2 p_1(\mathbf{v}) d\mathbf{v}.$
- $F_2 = \int_{\|\mathbf{z}\|_2 \geq \epsilon_0, \mathbf{z} \in S_\alpha} \|\text{Proj}(\mathbf{z}, F_C(\mathbf{x}_0))\|_2^2 p_\sigma(\mathbf{z}) d\mathbf{z}, \quad T_2 = \int_{\|\mathbf{v}\|_2 \geq \frac{\epsilon_0}{\sigma}, \mathbf{v} \in S_\alpha} \|\text{Proj}(\mathbf{v}, T_C(\mathbf{x}_0))\|_2^2 p_1(\mathbf{v}) d\mathbf{v}.$
- $F_3 = \int_{\mathbf{z} \in \bar{S}_\alpha} \|\text{Proj}(\mathbf{z}, F_C(\mathbf{x}_0))\|_2^2 p_\sigma(\mathbf{z}) d\mathbf{z}, \quad T_3 = \int_{\mathbf{v} \in \bar{S}_\alpha} \|\text{Proj}(\mathbf{v}, T_C(\mathbf{x}_0))\|_2^2 p_1(\mathbf{g}) d\mathbf{v}.$

The rest of the argument will be very similar to the proof of Proposition 5.2. We know the following from Proposition 4.1:

$$\begin{aligned} T_1 + T_2 + T_3 &= \mathbf{D}(T_C(\mathbf{x}_0)^*) \\ F_1 + F_2 + F_3 &= \mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2] \leq \sigma^2(T_1 + T_2 + T_3) \end{aligned}$$

To proceed, we will argue that, the contributions of the second and third terms are small for sufficiently small  $\sigma, \alpha, \epsilon > 0$ . Observe that:

$$T_3 \leq \int_{\mathbf{v} \in \bar{S}_\alpha} \alpha^2 \|\mathbf{v}\|_2^2 p_1(\mathbf{v}) d\mathbf{v} \leq \alpha^2 n$$

For  $T_2$ , we have:

$$T_2 \leq \int_{\|\mathbf{v}\|_2 \geq \frac{\epsilon_0}{\sigma}} \|\mathbf{v}\|_2^2 p_1(\mathbf{g}) d\mathbf{v} = C\left(\frac{\epsilon_0}{\sigma}\right)$$

Since  $\|\mathbf{g}\|_2$  has finite second moment, fixing  $\epsilon_0 > 0$  and letting  $\sigma \rightarrow 0$ , we have  $C(\frac{\epsilon_0}{\sigma}) \rightarrow 0$ . For  $T_1$ , from (E.1), we have:

$$F_1 \geq (1 - \epsilon)^2 \sigma^2 T_1$$

Overall, we found:

$$\frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2} \geq \frac{F_1}{\sigma^2} \geq (1 - \epsilon)^2 \frac{T_1}{T_1 + T_2 + T_3} \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$$

Writing  $T_1 = \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - T_2 - T_3 \geq \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - \alpha^2 n - C(\frac{\epsilon_0}{\sigma})$ , we have:

$$\frac{T_1}{T_1 + T_2 + T_3} \geq \frac{\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - \alpha^2 n - C(\frac{\epsilon_0}{\sigma})}{\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)}$$

Letting  $\sigma \rightarrow 0$  for fixed  $\alpha, \epsilon_0, \epsilon$ , we obtain:

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2 \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)} \geq (1 - \epsilon)^2 \frac{\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - \alpha^2 n}{\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)}$$

Since,  $\alpha, \epsilon$  can be made arbitrarily small, we obtain  $\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{x}^* - \mathbf{x}_0\|_2^2]}{\sigma^2 \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)} = 1$ . □

The next result shows that, as  $\sigma \rightarrow 0$ , we can exactly predict the cost of the constrained problem.

**Proposition E.1.** *Consider the setup in Theorem E.1. Let  $\mathbf{w}^*(\sigma \mathbf{v}) = \mathbf{x}^*(\sigma \mathbf{v}) - \mathbf{x}_0$ . Then,*

$$\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\sigma \mathbf{v} - \mathbf{w}^*(\sigma \mathbf{v})\|_2^2]}{\sigma^2} = \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0))$$

*Proof.* Let  $\mathbf{w}^* = \mathbf{w}^*(\sigma \mathbf{v})$  and  $\mathbf{z} = \sigma \mathbf{v}$ .  $\mathbf{z} - \mathbf{w}^*$  satisfies two conditions.

- From Lemma F.1,  $\|\mathbf{z} - \mathbf{w}^*\|_2 = \text{dist}(\mathbf{z}, F_{\mathcal{C}}(\mathbf{x}_0)) \geq \text{dist}(\mathbf{z}, T_{\mathcal{C}}(\mathbf{x}_0))$ .
- Using Lemma F.4,  $\|\mathbf{z} - \mathbf{w}^*\|_2^2 + \|\mathbf{w}^*\|_2^2 \leq \|\mathbf{z}\|_2^2$ .

Consequently, when  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{I})$ , we find:

$$n\sigma^2 = \mathbb{E}[\|\mathbf{z}\|_2^2] \geq \mathbb{E}[\|\mathbf{z} - \mathbf{w}^*\|_2^2] + \mathbb{E}[\|\mathbf{w}^*\|_2^2] \geq \sigma^2 \mathbb{E}[\text{Proj}(\mathbf{v}, T_{\mathcal{C}}(\mathbf{x}_0)^*)^2] + \mathbb{E}[\|\mathbf{w}^*\|_2^2]$$

Normalizing both sides by  $\sigma^2$  and subtracting  $\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)) = \mathbb{E}[\text{Proj}(\mathbf{v}, T_{\mathcal{C}}(\mathbf{x}_0)^*)^2]$  and  $\mathbb{E}[\|\mathbf{w}^*\|_2^2]$  we find:

$$\mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*) - \frac{\mathbb{E}[\|\mathbf{w}^*\|_2^2]}{\sigma^2} \geq \frac{\mathbb{E}[\|\mathbf{z} - \mathbf{w}^*\|_2^2]}{\sigma^2} - \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)) \geq 0$$

where we used Lemma A.2. Now, letting  $\sigma \rightarrow 0$  and using the fact that  $\lim_{\sigma \rightarrow 0} \frac{\mathbb{E}[\|\mathbf{w}^*\|_2^2]}{\sigma^2} = \mathbf{D}(T_{\mathcal{C}}(\mathbf{x}_0)^*)$ , we find the desired result. □

## F Approximation results on convex cones

### F.1 Standard observations

**Remark:** Throughout the section,  $\mathcal{C}$  will be a nonempty, closed and convex set in  $\mathbb{R}^n$ .

**Property F.1.** *Let  $\mathbf{x}_0 \in \mathcal{C}$  and  $\mathbf{y} = \mathbf{x}_0 + \mathbf{z} \in \mathbb{R}^n$ . From Lemma A.1, recall that,  $\text{Proj}(\mathbf{y}, \mathcal{C})$  is the unique vector that is equal to  $\arg \min_{\mathbf{u} \in \mathcal{C}} \|\mathbf{y} - \mathbf{u}\|_2$ . By definition of feasible set  $F_{\mathcal{C}}(\mathbf{x}_0)$ , we also have,  $\text{Proj}(\mathbf{y}, \mathcal{C}) = \text{Proj}(\mathbf{z}, F_{\mathcal{C}}(\mathbf{x}_0))$ .*

**Lemma F.1.** *For all  $\mathbf{z} \in \mathbb{R}^n$  and  $\mathbf{x}_0 \in \mathcal{C}$ , we have:*

$$\|\text{Proj}(\mathbf{z}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2 \leq \|\text{Proj}(\mathbf{z}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2$$

*Proof.* Setting  $f(\cdot) = 0$  and  $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$  in Lemma 4.2, we have:

$$\|\text{Proj}(\mathbf{z}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2 = \|\mathbf{x}^* - \mathbf{x}_0\|_2 \leq \text{dist}(\mathbf{z}, T_{\mathcal{C}}(\mathbf{x}_0)^*) = \|\text{Proj}(\mathbf{z}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2$$

□

The following lemma shows that projection onto the feasible cone is arbitrarily close to the projection onto the tangent cone as we scale down the vector. This is due to Proposition 5.3.5 of Chapter III of [88].

**Lemma F.2.** *Assume  $\mathbf{x}_0 \in \mathcal{C}$ . Then, for any  $\mathbf{z} \in \mathcal{C}$ ,*

$$\lim_{\epsilon \rightarrow 0} \frac{\text{Proj}(\epsilon \mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))}{\epsilon} \rightarrow \text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))$$

Hence,

- If  $\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0)) = 0$ , using Lemma F.1,  $\text{Proj}(\mathbf{z}, F_{\mathcal{C}}(\mathbf{x}_0)) = 0$ .
- If  $\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0)) \neq 0$ ,

$$\lim_{\epsilon \rightarrow 0} \frac{\|\text{Proj}(\epsilon \mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2}{\|\text{Proj}(\epsilon \mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2} = 1.$$

## F.2 Uniform approximation to the tangent cone

**Proposition F.1.** *Let  $\mathcal{C}$  be a closed and convex set including  $\mathbf{x}_0$ . Denote the unit  $\ell_2$ -sphere in  $\mathbb{R}^n$  by  $\mathcal{S}^{n-1}$  and let  $1 \geq \alpha > 0$  be arbitrary. Given  $\alpha, \epsilon > 0$ , there exists an  $\epsilon_0 > 0$  such that, for all  $\mathbf{w} \in \mathcal{S}^{n-1}$ ,  $\|\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2 \geq \alpha$  and for all  $0 < t \leq \epsilon_0$ , we have:*

$$\frac{\|\text{Proj}(t\mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2}{t\|\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2} \geq 1 - \epsilon \quad (\text{F.1})$$

In particular, setting  $\alpha = 1$ , given  $\epsilon > 0$ , there exists  $\epsilon_0 > 0$  such that, for all  $t \leq \epsilon_0$  and all  $\mathbf{w} \in T_{\mathcal{C}}(\mathbf{x}_0) \cap \mathcal{S}^{n-1}$ ,  $\|\text{Proj}(t\mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2 \geq (1 - \epsilon)t$ .

**Remark:** Note that, statements of Propositions F.1 and 5.1 are quite similar.

*Proof.* Given  $\alpha > 0$ , consider the following set:

$$S = \{\mathbf{w} \in \mathcal{S}^{n-1} \mid \|\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2 \geq \alpha\}$$

This set is closed and bounded and hence compact. Define the following function on this set:

$$c(\mathbf{w}) = \max\{c > 0 \mid \frac{\|\text{Proj}(c\mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2}{\|\text{Proj}(c\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2} \geq 1 - \epsilon\}$$

$c(\mathbf{w})$  is strictly positive due to Lemma F.2 and it can be as high as infinity. Furthermore, from Lemma F.3, we know that whenever  $c < c(\mathbf{w})$ :

$$\frac{\|\text{Proj}(c\mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2}{\|\text{Proj}(c\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2} \geq 1 - \epsilon$$

as well. Let  $s(\mathbf{w}) = \min\{1, c(\mathbf{w})\}$ . If  $s(\mathbf{w})$  is continuous, since  $\mathcal{S}^{n-1}$  is compact  $s(\mathbf{w})$  will attain its minimum which implies  $c(\mathbf{w}) \geq s(\mathbf{w}) \geq \epsilon_0 > 0$  for some  $\epsilon_0$ . Again, this also implies, for all  $\mathbf{w} \in \mathcal{S}^{n-1}$ , and  $0 < t \leq \epsilon_0$ ,

$$\frac{\|\text{Proj}(t\mathbf{w}, F_{\mathcal{C}}(\mathbf{x}_0))\|_2}{t\|\text{Proj}(\mathbf{w}, T_{\mathcal{C}}(\mathbf{x}_0))\|_2} \geq 1 - \epsilon$$

To end the proof, we will show continuity of  $s(\mathbf{w})$ .

**Claim:**  $s(\mathbf{w})$  is continuous.

*Proof.* We will show that  $\lim_{\mathbf{w}_2 \rightarrow \mathbf{w}_1} s(\mathbf{w}_2) = s(\mathbf{w}_1)$ . To do this, we will make use of the continuity of the functions  $\|\text{Proj}(c_1 \mathbf{w}, F_C(\mathbf{x}_0))\|_2$ ,  $\|\text{Proj}(c_1 \mathbf{w}, T_C(\mathbf{x}_0))\|_2$  and  $\frac{\|\text{Proj}(c_1 \mathbf{w}, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c_1 \mathbf{w}, T_C(\mathbf{x}_0))\|_2}$  when the denominator is nonzero. Given  $\mathbf{w}_1$ , let  $c_1 = \min\{2, c(\mathbf{w}_1)\}$ .

**Case 1:** If  $\frac{\|\text{Proj}(c_1 \mathbf{w}_1, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c_1 \mathbf{w}_1, T_C(\mathbf{x}_0))\|_2} > 1 - \epsilon$ , then  $c(\mathbf{w}_1) > 2$  and for all  $\mathbf{w}_2$  sufficiently close to  $\mathbf{w}_1$ ,  $\frac{\|\text{Proj}(c_1 \mathbf{w}_2, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c_1 \mathbf{w}_2, T_C(\mathbf{x}_0))\|_2}$  is more than  $1 - \epsilon$  and hence  $c(\mathbf{w}_2) \geq 2 > 1$ . Hence,  $s(\mathbf{w}_1) = s(\mathbf{w}_2)$ .

**Case 2:** Now, assume  $\frac{\|\text{Proj}(c_1 \mathbf{w}_1, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c_1 \mathbf{w}_1, T_C(\mathbf{x}_0))\|_2} = 1 - \epsilon$  which implies  $c_1 = c(\mathbf{w}_1)$ . Using the ‘‘strict decrease’’ part of Lemma F.3, for any  $\epsilon' > 0$  and  $c' = c_1 - \epsilon'$ ,  $\frac{\|\text{Proj}(c' \mathbf{w}_1, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c' \mathbf{w}_1, T_C(\mathbf{x}_0))\|_2} > 1 - \epsilon$ . Then, for  $\mathbf{w}_2$  sufficiently close to  $\mathbf{w}_1$ ,  $\frac{\|\text{Proj}(c' \mathbf{w}_2, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c' \mathbf{w}_2, T_C(\mathbf{x}_0))\|_2} > 1 - \epsilon$  which implies  $c(\mathbf{w}_2) \geq c'$ . Hence,  $c(\mathbf{w}_2) \geq c_1 - \epsilon'$  for arbitrarily small  $\epsilon' > 0$ . Conversely, for any  $\epsilon' > 0$  and  $c' = c_1 + \epsilon'$ ,  $\frac{\|\text{Proj}(c' \mathbf{w}_1, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c' \mathbf{w}_1, T_C(\mathbf{x}_0))\|_2} < 1 - \epsilon$ . Then, for  $\mathbf{w}_2$  sufficiently close to  $\mathbf{w}_1$ ,  $\frac{\|\text{Proj}(c' \mathbf{w}_2, F_C(\mathbf{x}_0))\|_2}{\|\text{Proj}(c' \mathbf{w}_2, T_C(\mathbf{x}_0))\|_2} < 1 - \epsilon$  which implies  $c(\mathbf{w}_2) \leq c'$ . Hence,  $c(\mathbf{w}_2) \leq c_1 + \epsilon'$  for arbitrarily small  $\epsilon' > 0$ . Combining these, we obtain  $c(\mathbf{w}_2) \rightarrow c(\mathbf{w}_1)$  as  $\mathbf{w}_2 \rightarrow \mathbf{w}_1$ . This also implies  $s(\mathbf{w}_2) \rightarrow s(\mathbf{w}_1)$ .  $\square$

This finishes the proof of the main statement (F.1). For the  $\alpha = 1$  case, observe that,  $\|\mathbf{w}\|_2 = 1$  and  $\|\text{Proj}(\mathbf{w}, T_C(\mathbf{x}_0))\|_2 = 1$  implies  $\mathbf{w} \in T_C(\mathbf{x}_0)$ .  $\square$

**Lemma F.3.** Let  $\mathbf{x}_0 \in \mathbb{R}^n$  and let  $\mathbf{w}$  have unit  $\ell_2$ -norm and set  $l_T = \|\text{Proj}(\mathbf{w}, T_C(\mathbf{x}_0))\|_2$ . Define the function,

$$g(t) = \begin{cases} \frac{\|\text{Proj}(t\mathbf{w}, F_C(\mathbf{x}_0))\|_2}{t} & \text{for } t > 0 \\ l_T & \text{for } t = 0 \end{cases}$$

Then,  $g(\cdot)$  is continuous and non increasing on  $[0, \infty)$ . Furthermore, it is strictly decreasing on the interval  $[t_0, \infty)$  where  $t_0 = \sup\{t > 0 \mid g(t) = l_T\}$ .

*Proof.* Due to Lemma F.1,  $g(t) \leq l_T$  and from Lemma F.2, the function is continuous at 0. Continuity at  $t \neq 0$  follows from the continuity of the projection (see Fact A.2). Next, if  $g(t) = l_T$ , using the fact that  $F_C(\mathbf{x}_0)$  contains 0, the second statement of Lemma F.4 gives,

$$\text{Proj}(t\mathbf{w}, T_C(\mathbf{x}_0)) = \text{Proj}(t\mathbf{w}, F_C(\mathbf{x}_0)) \in F_C(\mathbf{x}_0).$$

From convexity,  $\text{Proj}(t'\mathbf{w}, T_C(\mathbf{x}_0)) \in F_C(\mathbf{x}_0)$  for all  $0 \leq t' \leq t$ . Hence,  $g(t') = l_T$ . This implies  $g(t) = l_T$  for  $t \leq t_0$ .

Now, assume  $t_1 > t_0$  and  $t_1 > t_2 > 0$  for some  $t_1, t_2 > 0$ . Then,  $g(t_1) < l_T$ , hence, the third statement of Lemma F.4 applies. Setting  $\alpha = \frac{t_2}{t_1}$  in Lemma F.4, we find,

$$\|\text{Proj}(t_1 \mathbf{w}, F_C(\mathbf{x}_0))\|_2 < \frac{\|\text{Proj}(t_2 \mathbf{w}, F_C(\mathbf{x}_0))\|_2}{\frac{t_2}{t_1}},$$

which implies the strict decrease of  $\frac{\|\text{Proj}(t\mathbf{w}, F_C(\mathbf{x}_0))\|_2}{t}$  over  $t \geq t_0$ .  $\square$

For the rest of the discussion, given three points  $A, B, C$  in  $\mathbb{R}^n$ , the angle induced by the lines  $AB$  and  $BC$  will be denoted by  $\hat{A}BC$ .

**Lemma F.4.** Let  $\mathcal{K}$  be a convex and closed set in  $\mathbb{R}^n$  that includes 0. Let  $\mathbf{z} \in \mathbb{R}^n$  and  $0 < \alpha < 1$  be arbitrary, let  $\mathbf{p}_1 = \text{Proj}(\mathbf{z}, \mathcal{K})$ ,  $\mathbf{p}_2 = \text{Proj}(\alpha \mathbf{z}, \mathcal{K})$ . Denote the points whose coordinates are determined by 0,  $\mathbf{p}_1, \mathbf{p}_2, \mathbf{z}$  by  $O, P_1, P_2$  and  $Z$  respectively. Then,

- $Z\hat{P}_1O$  is either wide or right angle.
- If  $Z\hat{P}_1O$  is right angle, then  $\mathbf{p}_1 = \frac{\mathbf{p}_2}{\alpha} = \text{Proj}(\mathbf{z}, T_{\mathcal{K}}(0))$ .
- If  $Z\hat{P}_1O$  is wide angle, then  $\|\mathbf{p}_1\|_2 < \frac{\|\mathbf{p}_2\|_2}{\alpha} \leq \|\text{Proj}(\mathbf{z}, T_{\mathcal{K}}(0))\|_2$ .

*Proof. Acute angle:* Assume  $Z\hat{P}_1O$  is acute angle. If  $Z\hat{O}P_1$  is right or wide angle, then 0 is closer to  $\mathbf{z}$  than  $\mathbf{p}_1$  which is a contradiction. If  $Z\hat{O}P_1$  is acute angle, then draw the perpendicular from  $Z$  to the line  $OP_1$ . The intersection is in  $\mathcal{K}$  due to convexity and it is closer to  $\mathbf{z}$  than  $\mathbf{p}_1$ , which again is a contradiction.

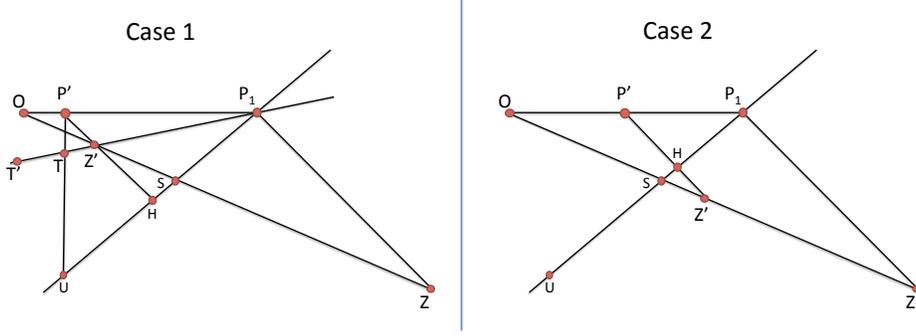


Figure 5: Possible configurations of the points in Lemma F.4.

**Right angle:** Now, assume  $Z\hat{P}_1O$  is right angle. Using Fact A.2, there exists a hyperplane  $H$  that separates  $\mathbf{z}$  and  $\mathcal{K}$  passing through  $P_1$  which is perpendicular to  $\mathbf{z} - \mathbf{p}_1$ . The line  $P_1O$  lies on  $H$ . Consequently, for any  $\alpha \in [0, 1]$ , the closest point to  $\alpha\mathbf{z}$  over  $\mathcal{K}$  is simply  $\alpha\mathbf{p}_1$ . Hence,  $\mathbf{p}_2 = \alpha\mathbf{p}_1$ . Now, let  $\mathbf{q}_1 := \text{Proj}(\mathbf{z}, T_{\mathcal{K}}(0))$ . Then,  $\text{Proj}(\alpha\mathbf{z}, T_{\mathcal{K}}(0)) = \alpha\mathbf{q}_1$ . If  $\mathbf{q}_1 \neq \mathbf{p}_1$  then,  $\|\mathbf{q}_1\|_2 > \|\mathbf{p}_1\|_2$  since  $\|\mathbf{z} - \mathbf{q}_1\|_2 < \|\mathbf{z} - \mathbf{p}_1\|_2$  and:

$$\|\mathbf{q}_1\|_2^2 = \|\mathbf{z}\|_2^2 - \|\mathbf{z} - \mathbf{q}_1\|_2^2 > \|\mathbf{z}\|_2^2 - \|\mathbf{z} - \mathbf{p}_1\|_2^2 \geq \|\mathbf{p}_1\|_2^2$$

where the last inequality follows from the fact that  $Z\hat{P}_1O$  is not acute. Then,

$$\lim_{\alpha \rightarrow 0} \frac{\|\text{Proj}(\mathbf{z}, T_{\mathcal{K}}(0))\|_2}{\|\text{Proj}(\mathbf{z}, \mathcal{K})\|_2} = \frac{\|\mathbf{q}_1\|_2}{\|\mathbf{p}_1\|_2} > 1$$

which contradicts with Lemma F.2.

**Wide angle:** Finally, assume  $Z\hat{P}_1O$  is wide angle. We start by reducing the problem to a two dimensional one. Obtain  $\mathcal{K}'$  by projecting the set  $\mathcal{K}$  to the 2D plane induced by the points  $Z, P_1$  and  $O$ . Now, let  $\mathbf{p}'_2 = \text{Proj}(\alpha\mathbf{z}, \mathcal{K}')$ . Due to the projection, we still have:

$$\|\mathbf{z} - \mathbf{p}'_2\|_2 \leq \|\mathbf{z} - \mathbf{p}_2\|_2 \leq \|\mathbf{z} - \alpha\mathbf{p}_1\|_2 \quad (\text{F.2})$$

and  $\|\mathbf{p}'_2\|_2 \leq \|\mathbf{p}_2\|_2$ . Next, we will prove that  $\|\mathbf{p}'_2\|_2 > \|\alpha\mathbf{p}_1\|_2$  to conclude. Figure 5 will help us explain our approach. Let the line  $UP_1$  be perpendicular to  $ZP_1$ . Assume, it crosses  $ZO$  at  $S$ . Let  $P'Z'$  be parallel to  $P_1Z_1$ . Observe that  $P'$  corresponds to  $\alpha\mathbf{p}_1$ .  $H$  is the intersection of  $P'Z'$  and  $P_1U$ . Denote the point corresponding to  $\mathbf{p}'_2$  by  $P'_2$ . Observe that  $P'_2$  satisfies the following:

- $P_1$  is the closest point to  $Z$  in  $\mathcal{K}$  hence  $P'_2$  lies on the left of  $P_1U$  (same side as  $O$ ).
- $P_2$  is the closest point to  $Z'$ . Hence,  $Z'\hat{P}_2P_1$  is not acute angle. Otherwise, we can draw a perpendicular to  $P_2P_1$  from  $Z'$  and end up with a shorter distance. This would also imply that  $Z'\hat{P}'_2P_1$  is not acute as well. The reason is, due to projection,  $|Z'P'_2| \leq |Z'P_2|$  and  $|P'_2P_1| \leq |P_2P_1|$  hence,

$$|Z'P_1| \geq |Z'P_2|^2 + |P_2P_1|^2 \geq |Z'P'_2|^2 + |P'_2P_1|^2 \quad (\text{F.3})$$

- $P'_2$  has to lie below or on the line  $OP_1$  otherwise, perpendicular to  $OP_1$  from  $Z'$  would yield a shorter distance than  $|P'_2Z'|$ .
- $\mathbf{p}_2 \neq \alpha\mathbf{p}_1$ . To see this, note that  $Z'\hat{P}'O$  is wide angle. Let  $\mathbf{q} \in \mathbb{R}^n$  be the projection of  $\alpha\mathbf{z}$  on the line  $\{c\mathbf{p}_1 | c \in \mathbb{R}\}$  and point  $Q$  denote the vector  $\mathbf{q}$ . If  $Q$  lies between  $O$  and  $P_1$ ,  $\mathbf{q} \in \mathcal{K}$  and  $|QZ'| < |P'Z'|$ . Otherwise,  $P_1$  lies between  $Q$  and  $P'$  hence  $|P_1Z'| < |P'Z'|$  and  $\mathbf{p} \in \mathcal{K}$ . This implies  $P_2, P'_2 \neq P'$ .

Based on these observations, we investigate the problem in two cases illustrated by Figure 5.

**Case 1 ( $S$  lies on  $Z'Z$ ):** Consider the lefthand side of Figure 5. If  $P'_2$  lies on the righthand side of  $P'U$ , this implies  $|P'_2O| > |P'O|$  which is what we wanted.

If  $P'_2$  lies on the region induced by  $OP'TT'$  then  $P_1\hat{P}'_2Z'$  is acute angle as  $P_1\hat{Z}'P'_2 > P_1\hat{Z}'P'$  is wide, which contradicts with (F.3).

If  $P'_2$  lies on the remaining region  $T'TU$ , then  $Z'\hat{P}'_2P_1$  is acute. The reason is,  $P'_2\hat{Z}'P_1$  is wide as follows:

$$P'_2\hat{Z}'P_1 \geq P'_2\hat{T}P_1 \geq U\hat{T}P_1 > U\hat{P}'P_1 = \frac{\pi}{2}$$

**Case 2 ( $S$  lies on  $OZ'$ ):** Consider the righthand side of Figure 5. Due to location restrictions,  $P'_2$  lies on either  $P_1P'H$  triangle or the region induced by  $OP'HU$ . If it lies on  $P_1P'H$  then,  $O\hat{P}'P'_2 \geq O\hat{P}'H$  (thus wide); which implies  $|OP'_2| > |OP'|$  as  $O\hat{P}'P'_2$  is wide angle and  $P' \neq P'_2$ .

If  $P'_2$  lies on  $OP'HU$  then,  $P_1\hat{P}'_2Z' < P_1\hat{H}Z' = \frac{\pi}{2}$  hence  $P_1\hat{P}'_2Z'$  is acute angle which contradicts with (F.3).

In all cases, we end up with  $|OP'_2| > |OP'|$  which implies  $\|\mathbf{p}_2\|_2 \geq \|\mathbf{p}'_2\|_2 > \alpha\|\mathbf{p}_1\|_2$  as desired.

Finally, apply Lemma F.1 on  $\alpha\mathbf{z}$  to upper bound  $\|\mathbf{p}_2\|_2$  by  $\alpha\|\text{Proj}(\mathbf{z}, T_{\mathcal{K}}(0))\|_2$ . □